

Université de Montréal

**Schémas de classification et repérage des documents administratifs
électroniques dans un contexte de gestion décentralisée des
ressources informationnelles**

par
Sabine Mas

École de bibliothéconomie et des sciences de l'information
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en sciences de l'information

Mai 2007

© Sabine Mas, 2007

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée :

Schémas de classification et repérage des documents administratifs électroniques
dans un contexte de gestion décentralisée des ressources informationnelles

présentée par :
Sabine Mas

a été évaluée par un jury composé des personnes suivantes :

Clément Arsenault, président-rapporteur
Louise Gagnon-Arguin, directrice de recherche
Michèle Hudon, codirectrice
James M. Turner, membre du jury
Luis Carlos Lopes, examinateur externe
Jian Yun Nie, représentant du doyen de la FES

Résumé

Les employés d'un organisme utilisent souvent un schéma de classification personnel pour organiser les documents électroniques qui sont sous leur contrôle direct, ce qui suggère la difficulté pour d'autres employés de repérer ces documents et la perte possible de documentation pour l'organisme. Aucune étude empirique n'a été menée à ce jour afin de vérifier dans quelle mesure les schémas de classification personnels permettent, ou même facilitent, le repérage des documents électroniques par des tiers, dans le cadre d'un travail collaboratif par exemple, ou lorsqu'il s'agit de reconstituer un dossier.

Le premier objectif de notre recherche était de décrire les caractéristiques de schémas de classification personnels utilisés pour organiser et classer des documents administratifs électroniques. Le deuxième objectif consistait à vérifier, dans un environnement contrôlé, les différences sur le plan de l'efficacité du repérage de documents électroniques qui sont fonction du schéma de classification utilisé. Nous voulions vérifier s'il était possible de repérer un document avec la même efficacité, quel que soit le schéma de classification utilisé pour ce faire.

Une collecte de données en deux étapes fut réalisée pour atteindre ces objectifs. Nous avons d'abord identifié les caractéristiques structurelles, logiques et sémantiques de 21 schémas de classification utilisés par des employés de l'Université de Montréal pour organiser et classer les documents électroniques qui sont sous leur contrôle direct. Par la suite, nous avons comparé, à partir d'une expérimentation contrôlée, la capacité d'un groupe de 70 répondants à repérer des documents électroniques à l'aide de cinq schémas de classification ayant des caractéristiques structurelles, logiques et sémantiques variées. Trois variables ont été utilisées pour mesurer l'efficacité du repérage : la proportion de documents repérés, le temps moyen requis (en secondes) pour repérer les documents et la proportion de documents repérés dès le premier essai.

Les résultats révèlent plusieurs caractéristiques structurelles, logiques et sémantiques communes à une majorité de schémas de classification personnels : macro-

structure étendue, structure peu profonde, complexe et déséquilibrée, regroupement par thème, ordre alphabétique des classes, etc. Les résultats des tests d'analyse de la variance révèlent des différences significatives sur le plan de l'efficacité du repérage de documents électroniques qui sont fonction des caractéristiques structurelles, logiques et sémantiques du schéma de classification utilisé. Un schéma de classification caractérisé par une macro-structure peu étendue et une logique basée partiellement sur une division par classes d'activités augmente la probabilité de repérer plus rapidement les documents. Au plan sémantique, une dénomination explicite des classes (par exemple, par utilisation de définitions ou en évitant acronymes et abréviations) augmente la probabilité de succès au repérage. Enfin, un schéma de classification caractérisé par une macro-structure peu étendue, une logique basée partiellement sur une division par classes d'activités et une sémantique qui utilise peu d'abréviations augmente la probabilité de repérer les documents dès le premier essai.

Mots-clés : schéma de classification, classification, repérage, document électronique, document administratif, archives, gestion personnelle de l'information, organisation des documents, théorie de la classification, principes archivistiques

Abstract

The employees of an organization often use a personal classification scheme to organize electronic documents residing on their own workstations. As this may make it hard for other employees to retrieve these documents, there is a risk for the organization of losing track of needed documentation. To this day, no empirical study has been conducted to verify whether personal classification schemes allow, or even facilitate the retrieval of documents created and classed by someone else, in collaborative work, for example, or when it becomes necessary to reconstruct a “dossier”.

The first objective of our research was to describe the characteristics of personal classification schemes used to organize and classify administrative electronic documents. Our second objective was to verify, in a controlled environment, differences as to retrieval effectiveness which would be linked to the characteristics of classification schemes. More precisely, we wanted to verify if it was possible to find a document with the same effectiveness, whatever the classification scheme used.

Two types of data collection were necessary to reach those objectives. We first identified the structural, logical and semantic characteristics of 21 classification schemes used by Université de Montréal employees to organize and classify electronic documents residing on their own workstations. We then compared, in a controlled experimentation, the capacity of 70 participants to find electronic documents with the help of five classification schemes exhibiting variations in their structural, logical and semantic characteristics. Three variables were used to measure retrieval effectiveness : the number of documents found, the average time needed (in seconds) to locate the documents and the number of documents found on the first try.

Results revealed many structural, logical and semantic characteristics common to a majority of personal classification schemes : extended macro-structures, shallow, complex and unbalanced structures, thematic grouping, alphabetical order of classes, etc. An analysis of variance revealed significant differences as to retrieval effectiveness that are related to the structural, logical and semantic characteristics of the classification

scheme. A classification scheme characterized by a narrow macro-structure and a logic based on classes of activities increases the probability of finding documents more rapidly. On the semantic level, more explicit denominations of classes (for example, by using definitions or avoiding acronyms and abbreviations) increases the probability of success in finding documents. Finally, a classification scheme characterized by a narrow macro-structure, a logic based on classes of activities, and a semantic that uses few abbreviations minimizes the risk of error and failure in retrieval.

Keywords : classification schemes, classification, retrieval, electronic records, administrative records, archives, personal information management, document organization, classification theory, archival principles

Table des matières

Liste des tableaux	xi
Liste des figures	xiii
Liste des sigles.....	xiv
1 Introduction	1
1.1 Problématique	1
1.2 Définitions.....	7
1.3 But et objectifs de la recherche	8
1.4 Questions de recherche	8
1.5 Plan de la thèse.....	9
2 Revue de la littérature	11
2.1 Introduction.....	11
2.2 Assises théoriques de la classification documentaire.....	12
2.2.1 Nature et fonctions de la classification documentaire.....	12
2.2.2 Objet de la classification documentaire	14
2.2.3 Outil de classification documentaire.....	15
2.3 Principes de classification des documents administratifs en archivistique.....	22
2.3.1 Nature et fonctions de la classification des documents administratifs.....	23
2.3.2 Objet de la classification des documents administratifs	26
2.3.3 Outil de classification des documents administratifs	27
2.4 Organisation personnelle des documents administratifs électroniques.....	39
2.4.1 Nature et fonction de l'organisation personnelle des documents administratifs électroniques	40
2.4.2 Objet de l'organisation personnelle des documents administratifs électroniques	41
2.4.3 Outil d'organisation personnelle des documents administratifs électroniques	41
2.5 Repérage des documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles	51
2.5.1 Nature et fonction du repérage des documents administratifs électroniques dans un contexte de gestion documentaire décentralisée.....	51
2.5.2 Objet du repérage des documents	52
2.5.3 Facteurs qui influencent la qualité du repérage.....	53

2.5.4	Schéma de classification et repérage.....	60
2.6	Modèle d'analyse des schémas de classification personnels	66
2.6.1	Indicateurs de mesure de la dimension structurelle des schémas de classification.....	67
2.6.2	Indicateurs de mesure de la dimension logique des schémas de classification.....	69
2.6.3	Indicateurs de mesure de la dimension sémantique des schémas de classification.....	70
2.7	Conclusion	74
3	Méthodologie	75
3.1	Introduction.....	75
3.2	Caractérisation des schémas de classification.....	76
3.2.1	Procédure d'échantillonnage et taille de l'échantillon	77
3.2.2	Collecte des données	81
3.2.3	Méthode d'analyse des données.....	90
3.2.4	Caractéristiques du schéma de classification personnel retenu pour des fins de manipulation (Schéma A).....	92
3.3	Manipulation des schémas de classification	95
3.3.1	Nature des manipulations.....	96
3.3.2	Produits de la manipulation.....	96
3.4	Hypothèses de recherche.....	107
3.5	Simulation du repérage	108
3.5.1	Procédure d'échantillonnage et taille de l'échantillon	109
3.5.2	Collecte des données	112
3.5.3	Méthode d'analyse des données.....	122
3.6	Qualité de la recherche.....	123
3.6.1	Validité de la recherche.....	123
3.6.2	Fiabilité de la recherche	126
3.7	Conclusion	128
4	Résultats	130
4.1	Introduction.....	130
4.2	Résultats de la caractérisation des schémas de classification personnels	131
4.2.1	Caractéristiques structurelles des schémas de classification observés.....	132
4.2.2	Caractéristiques logiques des schémas de classification observés.....	145
4.2.3	Caractéristiques sémantiques des schémas de classification observés	150
4.3	Résultats de la simulation du repérage.....	156
4.3.1	Observations générales	157
4.3.2	Tests d'hypothèses sur la proportion de documents repérés.....	162

4.3.3	Tests d'hypothèses sur le temps moyen du repérage	166
4.3.4	Tests d'hypothèses sur la proportion de documents repérés dès le premier essai	169
4.4	Conclusion	175
5	Discussion.....	177
5.1	Introduction.....	177
5.2	Caractéristiques des schémas de classification personnels observés	177
5.2.1	Caractéristiques structurelles des schémas de classification observés.....	178
5.2.2	Caractéristiques logiques des schémas de classification observés.....	185
5.2.3	Caractéristiques sémantiques des schémas de classification observés	191
5.2.4	Conclusion	195
5.3	Schémas de classification et repérage	197
5.3.1	Schémas de classification et proportion de documents repérés	198
5.3.2	Schémas de classification et temps moyen du repérage.....	203
5.3.3	Schémas de classification et proportion de documents repérés dès le premier essai	207
5.3.4	Conclusion	211
5.4	Limites de la recherche	213
5.5	Conclusion	214
6	Conclusion.....	217
6.1	Résumé de la recherche.....	217
6.2	Contributions dans le domaine des sciences de l'information	222
6.3	Recommandations pour des recherches futures	224
6.3.1	Valider la relation entre la classification personnelle et la constitution du fonds d'archives	224
6.3.2	Tester les schémas de classification dans un contexte réel d'utilisation...	225
6.3.3	Vérifier l'influence de facteurs socio-démographiques sur les caractéristiques des schémas de classification personnels	226
6.3.4	Étudier les schémas de classification personnels utilisés pour l'organisation et le repérage des documents papier et des courriels	227
6.3.5	Identifier plus précisément les facteurs d'influence de l'efficacité du repérage	228
6.3.6	Explorer la possibilité d'utiliser une classification à facettes pour l'organisation personnelle des documents administratifs électroniques ...	229
7	Bibliographie	231
	Annexe 1. Certificat d'éthique et formulaires de consentement	I

Annexe 2. Lettre d'appel à participation (employés).....	VII
Annexe 3. Questionnaire adressé aux techniciennes à la coordination de travail de bureau.....	VIII
Annexe 4. Questionnaire adressé aux coordonnateurs de stages	XXIII
Annexe 5. Classes du <i>Système officiel de classification</i> de l'Université de Montréal utilisées par quatre techniciennes.....	XL
Annexe 6. Protocole d'entrevue.....	XLI
Annexe 7. Exemple de saisie d'écran	XLIV
Annexe 8. Extrait d'un rapport produit par le logiciel <i>PDS</i>	XLV
Annexe 9. Caractéristiques des schémas de classification personnels observés..	XLVI
Annexe 10. Lettre d'appel à participation (Étudiants et diplômés)	LIII
Annexe 11. Exemple de grille d'observation (Groupe 1)	LIV
Annexe 12. Directives communiquées aux répondants au début de la séance de simulation.....	LIX
Annexe 13. Description des documents à repérer.....	LXI
Annexe 14. Préparation des données pour le SCMEQ (Extrait).....	LXIII
Annexe 15. Analyse de la variable « Succès ».....	LXIV
Annexe 16. Analyse de la variable « Temps ».....	LXXVI
Annexe 17. Analyse de la variable « Essai »	LXXX
Annexe 18. Modèle d'analyse des schémas de classification.....	LXXXVII
Annexe 19. Modélisation de la relation entre les variables Schéma de classification et Repérage.....	LXXXIX
Annexe 20. Synthèse des caractéristiques des schémas de classification personnels observés et du schéma A.....	XCI

Liste des tableaux

Tableau 2.1 Principaux critères de division logique recensés dans les écrits	47
Tableau 2.2 Principales caractéristiques de la structure arborescente des répertoires recensés dans les écrits	48
Tableau 2.3 Principales stratégies d'abréviation recensées dans les écrits et observées	72
Tableau 3.1 Caractéristiques des schémas de classification personnels étudiés et caractéristiques du schéma de classification personnel retenu pour des fins de manipulation (Schéma A)	93
Tableau 3.2 Schéma B : Manipulation de la sémantique	97
Tableau 3.3 Schéma C - Manipulation de la structure et de la sémantique	99
Tableau 3.4 Schéma D – Manipulation de la structure et de la logique.....	101
Tableau 3.5 Schéma E – Manipulation de la structure, de la logique et de la sémantique	103
Tableau 3.6 Caractéristiques structurelles, logiques et sémantiques des schémas de classification A, B, C, D et E	105
Tableau 3.7 Caractéristiques socio-démographiques des répondants pour chaque groupe expérimental.....	115
Tableau 3.8 Protocole de simulation basé sur le protocole expérimental de Hertzum & Frøkjær (1996, p. 146)	116
Tableau 3.9 Durée moyenne de la simulation (en minutes).....	121
Tableau 4.1 Critères de division logique observés aux 1er et 2ème niveaux hiérarchiques	146
Tableau 4.2 Typologie des différents ordres de succession des classes observés	147
Tableau 4.3 Proportion moyenne de classes redondantes dans les schémas de classification partagés et non partagés des techniciennes et des coordonnateurs de stages	155
Tableau 4.4 Fréquences pour le nombre de documents repérés.....	158
Tableau 4.5 Proportion de documents repérés et proportion de documents non repérés en rapport avec le nombre d'essais pour l'ensemble des schémas de classification.	160
Tableau 4.6 Schéma de classification et succès/échec/abandon	161
Tableau 4.7 Proportion de documents repérés pour chacun des schémas de classification selon le nombre d'essais	161
Tableau 4.8 Fréquences de la proportion de documents repérés par schéma de classification.....	163
Tableau 4.9 Proportion moyenne de documents repérés pour chacun des schémas de classification.....	164
Tableau 4.10 Test des effets intra répondants.....	164
Tableau 4.11 Comparaisons des moyennes par paires de schémas de classification	165
Tableau 4.12 Moyenne du temps moyen des réussites (mesuré en secondes) pour les répondants ayant repéré deux documents ou plus.....	167
Tableau 4.13 Moyenne du temps moyen requis pour repérer les documents ciblés (mesuré en secondes) pour les répondants ayant repéré trois documents ou plus	167

Tableau 4.14 Tests des effets intra-répondants	168
Tableau 4.15 Comparaisons des moyennes par paires de schémas de classification	168
Tableau 4.16 Données valides et manquantes pour la proportion de documents trouvés en un essai (au moins deux documents repérés) par schéma de classification	170
Tableau 4.17 Proportion de documents trouvés en un essai pour chacun des schémas de classification (un minimum de deux documents repérés)	170
Tableau 4.18 Tests des effets intra-répondants	171
Tableau 4.19 Comparaisons des moyennes par paires de schémas de classification	171
Tableau 4.20 Proportion de documents trouvés en un essai pour chacun des schémas de classification (un minimum de trois documents repérés).....	172
Tableau 4.21 Synthèse de résultats par rapport aux hypothèses spécifiques de recherche	173

Liste des figures

<i>Figure 2.1 Exemple de dénomination des classes dans un environnement analogique et dans un environnement électronique.</i>	37
<i>Figure 2.2 Indicateurs de mesure de l'étendue et de la profondeur d'une structure arborescente.</i>	68
<i>Figure 4.1 Nombre de classes principales observé dans les schémas de classification des techniciennes à la coordination de travail de bureau.</i>	133
<i>Figure 4.2 Nombre moyen de classes à chaque niveau hiérarchique observé dans les schémas de classification des techniciennes à la coordination de travail de bureau.</i>	134
<i>Figure 4.3 Nombre de niveaux maximum et nombre moyen de niveaux observés dans les schémas de classification des techniciennes à la coordination de travail de bureau.</i>	135
<i>Figure 4.4 Nombre total de classes observé dans les schémas de classification des techniciennes à la coordination de travail de bureau.</i>	136
<i>Figure 4.5 Écart-type de la moyenne du nombre de classes à chaque niveau pour chacun des schémas des techniciennes à la coordination de travail de bureau.</i>	137
<i>Figure 4.6 Nombre de classes principales observé dans les schémas de classification des coordonnateurs de stages.</i>	138
<i>Figure 4.7 Nombre moyen de classes à chaque niveau hiérarchique observé dans les schémas de classification des coordonnateurs de stages.</i>	139
<i>Figure 4.8 Nombre de niveaux maximum et nombre moyen de niveaux observés dans les schémas de classification des techniciennes à la coordination de travail de bureau.</i>	140
<i>Figure 4.9 Nombre total de classes observé dans les schémas de classification des coordonnateurs de stages.</i>	141
<i>Figure 4.10 Écart-type de la moyenne du nombre de classes à chaque niveau pour chacun des schémas de classification des coordonnateurs de stages.</i>	142
<i>Figure 5.1 Exemple de classes redondantes au premier niveau.</i>	194
<i>Figure 5.2 Exemple de classes redondantes au deuxième niveau.</i>	194
<i>Figure 5.3 Proportion des documents repérés à l'aide des schémas de classification.</i>	198
<i>Figure 5.4 Temps moyen requis pour repérer les documents à l'aide des schémas de classification.</i>	203
<i>Figure 5.5 Proportion de documents repérés dès le premier essai à l'aide des schémas de classification.</i>	207

Liste des sigles

ANOVA.....	Analyse de variance
CRÉPUQ.....	Conférence des recteurs et des principaux des universités du Québec
IR.....	Information Retrieval
PDS.....	Personal Document Space
SCMEQ...	Service de consultation en méthodes quantitatives de l'Université de Montréal
SPSS.....	Statistical package for social sciences

*À Éric,
sans qui cette belle aventure n'aurait jamais commencé*

*À ma mère,
modèle de pugnacité et de joie de vivre à toute épreuve*

*À mon père,
exemple de patience et de courage face à l'adversité*

Avec tout mon amour

Remerciements

Ce n'est pas sans émotion que j'écris ces lignes, les plus difficiles de ma thèse. La difficulté n'est pas dans l'acte de remercier, la difficulté est dans la formulation des remerciements. Comment remercier dans des mots? Comment traduire, sans trahir, toute ma gratitude? Comment contenir cette larme qui tombe sur ces lignes? Dans cette larme, il y a le souvenir de toutes ces années brillantes. Dans cette larme, il y a beaucoup de joies et quelques peines. Il y a des mots heureux et des gestes d'encouragement. Dans cette larme, il y a aussi de nombreux visages. Il y a les visages de mes proches, de mes professeurs, de mes collaborateurs, de mes répondants et de mes amis, gravés à jamais dans ma mémoire.

La première personne que je désire remercier est Louise Gagnon-Arguin, ma directrice de recherche, ma « lampe de sagesse ». Grâce à une expérience, une empathie, une intuition et une vision de l'archivistique hors du commun, Louise m'a habilement guidé dans mes travaux tout en me laissant la liberté nécessaire à la réflexion et à l'accomplissement de la recherche. Louise a su m'offrir non seulement un support intellectuel mais aussi un support humain lorsque le besoin s'en faisait le plus sentir. Merci Louise, pour m'avoir donné les moyens de réussir ce projet qui me tenait tant à cœur et pour m'avoir enseigné durant toutes ces années ce qui ne se trouve dans aucun livre.

S'il y a des personnes que l'on qualifie à tort de perspicaces et de savantes, ce n'est pas le cas de Michèle Hudon, qui m'a fait l'honneur d'être ma codirectrice. Tout en ne m'épargnant pas les critiques et tout en ayant un constant souci du détail, Michèle n'a jamais fait l'économie de judicieux conseils et d'encouragements opportuns tout au long de ces années. En tout temps, elle a manifesté à mon égard la patience de Job qui ne peut que m'encourager à être encore plus rigoureuse dans mes recherches.

Mes plus vifs remerciements vont aussi à la généreuse disponibilité et au regard avisé de Charles Ramangalahy, membre de mon comité de recherche. Les questions qu'il a soulevées lors de nos rencontres ont été une source d'inspiration. Ses remarques et commentaires toujours pertinents m'ont permis de clarifier certaines parties de ce manuscrit. J'éprouve un respect sincère pour son travail ainsi que pour ses qualités humaines.

Je tiens aussi à remercier le corps professoral et le personnel passé et présent de l'École de bibliothéconomie et des sciences de l'information (EBSI) qui a toujours su m'encourager et me guider dans ce projet doctoral. Merci à Suzanne Bertrand-Gastaldy et Pierrette Bergeron pour la qualité de leur enseignement dispensé dans le cadre des séminaires. Un merci tout particulier à Carol Couture pour avoir fait germer en moi l'idée d'effectuer un doctorat. À Marcel Lajeunesse qui m'a gentiment taquinée toutes ces années en me posant la sempiternelle question : « Pis, le doctorat ? ». Le doctorat se porte bien, merci Marcel. Merci aussi à Lisette, Maria, Céline, Diane, pour leur dévouement et leur efficacité. Merci à Gilles et Carole, à la Bibliothèque, pour avoir toujours gardé le sourire malgré mes incessantes demandes.

Le soutien financier n'est pas à négliger dans la persistance au sein d'un programme d'études supérieures. Un grand merci à l'EBSI pour les bourses d'excellence. Merci à la Faculté des études supérieures pour m'avoir octroyé une belle bourse qui m'a permis de terminer la dernière année de cette thèse dans d'excellentes conditions.

Pour m'avoir gracieusement offert leur expertise et leur collaboration tout au long de ce projet, merci à Claude Minotto, Diane Baillargeon, Taïk Bourhis, Nathalie Denis et Gilles Landry de la Division des archives de l'Université de Montréal. Pour le travail d'analyse et de vérification des données, merci à l'équipe du Service de consultation en méthodes quantitatives de l'Université de Montréal et plus particulièrement à Miguel Chagnon, pour ses précieux conseils. Je remercie également Daniel J. Gonçalves, et Joaquim A. Jorge, professeurs à l'Instituto Superior Técnico de Lisbonne, pour m'avoir

autorisée à utiliser leur logiciel pour la collecte de mes données. Que de temps sauvé grâce à votre petit logiciel !

Mille mercis aux techniciennes à la coordination de travail de bureau, aux coordonnateurs et coordonnatrices de stages de l'Université de Montréal, ainsi qu'aux étudiants et aux diplômés de l'EBSI, de Psychologie et de Mathématiques qui m'ont accordé leur confiance et fait don de leur temps en acceptant volontairement de participer à cette recherche. Sans vous, ce projet n'aurait pu se concrétiser.

Je tiens aussi à remercier très sincèrement Florence Arès pour son inconditionnel soutien. Elle a toujours été présente pour écarter les doutes, soigner les blessures et partager les joies. Merci aussi à Christine Dufour, Zeïneb Gharbi et Cynthia Couture pour leur amitié et leur écoute dans les moments difficiles. Cette thèse est un peu la leur aussi.

Pour terminer, un grand merci à toute la tribu des doctorants de l'EBSI qui m'a accompagnée et soutenue dans cette aventure. Permettez-moi maintenant de vous dire chers amis qu'il est vrai que : « *Ceux qui sèment avec larmes moissonneront avec chant de joie* ».

1 Introduction

1.1 Problématique

Le développement des technologies de l'information et des télécommunications a entraîné la création et l'utilisation d'un nombre toujours croissant de documents électroniques pour la réalisation des activités quotidiennes au sein des organismes (Berchtold, Biliris & Panagos, 1999; Chapman, 1999; Skupsky, 1999; Sprehe, McClure & Zellner, 2002; Waldron, 2002). De plus en plus de documents, qui existent uniquement sous forme électronique (Skupsky, 1999; Dhérent & collaborateurs, 2002; Patterson & Sprehe, 2002), possèdent une valeur administrative, légale, financière et sont essentiels à la réalisation des activités (Berlin *et al.*, 1993) ou à la défense des droits des organismes (Saulnier, 1997-1998). Parmi ces documents, certains ont également une valeur historique et font alors partie de la mémoire corporative (Bergeron, 1992).

Pour répondre au besoin de gérer, de retrouver et d'utiliser ces documents électroniques, les gestionnaires de documents préconisent l'utilisation du schéma de classification institutionnel déjà conçu pour l'organisation des documents sur support papier (Vigneau, 1996; Dhérent & collaborateurs, 2002; Guercio, 2002; Tough & Moss, 2003; Conférence des recteurs et des principaux des universités du Québec, 2004). Cet outil d'organisation basé sur une structure hiérarchique de catégories descriptives relatives aux fonctions et activités de l'institution permet de regrouper logiquement et physiquement tous les documents relatifs à une même activité ou à un même dossier. Il est établi que le schéma de classification institutionnel permet l'organisation des documents à plus ou moins long terme et facilite entre autres la localisation et le repérage des documents par l'ensemble des employés (Rousseau, 1980; Conférence des recteurs et des principaux des universités du Québec, 1994; Gagnon-Arguin & Bannouri, 1998; Lévesque, 1998-1999; Héon, 1999; Sénécal, 1999; Dhérent & collaborateurs, 2002; Sprehe, McClure & Zellner, 2002).

Malgré les recommandations à cet effet, il est démontré que le schéma de classification institutionnel est rarement utilisé par les employés pour organiser et retrouver les documents électroniques se trouvant sur leur poste de travail. Les employés privilégient des schémas de classification plus « personnels » des documents électroniques qui répondent davantage à leurs besoins pour la réalisation des activités quotidiennes qu'à une vision institutionnelle (Gagnon-Arguin & Bannouri, 1998; Sénécal, 1999; Skupsky, 1999; Peyrelong & Accart, 2002; Tough & Moss, 2003).

Cette situation peut s'expliquer tout d'abord par les changements survenus dans le rôle des employés au sein des organismes. L'utilisation répandue des micro-ordinateurs depuis le milieu des années 1980 a favorisé une décentralisation et une démocratisation des fonctions de création, d'organisation, de contrôle, de manipulation et de diffusion des documents en format numérique (Tough & Moss, 2003). Alors qu'il appartenait traditionnellement au responsable d'un poste de classement de procéder à l'analyse, à la classification et à l'indexation de ses documents ou de ses dossiers, c'est désormais à chaque employé que revient la responsabilité de manipuler et gérer les documents électroniques créés ou reçus à partir de son poste de travail informatique (Bergeron, 1992; Berlin *et al.*, 1993; Kaptelinin, 1996; Gagnon-Arguin & Bannouri, 1998; Roberge, 2002; Saffady, 2002; Roberge, 2004) :

« De façon générale, on peut affirmer aujourd'hui que, dans une organisation, toutes et tous sont des *gestionnaires de documents* : les secrétaires, plus particulièrement, mais aussi le personnel d'encadrement, le personnel professionnel et le personnel technique créent, reçoivent, identifient, classifient, classent, repèrent, utilisent, déclassent et éliminent les documents dont la conservation est de plus en plus décentralisée et déconcentrée » (Roberge, 2002, p. 6.6).

En fait, on observe un paradigme « individualiste » ou « d'informatique individuelle » qui contourne les pratiques de gestion traditionnelles (Bergeron, 1992, p. 59). Les documents sont « auto-gérés » par les créateurs et utilisateurs, qui ne sont plus uniquement des personnes employées à titre de secrétaires et qui n'ont souvent ni

mandat, ni formation, ni incitation à se préoccuper eux-mêmes de la gestion des documents électroniques.

Plusieurs facteurs institutionnels influencent et favorisent l'individualisme qui caractérise l'organisation des documents électroniques dans un contexte de gestion décentralisée. La gestion des documents ne constitue généralement pas une priorité pour les organismes en raison d'un manque de motivation institutionnelle ou de ressources (Hedstrom, 1996; Sprehe, McClure & Zellner, 2002). Cette situation est renforcée par le fait que de nombreuses personnes pensent que les documents électroniques ne sont pas régis par les lois sur les archives ou ont du mal à faire une distinction entre les documents d'archives électroniques et les autres sources d'information électronique telles que les ressources Web, par exemple (Sprehe, McClure & Zellner, 2002). Des enquêtes ont permis de constater dans la majorité des organismes la quasi-absence de politiques, normes, méthodes et procédures de gestion complète et intégrée des documents électroniques, ce qui entraîne une perte de contrôle sur une partie importante de la mémoire institutionnelle (Karjalainen *et al.*, 2000; Sprehe, McClure & Zellner, 2002). L'absence de formation ou d'encadrement des employés (Sprehe, McClure & Zellner, 2002), de directives ou de conseils (Conférence des recteurs et des principaux des universités du Québec, 1994) conduit à la manipulation « sans règle ni discipline » des documents électroniques par les employés (Skupsky, 1999, p. 30). En l'absence d'un cadre normalisé, un organisme se retrouve avec autant de façons de faire qu'il y a d'individus (Bergeron, 1992; Jones, 2007).

À tous ces éléments s'ajoute la « complexité inhérente » à l'utilisation du schéma de classification institutionnel (Université de Montréal, 1995; Archives nationales du Canada, 1996; Saulnier, 1997-1998; Gagnon-Arguin & Bannouri, 1998, p. 30; Lévesque, 1998-1999; Sénécal, 1999; Patterson & Sprehe, 2002). Celui-ci, tel qu'utilisé pour l'organisation des documents sur support papier, peut difficilement être appliqué dans l'environnement électronique en raison de plusieurs contraintes qui existent *a priori*. Tout d'abord, la terminologie souvent trop générique utilisée dans la structure du schéma ne correspond pas toujours à celle qu'utilisent les employés dans leur travail. Ces

derniers nomment leurs répertoires et leurs documents électroniques selon une sémantique personnelle, souvent reliée au domaine spécialisé dans lequel ils évoluent (Lévesque, 1998-1999; Sénécal, 1999). Ensuite, on peut souligner la structure hiérarchique trop rigide du schéma de classification institutionnel (Université de Montréal, 1995; Archives nationales du Canada, 1996; Saulnier, 1997-1998; Gagnon-Arguin & Bannouri, 1998; Sénécal, 1999) et sa mise à jour irrégulière, voire inexistante (Patterson & Sprehe, 2002), qui n'incitent pas les individus à l'utiliser dans l'environnement électronique. Enfin, l'effort cognitif dans l'acte de classier et le manque de temps (Malone, 1983; Boardman, 2001a; Henderson, 2003) incitent les individus à privilégier l'utilisation d'un schéma de classification personnel pour l'organisation des documents électroniques. Les individus n'ont plus à négocier ni à discuter les règles de division logique ou conventions classificatoires qu'ils utilisent; ils se rappellent de leurs propres règles classificatoires, des changements et des exceptions (Kwasnik, 1989; Berlin *et al.*, 1993; Saffady, 2002).

Toutes ces considérations ont conduit les gestionnaires de documents à revoir leurs méthodes et à proposer des solutions en vue d'uniformiser l'organisation des documents électroniques à caractère institutionnel. Depuis les années 1990, plusieurs organismes, groupes de travail et professionnels ont mené des réflexions, formulé des recommandations et des normes en vue d'améliorer l'organisation des documents électroniques sur un micro-ordinateur (Conférence des recteurs et des principaux des universités du Québec, 1994; Université de Montréal, 1995; Vigneau, 1996; Conseil international des Archives, 1997; Saulnier, 1997-1998; U.S. National Archives and Records Administration, 2001; Poivre, 2004). Ainsi diverses propositions ont été énoncées pour simplifier et faciliter l'implantation et l'utilisation du schéma de classification institutionnel dans l'environnement électronique : formulation des règles de dénomination de fichiers et répertoires, définition de métadonnées, diminution du nombre de niveaux hiérarchiques de la structure de classification et personnalisation des titres des catégories. Malgré les adaptations effectuées, les employés qui ont eu recours au schéma de classification institutionnel pour organiser leurs documents électroniques

se disent souvent insatisfaits en raison de la « logique de la classification » qui n'est pas la même que la leur (Gagnon-Arguin & Bannouri, 1998, p. 23). Les employés interrogés par Gagnon-Arguin & Bannouri (1998, p. 23) préfèrent utiliser soit un schéma de classification « spécifique » issu du schéma de classification institutionnel, soit un schéma de classification « parallèle », plus approprié à l'organisation des documents électroniques se trouvant sur leur poste de travail informatique.

L'exploitation de solutions technologiques a également été envisagée pour aider les individus à organiser leurs documents. Plusieurs méthodes statistiques (par exemple, fréquences des mots, pondération des mots), linguistiques (par exemple, sémantique, syntaxique) ou d'intelligence artificielle (par exemple, systèmes expert, réseaux neuronaux) ont été développées pour organiser automatiquement des données ou des documents électroniques (Savic, 1995). Toutefois, la classification automatique basée sur le traitement de la langue naturelle et la technologie des réseaux neuronaux n'est pas encore implantée dans les organismes (Sprehe, McClure & Zellner, 2002) et son efficacité reste à démontrer pour l'organisation des documents électroniques institutionnels (Savic, 1995; Patterson & Sprehe, 2002).

L'individualisme qui caractérise l'organisation des documents électroniques dans un contexte de gestion décentralisée pose des problèmes multiples pour les organismes. Il est reconnu que l'absence d'organisation systématique et uniforme des documents électroniques conduit à plusieurs problèmes dont les plus critiques sont :

1. l'éparpillement des documents relatifs à une même affaire (Skupsky, 1999; Guercio, 2002);

2. l'impossibilité d'assurer la défense des droits et la conservation de la trace des activités de l'organisme en raison de la perte possible d'un nombre important de documents (même ceux qui n'ont pas d'équivalents sur papier) (Hedstrom, 1996; Lévesque, 1998-1999; Skupsky, 1999; U.S. National Archives and Records Administration, 2001; Guercio, 2002);

3. les difficultés d'interprétation, de repérage, de partage et de réutilisation de l'information consignée liées aux modes d'organisation individuels basés sur des critères très personnels qui s'avèrent le plus souvent incompréhensibles aux autres utilisateurs (Kwasnik, 1991; Berlin *et al.*, 1993; Kidd, 1994; Gagnon-Arguin & Bannouri, 1998; Sénécal, 1999; Skupsky, 1999; Sellen & Harper, 2001; Saffady, 2002; Bergman, Beyth-Marom & Nachmias, 2003).

Alors que les créateurs et utilisateurs premiers des documents semblent s'y retrouver (Chapman, 1999; Boardman & Sasse, 2004), l'utilisation de stratégies personnelles d'organisation des documents électroniques nous amène à nous poser plusieurs questions : Quelles sont les caractéristiques des schémas de classification conçus et utilisés par les employés pour organiser leurs documents électroniques? Qu'en est-il de la capacité de tierces personnes à repérer les documents électroniques à l'aide de schémas de classification personnels avec lesquels ils ne sont pas familiers? Existe-t-il des différences quant à l'efficacité du repérage réalisé à l'aide de schémas de classification ayant des caractéristiques variées?

Les recherches portant sur les schémas de classification utilisés par des catégories particulières d'employés pour organiser leurs documents électroniques dans un contexte de travail sont peu nombreuses. Aucune étude empirique n'a été menée à ce jour afin de vérifier dans quelle mesure les schémas de classification personnels existants permettent, ou même facilitent, le repérage des documents électroniques par des tiers, dans le cadre d'un travail collaboratif par exemple, lorsqu'il s'agit de reconstituer un dossier ou lorsque des employés quittent l'organisation. L'objectif de notre recherche est d'augmenter nos connaissances sur les schémas de classification personnels utilisées en contexte de travail et de mesurer empiriquement la capacité de tierces personnes à repérer des documents à l'aide de schémas de classification personnels avec lesquels ils ne sont pas familiers.

1.2 Définitions

Dans le contexte de notre recherche, les principaux concepts sont définis de la manière suivante :

Document administratif

Information consignée sous une forme analogique ou numérique dont la production, la réception et l'utilisation s'inscrivent dans la réalisation des activités passées ou présentes d'un organisme.

Document administratif électronique

Information consignée sur un ou plusieurs fichiers informatiques et dont la production, la réception et l'utilisation s'inscrivent dans la réalisation des activités passées ou présentes d'un organisme.

Gestion décentralisée des ressources informationnelles

Gestion effectuée par un individu qui exerce un contrôle direct sur la création, l'organisation, la manipulation et la destruction des documents entreposés sur son espace de travail personnel (Pétilat, 1999; Bergman, Beyth-Marom, & Nachmias, 2003; Boardman & Sasse, 2004).

Repérage d'un document

Activité qui consiste à localiser et à identifier un document cible dans une base de documents plus ou moins volumineuse (Chapman, 1999; Toms & Campbell, 1999).

Schéma de classification

Structure généralement hiérarchique de catégories descriptives conçue et utilisée par un individu ou un organisme pour faciliter l'organisation, le repérage et l'utilisation des documents.

Schéma de classification institutionnel

Structure hiérarchique institutionnelle de catégories descriptives conçue par un professionnel en gestion documentaire et approuvée par les instances dirigeantes d'un organisme pour faciliter l'organisation, le repérage et l'utilisation, par le personnel, de l'ensemble des documents administratifs.

Schéma de classification personnel

Structure généralement hiérarchique de catégories descriptives conçue et utilisée par un individu pour faciliter l'organisation, le repérage et l'utilisation d'un ensemble de documents administratifs qui sont sous son contrôle direct.

1.3 But et objectifs de la recherche

Le but de notre recherche est d'étudier l'efficacité du repérage de documents électroniques à l'aide de schémas de classification ayant des caractéristiques variées. Plus précisément, notre recherche vise à identifier certains facteurs qui favorisent un repérage efficace d'un ensemble donné de documents électroniques.

Les objectifs spécifiques (OS) de notre recherche sont de :

OS1. : Décrire les caractéristiques de différents schémas de classification utilisés par des employés pour organiser et classer les documents administratifs électroniques qui sont sous leur contrôle direct.

OS2. : Vérifier les différences sur le plan de l'efficacité du repérage de documents électroniques dans un environnement contrôlé qui sont fonction du schéma de classification utilisé.

1.4 Questions de recherche

Notre recherche vise à répondre à la question générale de recherche suivante :

Peut-on repérer avec la même efficacité un document électronique à l'aide de schémas de classification ayant des caractéristiques variées?

Les questions de recherche spécifiques (QRS) sont les suivantes :

QRS1. Quelles sont les caractéristiques des schémas de classification utilisés par les employés pour organiser et classer les documents administratifs électroniques qui sont sous leur contrôle direct?

QRS2. Existe-t-il des différences sur le plan de la capacité à repérer tous les documents ciblés à l'aide de schémas de classification ayant des caractéristiques variées?

QRS3. Existe-t-il des différences sur le plan de la capacité à repérer rapidement les documents ciblés à l'aide de schémas de classification ayant des caractéristiques variées?

QRS4. Existe-t-il des différences sur le plan de la capacité à repérer les documents ciblés dès le premier essai à l'aide de schémas de classification ayant des caractéristiques variées?

1.5 Plan de la thèse

Le Chapitre 2 présente une recension des écrits et des résultats de recherches pertinents à notre sujet. La revue de la littérature est organisée en cinq sections :

1. la nature, les principales fonctions et l'objet de la classification documentaire, ainsi que les principes classificatoires à la base de l'élaboration des schémas de classification documentaire;
2. la particularité de la nature, des fonctions et de l'objet de la classification des documents administratifs, ainsi que les principes d'élaboration des schémas de classification institutionnels des documents administratifs;
3. la nature et les fonctions de l'organisation personnelle des documents administratifs électroniques, la nature de l'objet à organiser, les pratiques d'élaboration des schémas de classification personnels des documents administratifs;
4. la spécificité de la nature, des fonctions et de l'objet du repérage appliqué aux documents administratifs électroniques, les facteurs déterminant la qualité du repérage et la relation théorique qui existe entre schéma de classification et repérage;
5. le modèle d'analyse des schémas de classification personnels des documents électroniques conçu par la chercheuse sur la base d'indicateurs de mesure recensés dans les écrits reliés à la gestion personnelle de l'information, aux sciences et technologies de l'information et à la théorie de la classification.

Le Chapitre 3 présente la méthodologie utilisée. Les hypothèses de recherche sont présentées dans la section 3.4. La présentation de la méthodologie est articulée selon les trois étapes de notre protocole de recherche : la description des caractéristiques des schémas de classification personnels, la manipulation et la création artificielle de schémas de classification en vue de constituer un corpus de schémas de classification variés et la simulation du repérage d'un ensemble de documents ciblés à l'aide de ce corpus. La première partie de ce chapitre décrit la procédure d'échantillonnage des schémas de classification personnels et la taille de l'échantillon ainsi que les méthodes de collecte et d'analyse des données. La deuxième partie du Chapitre 3 décrit les différentes manipulations qui ont été réalisées sur un schéma de classification personnel afin de pouvoir tester, dans un environnement contrôlé, l'efficacité du repérage effectué à l'aide de cinq schémas de classification ayant des caractéristiques variées. La procédure de sélection des participants à la simulation, la taille de l'échantillon ainsi que les méthodes de collecte et d'analyse des données utilisées à cette troisième étape de notre recherche sont présentées dans la troisième partie du Chapitre 3. Les mesures prises pour assurer la qualité de la recherche en termes de validité et de fiabilité sont présentées dans la dernière partie du chapitre.

Le Chapitre 4 présente les caractéristiques des schémas de classification personnels observés et les résultats de la simulation du repérage.

Les résultats sont discutés dans le Chapitre 5 en regard de nos questions et hypothèses de recherche. Les limites théoriques et méthodologiques de la recherche sont présentées dans la dernière section du Chapitre 5.

Un résumé de notre recherche est présenté dans le dernier chapitre de la thèse. Par la suite, nous décrivons les retombées théoriques, méthodologiques et pratiques de notre recherche. Enfin, nous formulons des recommandations pour des recherches futures.

2 Revue de la littérature

2.1 Introduction

Ce chapitre présente une recension des écrits et des résultats de recherches pertinents à notre sujet. L'objectif de cette recension des écrits est de présenter les principaux concepts et les fondements théoriques sur lesquels s'appuie notre recherche. La recension des écrits s'articule autour de deux domaines de connaissances reliés à notre recherche sur les schémas de classification personnels des documents électroniques. Premièrement, la littérature des sciences de l'information comporte de nombreux écrits relatifs à la théorie de la classification, aux principes théoriques d'élaboration des schémas de classification et à leur relation avec le repérage des documents dans un contexte documentaire traditionnel et dans un contexte administratif plus pragmatique. Deuxièmement, la littérature dans le domaine de la gestion personnelle de l'information (« *Personal Information Management* ») nous informe des principales caractéristiques des schémas de classification conçus par les individus pour organiser leurs documents dans un environnement électronique.

La recension des écrits relatifs aux assises théoriques de la classification documentaire est développée dans la section 2.2. La section 2.3 présente la classification appliquée plus spécifiquement aux documents dans un contexte administratif et les principes sous-jacents à l'élaboration des schémas de classification de cette catégorie de documents. Le contexte général relatif à l'organisation personnelle des documents administratifs électroniques est décrit dans la section 2.4. La section 2.5 aborde la question du repérage des documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles et identifie les indicateurs retenus dans notre recherche pour mesurer l'efficacité de ce repérage. Pour terminer, la section 2.6 présente le modèle d'analyse des schémas de classification que nous avons conçu à partir des indicateurs de mesure recensés dans les écrits.

2.2 Assises théoriques de la classification documentaire

L'expression « théorie de la classification », souvent mentionnée dans les écrits en sciences de l'information, réfère essentiellement aux principes classificatoires (Smiraglia, 2001; Smiraglia, 2002; Mai, 2004). L'objectif de cette section est de présenter la nature et les principales fonctions de la classification documentaire (section 2.2.1), la nature de l'objet de cette classification (section 2.2.2) ainsi que les principes classificatoires à la base de l'élaboration des schémas de classification documentaire (section 2.2.3). Nous partons de l'hypothèse, déjà formulée par plusieurs chercheurs, selon laquelle l'application des principes théoriques de la classification documentaire constitue une base solide pour la conception d'un système efficace de classification des documents (Bates, 1977; Svenonius, 1983; Cochrane, 1985; Van der Walt, 2004). La connaissance de ces principes théoriques est pertinente et essentielle pour nous permettre ultérieurement de caractériser et comparer les schémas de classification personnels sur la base de critères inspirés de la théorie de la classification.

2.2.1 Nature et fonctions de la classification documentaire

Cette section définit le concept de classification documentaire, aborde certaines considérations terminologiques et dégage les principales fonctions de la classification documentaire.

2.2.1.1 La classification documentaire

La classification dans son sens le plus général désigne l'opération qui consiste à regrouper des entités en classes sur la base de qualités ou caractéristiques communes, de sorte que les entités semblables soient regroupées et séparées des entités non semblables (Maniez, 1987; Hunter, 2000; Hunter, 2002). Quand ces entités sont des documents, on parle de « classification documentaire » (Satija, 2000).

2.2.1.2 Considérations terminologiques

La littérature scientifique rend bien compte de la nature polysémique du concept de classification documentaire. Selon le qualificatif que l'on adjoint à sa dénomination, la classification documentaire prend un sens différent ou nous renseigne sur le niveau de

détail ou de spécificité de la classification (par exemple, classification « en profondeur » ou classification « sommaire »), sur son domaine d'application (par exemple, classification « universelle » ou classification « spécialisée »), ou sur son élaboration (par exemple, classification « professionnelle » ou classification « naïve ») (Satija, 2000; Beghtol, 2003). Ainsi, l'organisation hiérarchique des connaissances d'un domaine spécifique (par exemple, la Médecine) est désignée par les chercheurs par le terme de « classification spécialisée », alors que la classification de l'ensemble des connaissances est connue sous les termes de « classification universelle », « classification générale » ou « classification de type philosophique » (par exemple, la Classification Décimale de Dewey) (Maniez, 1987; Satija, 2000; Hudon, 2001). D'autre part, une classification formelle, effectuée à partir d'une structure imposée est qualifiée par Beghtol (2003) de « professionnelle » alors qu'une classification instinctive faite par les créateurs de documents eux-mêmes est considérée comme « naïve » (Beghtol, 2003).

Dans le cadre de notre recherche, nous nous intéressons plus particulièrement aux classifications « spécialisées » (milieu administratif) et « naïves » (conçues par des employés au sein d'un organisme).

2.2.1.3 Fonctions de la classification documentaire

La littérature fait état de multiples fonctions et utilisations pratiques de la classification dans un contexte documentaire (Marcella & Newton, 1994; Hunter, 2000; Guercio, 2002; Beghtol, 2003). Nous considérons ici deux fonctions principales de la classification documentaire qui sont pertinentes à notre recherche pour déterminer le rôle précis de la classification dans la gestion d'une collection de ressources documentaires.

Une des principales fonctions de la classification documentaire est de guider le regroupement systématique des documents concernant un sujet en vue de répondre à un besoin de classement physique et de localisation des documents (Marcella & Newton, 1994; Rolland-Thomas, 1994; Hunter, 2000). La présentation linéaire des sujets vise à permettre et à faciliter le « bouquinage », la stratégie de recherche la plus pratiquée par

la majorité des utilisateurs d'une bibliothèque ou d'un service d'information (Marcella & Newton, 1994).

La classification a également un rôle significatif à jouer dans le processus de recherche en « contextualisant » le sujet de la recherche, c'est-à-dire en amenant progressivement le chercheur vers les documents susceptibles de répondre pertinemment à sa question (Hunter, 2000; Hudon, 2001). Par exemple, une recherche de documents relatifs aux « roses » peut être élargie à des sujets plus généraux tels que « fleurs » ou « jardinage » auxquels ce sujet spécifique peut potentiellement être lié (Hunter, 2000).

En résumé, la classification documentaire est un processus essentiel pour l'organisation de documents ainsi que pour la recherche systématique de l'information qu'ils contiennent (Hunter, 2000). Dans un contexte documentaire, on classe avant tout pour regrouper logiquement une grande quantité de ressources documentaires portant sur le même sujet afin de pouvoir répondre de manière pertinente à différentes pratiques de recherche (par exemple, bouquinage et recherche systématique) et de façon à permettre de rechercher l'information avec un minimum de complication et d'effort (Sayers & Maltby, 1975; Satija, 2000).

2.2.2 Objet de la classification documentaire

Plusieurs chercheurs se sont intéressés à l'objet spécifique de la classification documentaire (Langridge, 1992; Rolland-Thomas, 1994; Mai, 2004). Alors que la classification scientifique est basée directement sur les objets tels qu'ils sont observés dans le monde, la classification documentaire prend plutôt en considération le sujet ou la connaissance véhiculée par le document.

Nous retiendrons que la classification documentaire organise les différentes disciplines ou domaines de la connaissance dans lesquels les sujets contenus dans les documents s'insèrent afin de permettre le regroupement de documents traitant de sujets similaires selon un système méthodique, systématique et logique de classes et de sous-classes (Maniez, 1987; Rolland-Thomas, 1994; Lin & Chan, 1999; Mai, 2000; Satija, 2000).

2.2.3 Outil de classification documentaire

Cette section aborde certaines considérations terminologiques, définit la nature et les principales fonctions de l’outil de classification documentaire, décrit ensuite les principaux modèles structurels et présente enfin les principes théoriques et les méthodes spécifiques à l’élaboration des schémas de classification documentaire de type hiérarchique, les plus connus et utilisés.

2.2.3.1 Considérations terminologiques

On constate que différents termes sont employés pour désigner l’outil de classification documentaire. Certains chercheurs utilisent les termes « langage classificatoire » (Maniez, 1987), « plan de classement » (Nasse-Kolmayer, 1997), voire même « classification », pour y faire référence (Maniez, 1987). Afin d’éviter toute ambiguïté liée à l’emploi du terme « classification » ou à celui de « classement », nous utiliserons dans notre thèse l’expression « schéma de classification », couramment utilisée dans les écrits pour désigner l’outil permettant la représentation de cette organisation intellectuelle (Broughton, 2000; Fallis & Mathiesen, 2000; Jones & Maier, 2003; Hertzum, 2004; Mai, 2004; Van der Walt, 2004).

2.2.3.2 Nature du schéma de classification documentaire

Il existe peu de définitions explicites et suffisamment complètes du concept de schéma de classification documentaire. On associe le plus souvent cet outil à une « structure classificatoire », à une « hiérarchie » ou à un « système de classification » (Sukiasyan, 1998) sans toujours en préciser la nature. Dans le cadre de notre recherche, nous retiendrons qu’un schéma de classification documentaire est une structure généralement hiérarchique de catégories descriptives visant à faciliter l’organisation, le repérage et l’utilisation d’une collection documentaire.

2.2.3.3 Fonctions du schéma de classification documentaire

Les schémas de classification sont utilisés comme cadres de classement uniformes et permettent d’organiser et de repérer les documents quand le nombre de ces derniers devient important. Il peut s’agir de milliers, voire de millions d’unités (Lin &

Chan, 1999). Afin de faciliter le repérage des documents, les domaines dans lesquels s'insèrent les nombreux sujets traités dans ces documents doivent être regroupés puis différenciés en séries de classes et de sous-classes. Pour que cette classification rationnelle se manifeste concrètement dans le rangement des documents, on attribue à ces derniers une cote de classement (adresse d'un document dans les rayons) afin que les documents traitant de sujets voisins soient classés méthodiquement à proximité les uns des autres (Maniez, 1987).

2.2.3.4 Structure du schéma de classification documentaire

La structure d'un schéma de classification est considérée par Sukiasyan (1998) comme une caractéristique de son organisation interne et de la relation mutuelle entre les différents éléments qui le composent.

Maniez affirme que la structure d'un schéma de classification documentaire possède une double fonction puisqu'elle doit permettre « à la fois la classification d'un domaine et le classement physique des documents, selon la nature de leur sujet » (Maniez, 1987, p. 43). Par conséquent, la structure doit être suffisamment stable, tout en étant souple, pour s'adapter aux inévitables changements dans un domaine des connaissances. Ces fonctions et contraintes expliquent le nombre limité de modèles structurels de schémas de classification documentaire recensés dans les écrits. Dans le cadre de notre recherche, nous présenterons brièvement seulement un d'entre eux : le modèle hiérarchique parce que c'est le plus répandu, même pour la classification des documents administratifs.

La structure d'un schéma de classification hiérarchique est déterminée par le principe de division logique, issu de la logique formelle et qui remonte à l'Antiquité grecque (Maniez, 1987; Maniez, 1991; Canonne, 1993). Selon Aristote, la nature est un ensemble unifié qui peut être subdivisé en classes « naturelles », c'est-à-dire en catégories basées sur des caractéristiques communes (Maniez, 1987), et chaque classe en sous-classes, et ainsi de suite. Ce processus suit de manière systématique et ordonnée un ensemble de règles d'association et de distinction (Richmond, 1990; Kwasnik, 1999).

Selon le principe de division logique, les différents objets ou sujets sont répartis dans un ordre logique à l'intérieur de classes qui s'emboîtent du général au plus précis. Dans de tels schémas de classification, toutes les rubriques doivent idéalement rendre compte de tous les concepts puisque le système de classification hiérarchique doit permettre d'intégrer toutes les connaissances (Richmond, 1990). Une politique constante de révision de ces schémas de classification est nécessaire afin que ces derniers tiennent compte de l'apparition de nouveaux sujets dans un domaine donné (Canonne, 1993). Malgré de nombreuses faiblesses relevées par plusieurs chercheurs (Hjørland & Albrechtsen, 1999; Beghtol, 2003; Mai, 2004), ce modèle hiérarchique est particulièrement efficace pour représenter les connaissances et amener progressivement et intuitivement le chercheur vers les documents susceptibles de contenir l'information recherchée (Hunter, 2000; Hudon, 2001).

2.2.3.5 Principes et méthodes d'élaboration de schémas de classification documentaire hiérarchiques spécialisés

Les principes et méthodes théoriques d'élaboration des schémas de classification documentaire hiérarchiques ont été décrits par de nombreux chercheurs dans le domaine des sciences de l'information (Sayers, 1915; Sayers & Maltby, 1975; Buchanan, 1979; Maniez, 1987; Langridge, 1992; Canonne, 1993; Iyer, 1995; Beghtol, 1998; Dahlberg, 1998; Hjørland & Albrechtsen, 1999; Mai, 2000; Smiraglia, 2001). Dans la recension de ces principes et méthodes visant à mener à bien l'élaboration d'un schéma de classification efficace, nous envisageons successivement les principes et méthodes de travail préconisées, le processus proprement dit d'élaboration, et la mise à jour régulière du schéma de classification ainsi conçu.

Principes et méthodes de travail

Le schéma de classification hiérarchique spécialisé doit être spécifiquement adapté à un fonds documentaire particulier en raison de sa double fonction qui est de permettre, rappelons-le, l'organisation logique d'un domaine spécialisé et l'organisation physique de documents qui lui est subordonnée (Langridge, 1992; Canonne, 1993; Iyer, 1995; Hudon, 2001). Pour ce faire, il faut tenir compte à la fois : (1) du nombre de

documents à organiser et (2) de la structure logique du domaine à classifier (Maniez, 1987).

D'une part, la garantie d'existence bibliographique (« *literary warrant* ») consiste à tenir compte du nombre de documents offerts sur chaque sujet pour s'assurer d'élaborer un schéma de classification spécifique à l'ensemble documentaire (Langridge, 1992; Canonne, 1993; Iyer, 1995). Autrement dit, une nouvelle classe n'est créée que si la quantité de documents portant sur un même sujet l'exige. Cette approche inductive (« *bottom-up* ») basée sur une référence constante aux documents à organiser permet de vérifier la pertinence du choix des classes et d'éviter de créer un nombre trop élevé de classes qui seront inutilisées par la suite (Langridge, 1992; Canonne, 1993; Iyer, 1995).

D'autre part, on subdivise le domaine conceptuel à structurer selon les règles de division logique, indépendamment du contenu du fonds documentaire. L'approche déductive (« *top-down* ») sous-jacente à cette structuration logique du domaine assure la cohérence et les possibilités d'évolution du schéma de classification (Maniez, 1987; Fallis & Mathiesen, 2000).

Hudon (2001) note que les schémas d'organisation les plus performants sont développés à partir d'une combinaison de ces deux approches qui tiennent compte à la fois de l'évolution bibliographique et de l'évolution scientifique.

Le processus d'élaboration

Plusieurs chercheurs se sont intéressés au processus d'élaboration et à l'identification d'un certain nombre de règles de cohérence visant à construire des schémas de classification plus prévisibles et donc plus efficaces (Maniez, 1987; Fallis & Mathiesen, 2000). Nous présentons principalement deux grandes étapes : (1) l'identification et l'ordre des classes et (2) le choix de la notation.

Identification et ordre des classes

Au niveau de la macro-structure

La détermination de la première caractéristique de division (ou critère de division) est la décision la plus importante à prendre lors de l'élaboration d'un schéma de classification puisque c'est elle qui détermine la macro-structure (ou structure de base) et décrit le domaine à classer (Kwasnik, 1999). Nous avons vu que la caractéristique de division permet de diviser une classe en sous-classes et d'organiser les connaissances en classes logiques. Cependant, même les documents les plus simples ont plusieurs caractéristiques (par exemple, le type de document ou le thème représenté) ce qui oblige à devoir arbitrairement privilégier comme base de classification, par exemple, tantôt le thème du document, tantôt le type de document, tantôt l'activité dans laquelle il s'insère, alors que le document pourrait de fait être classifié à plusieurs endroits. Ainsi, les possibilités de division peuvent être infinies et les mêmes objets ou idées peuvent être classifiés de différentes manières (Langridge, 1992; Mai, 2000).

Par conséquent, le choix des caractéristiques de division doit être basé sur deux impératifs : Premièrement, une seule caractéristique (par exemple, domaine de connaissance) devrait être utilisée à chaque étape de la division en vue d'assurer l'exclusivité des classes (i.e., classes distinctes les unes des autres) (Iyer, 1995; Kwasnik, 1999; Bowker & Star, 2000). Deuxièmement, les caractéristiques de division doivent être pertinentes à l'objectif de la classification (Iyer, 1995). Comme la logique ne permet pas de choisir les caractéristiques des espèces définies, c'est la connaissance des sujets et le but de la classification qui en guide le choix (Langridge, 1992).

Le choix de l'ordre de succession des classes de premier niveau est, en théorie, relativement indifférent puisque ce qui importe c'est que ces classes ne se chevauchent pas, dans le respect du principe de division logique (Maniez, 1987). Toutefois, l'ordre de succession des classes est généralement basé sur des principes, des modèles et une philosophie sous-jacents qui peuvent varier d'une classification à l'autre. Dans la classification Dewey, par exemple, les disciplines qui sont basées, selon Bacon (cité par

Iyer, 1995), sur la « raison » (Philosophie, Religion, Sciences sociales, etc.) apparaissent au début du schéma de classification, suivies des disciplines basées sur « l'imagination » (Beaux-Arts, Littérature) et la « mémoire » (Histoire et géographie).

Au niveau de la micro-structure

Comme une seule caractéristique de division n'est pas suffisante pour diviser un domaine de la connaissance, des sous-classes sont nécessaires. Quel que soit le type de structure choisi, il est important de décider de l'ordre dans lequel les règles de différenciation doivent être invoquées. Au niveau de la micro-structure (ou structure des classes subordonnées), il n'existe pas de règle absolue pour déterminer de manière significative la séquence des caractéristiques (Bowker & Star, 2000). Toutefois, quel que soit l'ordre des caractéristiques choisi, celui-ci doit être suivi de façon constante afin d'assurer l'homogénéité et la prévisibilité de la classification (Iyer, 1995).

Plusieurs principes ou conventions visent à guider le choix de l'ordre des caractéristiques. D'après le principe de majorité (« *majority principle* »), il faut concevoir un schéma qui convienne aux attentes de la majorité des utilisateurs ou qui repose sur la tradition (Iyer, 1995; Kwasnik, 1999; Fallis & Mathiesen, 2000). Ce principe permet de s'assurer que le schéma de classification sera compréhensible et utile pour la majorité des utilisateurs.

Un des impératifs reliés à ce principe préconise que l'ordre des caractéristiques doit refléter la vue contemporaine qui prévaut. D'emblée, le principe du « général au spécifique » permet de répondre généralement aux attentes des utilisateurs et fournit un élément de cohérence important (Langridge, 1992). De même, l'utilisation d'une terminologie familière et adaptée à la majorité des utilisateurs permet de s'assurer que le schéma sera utilisé efficacement.

Selon le principe d'exhaustivité, toutes les possibilités doivent être énumérées à chaque étape de la division (Langridge, 1992). L'exhaustivité suppose la certitude concernant ce qui existe déjà et la prédiction des découvertes et des créations futures

(Langridge, 1992). Toutefois, en raison de la nature dynamique de la connaissance, il n'est généralement pas possible pour le concepteur d'un schéma de classification de connaître ou de prédire toutes les possibilités.

Choix de la notation

La notation est un ensemble de symboles (par exemple, chiffres, lettres ou symboles typographiques) qui permet une représentation synthétique et abstraite des sujets dans un schéma de classification (Iyer, 1995; Maniez, 1987). Alors que la notation n'est pas indispensable dans une classification des connaissances, elle joue un rôle essentiel pour l'organisation physique des documents sur les étagères, dans les classeurs et comme clé d'accès dans les catalogues en ligne.

Les règles de construction de cet élément traditionnellement important d'un schéma de classification documentaire sont nombreuses. Nous résumerons ici les qualités qui sont mentionnées de manière récurrente dans la littérature (Canonne, 1993; Iyer, 1995; Maniez, 1987; Maniez, 1991) :

1. la notation doit utiliser des symboles familiers à l'utilisateur et si possible faciles à mémoriser, à lire et à prononcer;
2. la notation doit être logique c'est-à-dire refléter l'ordre de la classification, et expressive, c'est-à-dire exprimer le niveau de spécificité des rubriques;
3. la notation doit être suffisamment flexible pour permettre l'expansion du schéma par l'intégration de nouveaux concepts.

Mise à jour

Le schéma de classification hiérarchique spécialisé doit refléter le fond documentaire et le domaine à classer. Comme celui-ci évolue le schéma doit être régulièrement révisé afin de permettre de supprimer ou d'intégrer de nouvelles classes. Ce processus de mise à jour permet d'assurer l'utilisation et la survie d'une classification (Canonne, 1993). Ainsi, les structures des classifications encyclopédiques comme celles

de la Classification décimale de Dewey et de la bibliothèque du Congrès telles que développées à la fin du XIX^{ème} siècle ne reflèteraient plus les connaissances actuelles si elles n'étaient régulièrement mises à jour. Pour continuer à être utilisées, ces classifications, et plus précisément leur micro-structure, sont modifiées et adaptées selon les besoins, tout en évitant de bouleverser la macro-structure de classification déjà existante (De Grolier, 1988).

En résumé, la classification documentaire est un acte mental qui consiste à regrouper des entités en classes sur la base de caractéristiques communes. Plusieurs principes classificatoires sont à la base de la conception d'un système efficace de classification hiérarchique des documents. L'application du principe de division logique permet de représenter méthodiquement les connaissances et d'amener progressivement et intuitivement le chercheur vers les documents susceptibles de contenir l'information recherchée. La garantie d'existence bibliographique, quant à elle, consiste à tenir compte du nombre de documents offerts sur chaque sujet pour s'assurer d'élaborer un schéma de classification spécifique à l'ensemble documentaire.

Après avoir présenté ce qu'est la classification documentaire et étudié les caractéristiques d'un schéma de classification documentaire, il convient maintenant de s'intéresser à la notion de classification telle qu'elle est comprise et pratiquée dans les milieux administratifs et telle qu'elle est décrite dans les écrits spécialisés en archivistique.

2.3 Principes de classification des documents administratifs en archivistique

Si on constate la présence de multiples principes théoriques classificatoires dans un contexte documentaire général, on parle plutôt d'un « ensemble de conventions, de méthodes et de règles de procédures » dans le discours archivistique relatif à la classification des documents administratifs (Couture, 1999, p. 18; Boutet, 2000). Les principes classificatoires en archivistique sont généralement formulés avec moins de

rigueur intellectuelle que les principes théoriques classificatoires dans un contexte documentaire général en raison des lacunes de la recherche dans le domaine de la classification des documents administratifs. C'est surtout à partir d'une approche empirique que les archivistes ont su discerner des principes ou méthodes de classification qui leur sont propres. Comme la recherche dans le domaine de la classification des documents administratifs est encore embryonnaire, ces principes et méthodes classificatoires sont encore en développement dans la communauté archivistique.

L'objectif de cette section est de présenter, à partir d'une recension des écrits spécialisés en archivistique, la particularité de la classification et des schémas de classification institutionnels des documents administratifs dans un contexte de gestion centralisée des ressources informationnelles. Nous rappelons d'abord la nature et les fonctions de la classification des documents administratifs (section 2.3.1). Nous présentons ensuite brièvement la nature de l'objet de la classification dans un contexte administratif. L'analyse de la nature particulière des documents administratifs permet de comprendre l'importance de préserver le contexte de création de ces documents et l'emploi de techniques propres pour les classer (section 2.3.2). Enfin, nous abordons de manière plus approfondie les règles et principes d'élaboration des schémas de classification institutionnels des documents administratifs (section 2.3.2).

2.3.1 Nature et fonctions de la classification des documents administratifs

Cette section définit la nature de la classification des documents administratifs, aborde certaines considérations terminologiques et dégage les principales fonctions de la classification des documents administratifs.

2.3.1.1 La classification des documents administratifs

Il existe de nombreuses définitions de la fonction « classification » en archivistique (Schellenberg, 1956; Evans, Harrison & Thompson, 1974; Delsalle, 1996; Couture, 1999). Dans le cadre de notre recherche, nous retiendrons la définition proposée par Couture (1999) qui décrit la classification comme « un processus intellectuel

d'identification et de regroupement systématique d'articles¹ semblables, d'après des caractéristiques communes pouvant faire par la suite l'objet d'une différenciation si la quantité l'exige » (Couture, 1999, p. 18).

2.3.1.2 Considérations terminologiques

La nature polysémique du concept de classification, déjà observée pour la classification documentaire est également récurrente en archivistique. On constate que la notion de classification a été souvent désignée en archivistique par le terme « classement » (Rousseau, 1980; Delsalle, 1996; Duchain, 1998). Le classement et la classification sont pourtant deux opérations logiquement indépendantes puisqu'un système de classement peut être sans coïncidence avec la classification (Maniez, 1987).

Le terme « classification » dans les écrits de la discipline archivistique peut également désigner l'opération intellectuelle par laquelle le personnel administratif analyse les documents en vue d'en déterminer le sujet, pour ensuite les associer à un système de classification prédéfini (Evans, Harrison & Thompson, 1974; Casault, 1995). Vue dans ce sens, la classification désigne une opération intellectuelle qui vise l'identification et l'intégration de tout document à l'ensemble des documents déjà classifiés.

Par ailleurs, la notion de classification est utilisée aussi pour désigner de manière plus pragmatique le système de codification ou la procédure administrative selon laquelle des documents sont déclarés « secrets » ou « non communicables » pour des raisons de sécurité nationale ou de confidentialité des documents (Direction des archives de France, 1993; Faillès, 1997; Héon, 1999; Kaiser, 2003).

Afin d'éviter toute confusion terminologique et conceptuelle, nous utilisons dans le cadre de notre recherche le terme « classification » pour désigner le processus intellectuel de regroupement des documents administratifs en classes et le terme

¹ Dans le milieu des archives, l'article désigne l'unité matérielle de rangement (chemises de classement, boîtes, classeurs, registres, etc.).

« classement » pour désigner l'opération matérielle de rangement sur des étagères ou de sauvegarde dans des répertoires électroniques (Bellot, 2000).

2.3.1.3 Fonctions de la classification des documents administratifs

Suite à l'accroissement du nombre de documents dans les organismes, au développement des principes de base de la discipline archivistique (par exemple, le principe de provenance), à l'apparition d'exigences légales propres à la gestion des documents et à la reconnaissance des particularités de la gestion de l'information administrative, la classification des documents administratifs est apparue comme un processus susceptible de remplir diverses fonctions administratives (Couture & Rousseau, 1982; Roberge, 1985; Héon, 1995; Gagnon-Arguin & Bannouri, 1998; Héon, 1999; Roberge, 2002; Roberge, 2004).

La classification est indispensable à une gestion efficace des documents administratifs et fait partie intégrante d'un système de gestion documentaire (Roberge, 2004). Les ressources informationnelles étant le « cœur fonctionnel des organismes » (Héon, 1999, p. 223) il est important que ceux-ci se dotent du système de gestion de ces ressources qui soit le plus performant.

La classification des documents administratifs permet aux organismes du secteur public de respecter la législation en vigueur et au public d'exercer son droit d'accès à l'information (Héon, 1995). Ainsi, depuis 1982, la législation québécoise oblige les organismes publics à classer leurs documents « de manière à en permettre le repérage » et à « établir et tenir à jour une liste de classement² indiquant l'ordre selon lequel les documents sont classés » (*Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels et d'autres dispositions législatives*, L.R.Q., c. A-2.1, art. 16).

² Le terme « classement » correspond ici à la « classification » appliqué aux documents administratifs et signifie l'organisation systématique de ces documents selon un schéma conçu pour les rendre accessibles en vue d'une utilisation courante (Schellenberg, 1956).

Cette obligation légale est renforcée par la *Loi sur les archives* (L.R.Q., A-21.1), adoptée en 1983, qui met la gestion des documents actifs sous la responsabilité des Archives nationales du Québec. Ce contexte juridique confirme la nécessité d'appliquer des méthodes de contrôle plus rigoureuses concernant le classement et oblige l'archiviste à normaliser sa pratique en matière de classification (Héon, 1995).

Enfin, une autre des fonctions de la classification des documents administratifs dans un contexte de gestion centralisée est de respecter le premier degré du principe de respect des fonds qui est de « s'assurer que tous les documents sont rattachés à leur fonds d'origine » (Couture & Rousseau, 1982; Héon, 1999, p. 221). En archivistique, le principe de respect des fonds ou principe de provenance, en exigeant le regroupement des documents et des dossiers représentant les fonctions et activités de l'organisme, permet de préserver le contexte administratif de création d'un document et assure une meilleure interprétation, authenticité et valeur probante de l'information contenue dans ces documents (Duchain, 1977; MacNeil, 1992; Lambert, 1994; Chabin, 1997; Krawczyk, 1999). Par conséquent, la classification est à la base de la constitution de la première unité archivistique : le fonds d'archives d'un organisme (Gagnon-Arguin & Bannouri, 1998; Héon, 1999).

2.3.2 Objet de la classification des documents administratifs

Plusieurs chercheurs se sont intéressés à l'objet spécifique de la classification des documents administratifs (Schellenberg, 1956; Schellenberg, 1965; Duchain, 1977; Roberge, 1985; Bearman, 1989; Héon, 1995; Bearman, 1997; Duchain, 1998; Héon, 1999; Roberge, 2002; Roberge, 2004). Alors que la classification documentaire est une classification de documents sur la base de la similitude de leur *sujet* ou de la *connaissance* véhiculée (par exemple, les ressources humaines), la classification des documents administratifs est une classification de documents en fonction de l'*activité* dans le cadre de laquelle ils ont été créés ou reçus (par exemple, la *gestion* des ressources humaines). Autrement dit, la classification des documents administratifs ne s'intéresse pas tant au contenu ou sujet véhiculé par les documents qu'aux circonstances originales de création, de réception ou d'utilisation dans lesquelles ils s'insèrent (Bearman, 1989;

Bearman, 1997; Guercio, 2001). Par conséquent, la classification effectuée dans un contexte administratif est spécifique à chaque organisme même si on peut y retrouver des activités communes à tous les organismes (par exemple, la gestion des ressources financières ou la gestion des ressources humaines).

Pour expliquer cette distinction fondamentale entre l'objet de la classification documentaire et celui de la classification des documents administratifs, les chercheurs se sont traditionnellement appuyés sur les particularités de ce dernier (Schellenberg, 1956; Schellenberg, 1965; Roberge, 1985; Bearman, 1997; Guercio, 2001; Roberge, 2002; Roberge, 2004). Les documents administratifs se caractérisent par le fait d'avoir été créés ou reçus pour accomplir un but (l'accomplissement des affaires officielles, par exemple), directement en relation avec les activités fonctionnelles régulières d'un organisme (Schellenberg, 1956; Roberge, 1985; Bertrand-Gastaldy, 1990; Sénécal, 1997-1998; Roberge 2002; Roberge, 2004). Leur signification dépend de leur relation organique et de leur relation entre eux. La préservation de cette signification est importante en archivistique où les documents n'ont pas seulement une valeur documentaire mais également, et surtout, une valeur administrative, légale ou financière. Cette valeur ne peut être déterminée qu'en sachant d'où proviennent les documents, qu'en conservant leur contexte de création et d'utilisation. Ainsi, en établissant le lien entre les documents et une activité, un projet ou une affaire, la classification possède une visée pragmatique double, celle d'organiser les documents et de leur donner un sens (Gagnon-Arguin & Bannouri, 1998).

2.3.3 Outil de classification des documents administratifs

Cette section aborde certaines considérations terminologiques, définit la nature et les principales fonctions de l'outil de classification des documents administratifs, décrit le modèle structurel le plus utilisé pour classer les documents dans un contexte administratif et présente les principes et méthodes d'élaboration des schémas de classification des documents administratifs.

2.3.3.1 Considérations terminologiques

L'outil permettant la représentation de la classification des documents administratifs possède plusieurs désignations. Certains archivistes utilisent les termes « plan de classification » (Leclanche, 1994; Casault, 1995; Archives nationales du Québec, 1996; Guercio, 2002), « plan de classement » (Éthier, 1989; International Organization for Standardization, 2001; Poivre, 2004), ou « schéma corporatif de classification » (Roberge, 2004). Dans le cadre de notre recherche nous avons choisi d'utiliser l'expression « schéma de classification » pour désigner l'outil de classification des documents administratifs en raison de son utilisation prépondérante dans la littérature scientifique récente (Jones, 2007). Un schéma de classification est à distinguer d'un « guide de classification » qui désigne non seulement un schéma de classification des documents administratifs mais aussi l'index et les procédures d'utilisation qui l'accompagnent.

On parle difficilement de schémas de classification à portée « universelle » en archivistique puisqu'ils ne sont pas destinés à tous les utilisateurs possibles dans le monde, ni à l'organisation et au repérage de tous les documents. Le schéma de classification des documents administratifs devant être adapté aux besoins de chaque organisme, on parle plutôt de schémas de classification « institutionnels », « corporatifs », « sectoriels » ou « spécialisés ». Bien sûr, il existe des schémas de classification « type » (par exemple, le *Plan de classification type des documents de gestion gouvernementaux* des Archives nationales du Québec) qui servent pour l'élaboration des schémas de classification communs à certains organismes de même nature. Ces « modèles de classification » sont lacunaires par rapport à l'ensemble de la production documentaire puisqu'ils ne présentent que les fonctions et activités de gestion communes à tous les organismes et ignorent les activités d'exploitation propres à la mission d'un organisme particulier (Casault, 1995; Héon, 1999).

2.3.3.2 Nature du schéma de classification des documents administratifs

Qu'est-ce qu'un schéma de classification des documents administratifs? La littérature spécialisée en archivistique en présente quelques définitions (Couture &

Rousseau, 1982; Roberge, 1985; Casault, 1995; Héon, 1995; Archives nationales du Québec, 1996; International Organization for Standardization, 2001; Benfell, 2002; Poivre, 2004; Roberge, 2004). Les Archives nationales du Québec (1996, p. 154) proposent la définition suivante : « Structure hiérarchique et logique permettant le classement et le repérage de pièces d'archives ou d'ensembles documentaires ». Poivre (2004) le définit comme une « description méthodique et structurée de manière hiérarchique de la production documentaire » (Poivre, 2004, p. 15). Pour Benfell (2002), il s'agit d'un cadre ou d'une structure conceptuelle pour l'organisation des documents. Selon Roberge (2004), un schéma de classification est un « sous-système corporatif pour l'identification, la classification et la codification des documents et des dossiers au moment où ceux-ci sont produits ou reçus » (Roberge, 2004, p. 8.8).

Dans le cadre de notre recherche, nous retiendrons que le schéma de classification des documents administratifs est une représentation structurée des fonctions et activités administratives courantes d'un organisme réparties en classes, sous-classes et organisées hiérarchiquement. Nous considérons le schéma de classification comme une source de métadonnées descriptives extrinsèque à un document puisqu'il en décrit le contenu conceptuel et documente son contexte de production et d'utilisation.

2.3.3.3 Fonctions du schéma de classification des documents administratifs

L'utilisation d'un schéma de classification des documents administratifs vise plusieurs fonctions et objectifs (Casault & Lévesque, 1997; Héon, 1999; International Organization for Standardization, 2001; Roberge, 2004). Casault & Lévesque (1997, p. 5) identifient six fonctions :

1. fournir une méthode uniforme de classification et de classement;
2. ordonner la masse documentaire;
3. faciliter et accélérer le classement et le repérage des documents;
4. accélérer la prise de décision;

5. assurer la continuité de l'administration;
6. diminuer l'impact d'un changement de personnel.

La plupart de ces fonctions sont reliées entre elles. Par exemple, en fournissant un cadre uniforme de classification et de classement, le schéma de classification des documents administratifs facilite les démarches de recherche et diminue l'impact d'un changement de personnel (Héon, 1999; Benfell, 2002).

Nous retiendrons, en raison de leurs effets directement observables, qu'un schéma de classification des documents administratifs possède deux fonctions principales. La fonction première du schéma de classification est de regrouper logiquement et physiquement les documents relatifs à une même activité ou à une même affaire en vue d'organiser l'information et d'en faciliter le repérage. La relation entre le schéma de classification et le repérage est décrite à la section 2.5.4 Schéma de classification et repérage.

Une deuxième fonction importante d'un schéma de classification des documents administratifs est de fournir le contexte de création des documents afin de garantir l'interprétation de l'information consignée. Selon Sénécal (1999), il s'agit d'une fonction « cognitive » de représentation de l'organisme dans son ensemble.

« Par ce moyen, le schéma de classification permet autant à la personne qui classe qu'à celle qui recherche un document de mettre en contexte cette documentation. C'est-à-dire qu'il permet un positionnement relatif de l'utilisateur au sein des différentes structures en jeu dans la représentation de l'action administrative : structure documentaire, activités, unités administratives, etc. » (Sénécal, 1999, p. 4).

Cette représentation de l'organisme s'appuie sur des catégories descriptives relatives à ses fonctions et activités administratives (Sénécal, 1999; International Organization for Standardization, 2001). Elle permet l'interprétation des documents par la reconnaissance du contexte de création et d'utilisation (et donc de révéler la valeur

associée au document) « par l'ensemble du personnel de l'organisation ou par des utilisateurs éventuels » (Gagnon-Arguin & Bannouri, 1998, p. 3).

2.3.3.4 Structure du schéma de classification des documents administratifs

La structure d'un schéma de classification des documents administratifs doit permettre de représenter l'organisation des fonctions et activités administratives d'un organisme ainsi que de classer physiquement des documents selon l'activité à laquelle ils se rattachent (Roberge, 2004).

Pour ces raisons, il est aisé de comprendre que le modèle structurel privilégié pour la classification des documents administratifs dans un contexte de gestion centralisée des ressources informationnelles soit de type hiérarchique parce qu'il est le plus simple à construire et à utiliser par le personnel administratif.

2.3.3.5 Principes et méthodes d'élaboration du schéma de classification des documents administratifs

La conception d'un schéma de classification des documents administratifs suit désormais un processus reconnu (International Organization for Standardization, 2001). La présente section porte sur les règles, normes, conventions et exigences généralement retenues dans le processus d'élaboration de schémas de classification des documents administratifs. Comme ces principes varient selon l'environnement technologique, nous les aborderons d'abord dans un contexte analogique traditionnel puis dans un contexte électronique.

Principes d'élaboration dans l'environnement analogique

Les méthodes de travail

Nous avons mentionné que le schéma de classification des documents administratifs est spécifique à chaque organisme et reflète l'ensemble de ses activités administratives (Éthier, 1989). De plus, le schéma de classification correspond à la réalité documentaire de l'organisme (Roberge, 2004). Par conséquent, on recommande d'élaborer le schéma de classification à partir de la connaissance la plus complète

possible des *activités* et des *documents* de l'organisme (Garon, 1992; Gagnon-Arguin & Bannouri, 1998).

Cette construction est basée sur la combinaison de deux approches déjà identifiées dans un contexte de classification documentaire (section 2.2.3.5 Principes et méthodes d'élaboration de schémas de classification documentaire hiérarchique) : l'approche déductive (connaissances des fonctions et activités administratives) et l'approche inductive (étude de l'ensemble documentaire à organiser) (Roberge, 1985; Héon, 1995). Ces deux approches permettent de recueillir des données sur le fonds documentaire de l'organisme et d'identifier les classes et sous-classes qui formeront la structure interne du schéma de classification (Garon, 1992; Héon, 1999).

Le processus d'élaboration du schéma de classification des documents administratifs

Identification et ordre des classes

Au niveau de la macro-structure

Héon (1995), Vigneau (1996) et Guercio (2001) identifient les différentes approches appliquées en archivistique pour définir la macro-structure (ou structure de base) du schéma de classification des documents administratifs.

Dans un premier temps, le schéma de classification fut basé sur les structures administratives ou organisationnelles de l'organisme, autrement dit sur son organigramme. L'intérêt de cette approche était sa facilité de conception et d'utilisation. Ce schéma de classification s'est révélé peu flexible en raison d'une grande instabilité provoquée par les nombreuses modifications de l'organigramme à travers le temps (Rousseau, Couture & collaborateurs, 1994 cité dans Gagnon-Arguin & Bannouri, 1998; Vigneau, 1996; Héon, 1995; Héon, 1999; Benfell, 2002).

L'approche préconisée par Jacques Renaud (1975) dans les années soixante-dix fait aujourd'hui l'unanimité. Renaud fut le premier québécois à distinguer les fonctions

de gestion³ communes à la plupart des organismes des fonctions d'exploitation⁴ spécifiques à chaque organisme et à définir les classes de premier niveau à partir de ces « grandes fonctions » (Héon, 1995). Les classifications actuelles reposent encore sur les fonctions de l'organisme (Doyle, 1990; Gagnon-Arguin & Bannouri, 1998; International Organization for Standardization, 2001; Sabourin, 2001; Roberge, 2004). La relative stabilité des fonctions et activités d'un organisme assure une certaine pérennité du schéma de classification (Garon, 1992; Benfell, 2002). Cette pratique permet également de minimiser les répercussions que les fréquents changements organisationnels pourraient avoir sur le schéma de classification puisque la mise à jour se fait uniquement à partir de l'ajout de nouvelles fonctions et activités et non plus à chaque modification de l'organigramme (Doyle, 1990).

En ce qui concerne l'ordre des classes principales, il est sans importance puisqu'il n'affecte en rien la logique du schéma de classification. Toutefois, la consultation de plusieurs schémas de classification des documents administratifs gouvernementaux dans le cadre d'une précédente étude (Hudon & Mas, 2001) nous permet d'affirmer que la coutume consiste à disposer les classes relatives aux fonctions de gestion (par exemple, la gestion des ressources financières) avant les classes relatives aux fonctions d'exploitation (par exemple, la gestion de l'enseignement).

Au niveau de la micro-structure

À l'intérieur de chacune des classes principales, le schéma de classification représente dans un ordre logique, alphabétique, chronologique ou autre toutes les activités et sous-activités constitutives d'une fonction administrative.

Exemple : Gestion des ressources humaines (classe principale)
 Embauche
 Conditions de travail
 Etc.

³ Exemples de fonctions de gestion : organisation et management, législation et affaires juridiques, gestion des ressources humaines, gestion des ressources financières, etc.

⁴ Exemple de fonctions d'exploitation : Production, Enseignement, etc.

À tous les niveaux de division, il est possible d'augmenter la précision de certaines classes par l'ajout d'autres classes subordonnées.

Exemple : Engagement
 Engagement de personnel occasionnel
 Engagement de contractuels
 Engagement d'étudiants

Comme le constate Vigneau (1996), il n'y a pas de consensus établi sur les niveaux inférieurs (secondaires et tertiaires) du schéma de classification puisque les subdivisions sont basées, par exemple, sur des processus d'affaires, des catégories d'employés, des sujets, des types de document, ou sur la provenance.

Le principe théorique de division logique de la classification documentaire qui commande qu'une seule caractéristique de division soit rigoureusement appliquée à chaque niveau pour faciliter l'ajout de nouvelles unités de classification à quelque niveau que ce soit n'est pas toujours respecté (Hudon & Mas, 2001). Par exemple, on peut observer à un même niveau hiérarchique des classes représentant des activités, des sujets ou des types de documents comme dans l'exemple suivant :

Ressources humaines		
	Engagement	(Activité)
	Conditions de travail	(Sujet)
	Dossiers de personnel	(Type de document)

Nous étudierons dans la section 2.5.4 Schéma de classification et repérage l'influence de ces pratiques sur le repérage des documents administratifs.

Choix de la notation

La notation permet de regrouper physiquement et intellectuellement les documents administratifs ayant une relation entre eux et se rattachant à une activité

administrative commune. La notation est propre à chaque organisme, qui est libre de choisir son modèle (par exemple, numérique, décimal, alphabétique ou alphanumérique) (Roberge, 1985; Daum, 1997). L'examen de quelques schémas de classification des documents administratifs montre l'utilisation privilégiée d'une notation de type numérique (Hudon & Mas, 2001). Généralement, c'est la simplicité qui prévaut dans le choix de la notation pour les schémas de classification des documents administratifs afin que la notation puisse être aisément utilisée par le personnel de l'organisme. Enfin, quel que soit le choix de la notation, la cote doit être la plus significative possible, c'est-à-dire correspondre au découpage logique des activités du schéma de classification et respecter les niveaux hiérarchiques (Éthier, 1989; Héon, 1995).

Mise à jour du schéma de classification des documents administratifs

Les schémas de classification des documents administratifs doivent, comme tout schéma de classification documentaire, refléter le fond documentaire et le domaine à classifier. Les schémas de classification des documents administratifs doivent être mis à jour lors de l'ajout ou du retrait d'activités administratives ou, plus rarement, de fonctions administratives (Héon, 1999; Roberge, 2004). Cette mise à jour du schéma de classification peut entraîner des modifications de la micro-structure mais aussi de la macro-structure, plus rares mais possibles lors d'un changement dans la mission de l'organisme par exemple.

Principes d'élaboration dans l'environnement électronique

Approche générale

L'utilisation d'un seul schéma de classification est recommandée par les gestionnaires de documents et les institutions normatives nationales ou internationales pour organiser les documents administratifs dans l'environnement analogique et électronique. L'application de cette recommandation vise à simplifier la tâche des personnes responsables de la classification et à limiter le nombre de systèmes de classification à maîtriser (Vigneau, 1996; International Organization for Standardization, 2001; Benfell, 2002; Tough & Moss, 2003; Conférence des recteurs et des principaux

des universités du Québec, 2004; Poivre, 2004). Bien que l'utilisation d'un schéma de classification unique soit théoriquement possible dans n'importe quel environnement (analogique ou électronique), certaines caractéristiques du schéma de classification originellement conçu pour organiser les documents sur support analogique sont souvent modifiées pour être adaptées à l'environnement électronique. La section suivante présente le processus d'élaboration d'un schéma de classification adapté à l'organisation des documents électroniques.

Le processus d'élaboration du schéma de classification des documents administratifs électroniques

Identification et ordre des classes

Les systèmes d'exploitation informatiques utilisés dans les organismes, par exemple Windows ou Unix, se prêtent très bien à l'implantation du schéma de classification comme outil d'organisation des documents administratifs électroniques. La structure arborescente des répertoires du système de gestion des fichiers (« *filesystem* ») permet la migration relativement facile de la structure logique et hiérarchique du schéma de classification des documents administratifs (Poivre, 2004), la dénomination des répertoires et sous-répertoires pouvant correspondre aux intitulés des différentes classes et sous-classes du schéma de classification (Dhérent & collaborateurs, 2002; Tough & Moss, 2003; Roberge, 2004).

La structure du schéma de classification des documents administratifs électroniques est généralement moins profonde que la structure originale conçue pour organiser les documents administratifs sur support analogique (Sabourin, 2001). Beaupré (2004) et Chamard (2004) mentionnent l'utilisation, au sein de leur organisme, de schémas de classification des documents administratifs électroniques réduits à trois ou quatre niveaux correspondant à la réalité de l'ensemble documentaire électronique à organiser.

Choix de la notation et des énoncés

L'utilisation de la notation et des énoncés des classes du schéma de classification conçus pour organiser les documents dans l'environnement analogique est également encouragée pour nommer les répertoires électroniques correspondants aux niveaux supérieurs de la hiérarchie (fonctions, activités, sous-activités administratives) (Poivre, 2004). Comme l'indique l'exemple de la Figure 2.1, la dénomination des répertoires dans l'environnement électronique est le plus souvent une adaptation plutôt qu'une reproduction strictement fidèle des énoncés des classes du schéma de classification utilisé pour organiser et classer les documents sur support papier.

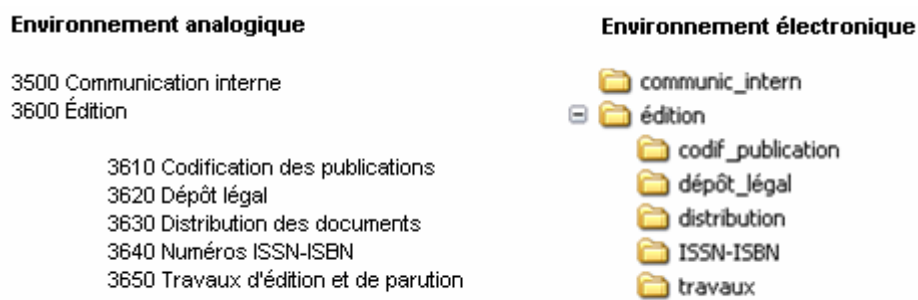


Figure 2.1 Exemple de dénomination des classes dans un environnement analogique et dans un environnement électronique.

Dans certains organismes, la notation est supprimée dans l'environnement électronique et les noms des répertoires continuent d'être limités à huit caractères afin d'offrir une meilleure ergonomie ou pour assurer la compatibilité entre les différents systèmes d'exploitation ou les différentes versions de logiciels utilisés (Chamard, 2004).

Mise à jour

Il n'existe pas d'écrits sur la mise à jour du schéma de classification des documents administratifs électroniques. On présume que la pratique est la même que dans l'environnement analogique : la mise à jour du schéma de classification des

documents administratifs électroniques dans un contexte de gestion centralisée est réalisée par un archiviste ou un gestionnaire de documents lors de l'ajout ou du retrait de fonctions ou d'activités administratives.

En résumé, la classification des documents administratifs ne s'intéresse pas tant au contenu ou au sujet véhiculé par les documents qu'aux circonstances originales de création, de réception ou d'utilisation dans lesquelles ces documents s'insèrent. En établissant un lien entre un document et une activité, autrement dit en conservant le contexte de création et d'utilisation du document, la classification permet de donner un sens et de déterminer la valeur des documents administratifs. En raison de cette fonction particulière, la classification des documents administratifs est spécifique à chaque organisme même si elle peut se baser en partie sur des fonctions et activités administratives communes à tous les organismes (par exemple, la gestion des ressources humaines). Cela dit, quelques principes d'élaboration d'un schéma de classification dans un contexte de gestion centralisée des documents administratifs sont communs aux règles de conception des schémas de classification documentaires; l'élaboration des schémas de classification des documents administratifs est basée à la fois sur une approche déductive (connaissance des fonctions et activités administratives) et inductive (étude de l'ensemble documentaire à organiser).

Après avoir présenté ce qu'est la classification des documents administratifs et étudié l'outil dans un contexte de gestion centralisée, il convient maintenant de s'intéresser à la notion d'organisation personnelle ou décentralisée des documents électroniques telle qu'elle est pratiquée par les individus au sein des organismes.

2.4 Organisation personnelle des documents administratifs électroniques

L'organisation personnelle des documents électroniques s'inscrit dans le concept plus large de gestion personnelle ou décentralisée de l'information (« *Personal Information Management* »), utilisé pour décrire la création, l'acquisition, l'entreposage, l'organisation et le repérage d'objets numériques (par exemple, les documents, les courriels, et les favoris Web) par un individu dans son environnement informatique personnel (Lansdale, 1988; Barreau, 1995; Whittaker & Hirschberg, 2001; Boardman & Sasse, 2004; Jones, 2004).

L'organisation personnelle des documents électroniques en milieu de travail est effectuée par des millions d'individus, plusieurs fois par jour (Whittaker, Terveen & Nardi, 2000). Elle est fréquemment considérée comme un fardeau (Malone, 1983; Lansdale, 1988; Whittaker & Sidner, 1996) et une source de problèmes et de frustrations (Boardman & Sasse, 2004). Si certains individus qui ont eu recours au schéma de classification des documents administratifs pour organiser leurs documents sur support électronique sont satisfaits des résultats, d'autres semblent éprouver des difficultés et préfèrent recourir à des stratégies personnelles d'organisation des documents plus adaptées à leurs besoins (Gagnon-Arguin & Bannouri, 1998).

L'objectif de cette section est de présenter, à partir d'une recension des écrits spécialisés dans le domaine de la gestion personnelle de l'information, les caractéristiques particulières de l'organisation personnelle des documents administratifs électroniques. Nous présentons d'abord la nature et les fonctions de l'organisation personnelle des documents administratifs électroniques. Nous décrivons ensuite brièvement la nature de l'objet à organiser avant d'aborder de manière plus approfondie les pratiques d'élaboration des schémas de classification personnels des documents administratifs.

2.4.1 Nature et fonction de l'organisation personnelle des documents administratifs électroniques

Cette section définit la nature de l'organisation personnelle des documents administratifs électroniques et dégage ses principales fonctions du point de vue des individus.

2.4.1.1 L'organisation personnelle des documents administratifs électroniques

L'organisation personnelle des documents administratifs électroniques est définie comme étant tout processus mis en place par un individu pour faciliter l'organisation et le repérage des documents qui sont sous son contrôle direct en vue de répondre à ses besoins personnels (Lin & Chan, 1999; Beyth-Marom & Nachmias, 2003; Boardman & Sasse, 2004).

Le concept d'organisation « personnelle » ou « décentralisée » des documents administratifs électroniques est donc distinct de celui d'organisation « institutionnelle » ou « centralisée » des documents administratifs électroniques abordé dans la section 2.3 Principes de classification des documents administratifs en archivistique. Dans un contexte de gestion centralisée, c'est un professionnel en gestion documentaire qui conçoit et contrôle un système d'organisation des documents pour une variété d'individus ayant des besoins différents à l'intérieur d'un même organisme ou d'une même unité (Bergman, Beyth-Marom & Nachmias, 2003).

2.4.1.2 Fonctions de l'organisation personnelle des documents administratifs électroniques

L'organisation personnelle des documents électroniques remplit trois fonctions principales. Elle permet de :

1. mettre en ordre de grandes quantités de documents qui peuvent être réutilisés ultérieurement, pour des besoins de référence par exemple (Barreau, 1995; Whittaker & Sidner, 1996; Bergman, Beyth-Marom & Nachmias, 2003; Henderson, 2003);

2. accéder très rapidement aux documents d'utilisation courante en les ayant « à la portée de la main » (Hertzum, 1999);
3. rappeler des choses urgentes à faire (Malone, 1983; Hertzum, 1993; Kidd, 1994; Barreau & Nardi, 1995; Whittaker & Sidner, 1996; Sellen & Harper, 2001).

La connaissance des principales fonctions sous-jacentes à l'organisation personnelle des documents administratifs est importante car ces fonctions déterminent les stratégies et les outils mis en place par les individus pour organiser les documents électroniques se trouvant sur leur poste de travail informatique (Kwasnik, 1989; Kwasnik, 1991; Barreau, 1995).

2.4.2 Objet de l'organisation personnelle des documents administratifs électroniques

Plusieurs chercheurs (Boardman & Sasse, 2004; Henderson, 2005; Khoo *et al.*, 2007) se sont intéressés à l'objet de l'organisation personnelle des documents administratifs électroniques, c'est-à-dire aux règles explicites ou implicites de regroupement et de différenciation des documents en classes. Ces chercheurs ont procédé à une analyse conceptuelle des répertoires utilisés pour organiser les documents électroniques se trouvant sur les postes de travail des individus. Les résultats de leur recherche nous apprennent que ce n'est ni le sujet ou la connaissance véhiculée par le document, ni uniquement les fonctions et activités administratives dans lesquelles il s'insère qui sont objet de l'organisation. La classification est plutôt basée sur des *projets*, des *types de documents* et des *tâches professionnelles* propres à l'individu. La présentation détaillée de ces critères de division est réalisée dans la section 2.6.2 Indicateurs de mesure de la dimension logique des schémas de classification.

2.4.3 Outil d'organisation personnelle des documents administratifs électroniques

Cette section aborde certaines considérations terminologiques et définit la nature, les principales fonctions et le modèle structurel du schéma de classification utilisé par les

individus pour organiser leurs documents électroniques dans un contexte de travail. Pour terminer, les pratiques d'organisation personnelle des documents administratifs électroniques recensées dans les écrits sont présentées.

2.4.3.1 Considérations terminologiques

L'outil communément utilisé par les individus pour organiser les documents administratifs électroniques qui sont sous leur contrôle direct possède des appellations variées : « système hiérarchique de fichiers » ou « structure organisationnelle » (Dumais *et al.*, 2003), « hiérarchie » ou « hiérarchie sémantique » (Abrams, Baecker, & Chignell, 1998) « système de classement personnel » (Lansdale, 1988), « taxonomie personnelle » (Jones, 2004) ou « schéma de classification personnel » (Jones & Maier, 2003; Ferraioli, 2005).

Un schéma de classification personnel des documents administratifs désigne un arrangement hiérarchique de catégories conceptuelles permettant le regroupement de documents selon des règles d'application souvent implicites, relatives par exemple à des projets, à des activités ou à des types de documents (Boardman & Sasse, 2004). L'organisation de ces catégories conceptuelles est qualifiée de « personnelle » car elle se base sur la représentation du monde que se fait une personne et vise à répondre à ses propres besoins (Jones & Maier, 2003; Jones, 2004).

2.4.3.2 Nature du schéma de classification personnel des documents administratifs électroniques

Dans le cadre de leurs activités quotidiennes, les individus passent beaucoup de temps à interagir avec une interface technologique (logicielle et matérielle) également désignée dans la littérature sous le vocable d'espace de travail virtuel (« *virtual workspace* ») (Kaptelinin, 1996). Cette interface est basée sur la métaphore du bureau (« *desktop metaphor* »), c'est-à-dire qu'elle reproduit dans le système informatique l'organisation des documents telle qu'elle est pratiquée dans un bureau physique pour faciliter l'accessibilité au système par des individus ayant des compétences

informatiques variées (Fertig, Freeman & Gelernter, 1996; Nardi & Barreau, 1997; Ravasio, Schar & Krueger, 2004).

Dans le cadre de notre recherche, l'expression « schéma de classification personnel des documents administratifs électroniques » désigne la structure arborescente des répertoires et sous-répertoires électroniques utilisée par des individus pour organiser les documents électroniques qui font partie de leur espace documentaire personnel.

2.4.3.3 Fonctions du schéma de classification personnel des documents administratifs électroniques

Selon Henderson (2004), le schéma de classification personnel des documents administratifs électroniques est utilisé par les individus pour *organiser* et *repérer* visuellement (sur le bureau virtuel et à l'intérieur des répertoires), temporellement (par date de création ou modification) et logiquement (à partir des noms des répertoires et des fichiers) les documents qui forment leur espace documentaire personnel à l'aide d'une structure arborescente de répertoires et sous-répertoires (Henderson, 2004).

2.4.3.4 Structure du schéma de classification personnel des documents administratifs électroniques

Depuis plusieurs décennies, le « paradigme organisationnel » principalement proposé et utilisé dans l'interface technologique est hiérarchique (Henderson, 2003; Nelson, 2003). La structure hiérarchique possède des avantages cognitifs pour l'individu puisqu'elle lui permet de repérer un document en privilégiant la navigation locale à travers la structure arborescente de répertoires et la localisation spatiale du dossier dans lequel se trouve classé le document (Berlin *et al.*, 1993; Barreau & Nardi, 1995; Berchtold, Biliris & Panagos, 1999). Par ailleurs, la position du document dans la hiérarchie et l'obligation de nommer les répertoires et fichiers fournissent de l'information sémantique sur le document (Henderson, 2003).

Cette structure d'organisation hiérarchique qui apparaît idéale pour des fins d'archivage ne semble toutefois pas toujours comprise et utilisée par les individus

(Marsden & Cairns, 2003). Il n'a pas été non plus clairement démontré que l'utilisation d'une classification strictement hiérarchique est appropriée pour organiser les documents en raison de son caractère arbitraire (Barreau, 1995; Sénécal, 1999; Dourish *et al.*, 2000; Nelson, 2003; Henderson, 2004). Beaucoup de recherches ont été réalisées et continuent à être menées pour tenter de changer ce paradigme hiérarchique (par exemple, en proposant de classer les documents à l'aide de différents attributs) (Dourish *et al.*, 1999) et d'améliorer l'interface technologique de gestion personnelle de l'information (Freeman & Gelernter, 1996; Robertson *et al.*, 1998; Dourish *et al.*, 1999; Rekimoto, 1999; Gemmell *et al.*, 2002; Bellotti *et al.*, 2003; Bergman, Beyth-Marom & Nachmias, 2003; Boardman, Spence & Sasse, 2003; Dumais *et al.*, 2003; Kaptelinin, 2003).

2.4.3.5 Pratiques d'organisation personnelle

Nous décrivons successivement les approches générales d'organisation personnelle des documents, le processus proprement dit d'élaboration et la méthodologie pour la mise à jour régulière du schéma de classification personnel ainsi conçu.

Approches générales d'organisation personnelle des documents

Plusieurs études ont proposé des classifications des comportements des individus dans l'utilisation des différents outils de gestion personnelle de l'information, en mettant l'accent sur l'activité d'organisation (Barreau, 1995; Barreau & Nardi, 1995; Whittaker & Snider, 1996; Bälter, 1997; Abrams, Baecker, & Chignell, 1998; Chapman, 1999; Boardman, 2001a; Henderson, 2003; Ravasio, Schar & Krueger, 2004). Ces études ont été guidées par les travaux de Cole (1982) et Malone (1983) qui ont identifié deux stratégies fondamentales d'organisation personnelle des documents : l'« empilement » (« *piling* ») et le « classement » (« *filing* ») des documents. Des recherches ultérieures ont montré que ces deux stratégies observées dans l'environnement papier sont aussi valables dans l'environnement électronique (Barreau, 1995; Barreau & Nardi, 1995; Henderson, 2003; Boardman & Sasse, 2004).

L'empilement des documents (« *piling* ») consiste à regrouper, sans ordre particulier et sans titre explicite, au même endroit physique ou spatial, le plus souvent sous forme de piles, des documents n'ayant le plus souvent aucune relation entre eux (Cole, 1982). Ce comportement permet de se rappeler des choses urgentes à faire, d'accéder très rapidement aux documents en les ayant « à la portée de la main », ou bien de « déposer » temporairement sur le bureau (virtuel ou non) les documents reçus (Malone, 1983; Hertzum, 1993; Kidd, 1994; Barreau & Nardi, 1995; Whittaker & Sidner, 1996; Sellen & Harper, 2001; Ravasio, Schar & Krueger, 2004).

Le classement des documents (« *fling* ») consiste à fournir un effort régulier pour classer les documents à l'aide d'un schéma de classification personnel ou dans des piles de documents relativement bien ordonnées et identifiées (Malone, 1983; Whittaker & Snider, 1996). Le schéma de classification personnel peut être conçu selon une approche déductive (« *top-down* »), c'est-à-dire sur la base d'une branche du schéma de classification des documents administratifs institutionnel, ou à partir d'une approche inductive (« *bottom-up* »), c'est-à-dire de manière incrémentale et informelle, davantage ajustée aux besoins d'un individu (Abrams, Baecker, & Chignell, 1998). Dans un environnement de travail, cette structure classificatoire personnelle, qui est généralement peu profonde, présente des avantages cognitifs pour les employés. Selon Jones *et al.* (2005), cette pratique leur permet de conserver un certain contrôle (par exemple, s'assurer que tous les documents sont à une seule place) et d'établir des liens entre les documents.

Le choix de la stratégie de classification mise en place par les individus peut être influencé, entre autres, par la nature des tâches à réaliser et l'ancienneté dans l'emploi actuel (Kidd, 1994; Henderson, 2003). La nature des tâches à réaliser suggère qu'un individu peut avoir besoin d'accéder plus fréquemment à certains types d'information que d'autres, obligeant une organisation plus systématique de ses documents électroniques. L'ancienneté (ou les *années d'expérience*) dans le travail actuel est également considérée comme un facteur pouvant entraîner une accumulation d'une plus grande quantité de documents incitant l'employé à faire l'effort d'implanter un schéma

de classification plus élaboré (Lansdale, 1988; Barreau, 1995; Whittaker & Snider, 1996; Jones, Dumais & Bruce, 2002; Henderson, 2003). L'absence de politique et de formation en matière de gestion des documents est également un facteur qui encouragerait l'individualisme en matière de gestion décentralisée des documents (Gagnon-Arguin & Bannouri, 1998).

Processus d'élaboration d'un schéma de classification personnel

Les résultats exploratoires des recherches ayant porté sur les caractéristiques des schémas de classification personnels des documents administratifs électroniques peuvent être regroupés selon trois aspects principaux : (1) les critères de division et de regroupement logique des répertoires (ou classes) (Boardman & Sasse, 2004; Henderson, 2005; Khoo *et al.*, 2007), (2) la structure arborescente des répertoires (classes) et sous-répertoires (sous-classes) électroniques (Boardman, 2001a; Henderson, 2003; Gonçalves & Jorge, 2003; Boardman & Sasse, 2004), et (3) la dénomination des répertoires et des fichiers (Carroll, 1982; Chapman, 1999; Gonçalves & Jorge, 2003). Les sections suivantes présentent successivement ces trois aspects.

Les critères de division et de regroupement logique

Boardman & Sasse (2004), Henderson (2005) et Khoo *et al.* (2007) ont réalisé une analyse des concepts représentés dans la dénomination personnelle des répertoires électroniques. Les résultats de leurs travaux ont permis d'identifier 18 critères de division et de regroupement différents dont trois consensuels (i.e., *type de document*, *projet* et *thème*) (Voir Tableau 2.1).

Tableau 2.1 Principaux critères de division logique recensés dans les écrits

Critères de division logique	Boardman & Sasse (2004)	Henderson (2005)	Khoo <i>et al.</i> (2007)
Genre/Type de document	xxx	xxx	xxx
Tâche/Projet	xxx	xxx	x
Sujet/Thème	x	xxx	x
Temps		xxx	x
Activités/Rôle	xxx		
Inconnu		xxx	
Personne/Groupe/Organisation/Contact	x	x	x
Personnel/Confidentiel/Privé		x	x
Statut temporel (p. ex., temporaire)		x	xxx
Divers/ Général/ Multiples		x	x
Application/Format de fichier		x	x
Fonction/Structure organisationnelle			xxx
Région géographique			x
Version du document			x
Priorité			x
Données statistiques			x
Cours		x	
Source/Auteur du document		x	

xxx : Concepts les plus souvent représentés dans les noms des répertoires.

L'analyse conceptuelle de Boardman & Sasse (2004) a mis en évidence une classification essentiellement basée sur le *type de document* (par exemple, « Lettres »), le *projet* (par exemple, « *ucl presentation* ») et l'*activité à long terme* (par exemple, « Enseignement »). Les résultats des travaux de Henderson (2005) indiquent également une classification principalement par type de document ou *genre* (par exemple, « Rapport »), par projet ou *tâches* (par exemple, « PhD »), et par *thème* (par exemple, « JavaScript »). Khoo *et al.* (2007) ont quant à eux identifié 14 catégories différentes de regroupement des documents. Leurs résultats révèlent un regroupement par *tâche* et par *type de document*.

La structure arborescente des répertoires

Plusieurs chercheurs ont analysé la structure arborescente des répertoires utilisée par des individus dans un contexte de gestion personnelle des ressources

informationnelles (Gonçalves & Jorge, 2003; Boardman & Sasse, 2004; Henderson, 2004; Khoo *et al.*, 2007). Nous présentons leur conclusion sous forme de tableau (Voir Tableau 2.2).

Tableau 2.2 Principales caractéristiques de la structure arborescente des répertoires recensées dans les écrits

Caractéristiques de la structure arborescente	Gonçalves & Jorge (2003)	Boardman & Sasse (2004)	Henderson (2004)	Khoo <i>et al.</i> (2007)
Structure étendue (<i>broad</i>) (grand nombre de classes au premier niveau de la hiérarchie) et peu profonde (<i>shallow</i>) (petit nombre de niveaux hiérarchiques)		Moyenne de 3.3 niveaux hiérarchiques		> 19 classes principales Max. 3 niveaux hiérarchiques
Structure peu étendue (<i>narrow</i>) (petit nombre de classes au premier niveau de la hiérarchie, petit nombre moyen de classes à chaque niveau) et profonde (<i>deep</i>) (grand nombre de niveaux hiérarchiques)	Moyenne de 8.45 niveaux hiérarchique (écart-type=2.9) Nombre moyen de classes à chaque niveau = 1.84 (écart-type=0.187) (« <i>Branching Factor</i> »)		Moyenne entre 2.6 et 6.2 niveaux hiérarchiques (Max. entre 6 et 16 niveaux hiérarchiques)	< 10 classes principales Max. 8 niveaux hiérarchiques
Structure équilibrée (faible écart-type de la moyenne du nombre de classes à chaque niveau)	Écart-type de la moyenne du nombre de classes à chaque niveau = 3.61 (écart-type=0.67) (« <i>Tree Skewness</i> »)			
Complexité (nombre total de classes)		Entre 5 et 218 classes	Entre 211 et 7200 classes	

Ces recherches ont mis en évidence deux formes principales d'arborescence, soit des structures étendues et peu profondes et des structures peu étendues et profondes. Toutes ces structures apparaissent généralement équilibrées et plus ou moins complexes.

La dénomination des répertoires

Nous considérons la dénomination des répertoires électroniques comme une tâche organisationnelle importante car elle possède une fonction descriptive commune à tout énoncé des classes des schémas de classification. Les termes utilisés pour nommer les

répertoires peuvent être considérés comme des descripteurs et jouent par conséquent un rôle important au moment du repérage documentaire. Il apparaît que la dénomination des répertoires électroniques est le plus souvent une création originale de l'individu plutôt qu'une reproduction strictement fidèle ou, à tout le moins, une adaptation de l'énoncé des classes du schéma de classification institutionnel. Les raisons qui motivent cette personnalisation peuvent être d'ordre pragmatique (par exemple, sauvegarder rapidement son travail) ou sémantique (par exemple, décrire le contenu du répertoire en utilisant des termes ou abréviations jugés personnellement utiles) (Chapman, 1999; Jones, 2004).

La dénomination des répertoires électroniques, qui est le plus souvent hâtive, ne rend pas toujours compte du contenu d'un répertoire. Par exemple, Jones (2004) rapporte qu'un des participants à son étude possédait des dossiers intitulés « *stuff* », « *more stuff* » et « *still more stuff* ». Il s'agit d'un cas extrême mais comme nous le rappelle le chercheur : combien parmi nous avons des répertoires avec des noms semblables et une absence complète de définition? (Jones, 2004, p. 17).

La deuxième fonction de la dénomination des répertoires électroniques relevée par Chapman (1999) est de permettre le repérage et l'identification ultérieure des documents. Les répertoires sont généralement nommés par les individus pour répondre à leurs besoins personnels et pour refléter la signification de leur contenu (Carroll, 1982; Chapman, 1999; Ravasio, Schar & Krueger, 2004). Ainsi, un répertoire dont le nom n'est pas suffisamment significatif peut être renommé ultérieurement (Chapman, 1999). Un nom qui est significatif pour le créateur peut ne pas l'être pour d'autres personnes; une attention particulière est donc généralement apportée au nom des répertoires destinés à être communiqués ou partagés par d'autres personnes en vue d'augmenter la fonction descriptive des classes (Chapman, 1999, p. 52).

Mise à jour

Ravasio, Schar & Krueger (2004) rapportent que la mise à jour du système d'organisation des documents électroniques est une activité qui est réalisée, pour la majorité de leurs répondants, lors d'événements clés (par exemple, lorsqu'un projet est

terminé) ou lorsque le nombre de documents devient trop important. Cette mise à jour signifie le plus souvent le retrait de la structure classificatoire des répertoires électroniques qui sont considérés comme inutiles ou la création de nouveaux répertoires (Ravasio, Schar & Krueger, 2004).

Nous avons mentionné dans la section 2.4.3.5 Pratiques d'organisation personnelle que plusieurs recherches ont permis d'obtenir des données sur les facteurs qui influencent les décisions classificatoires des individus dans un environnement papier ou électronique (Kwasnik, 1989; Kwasnik, 1991; Nardi, Anderson & Erickson, 1994; Barreau, 1995; Kaptelinin, 1996; Chapman, 1999). En raison de ces différents facteurs d'influence, les stratégies classificatoires d'un individu sont flexibles et peuvent changer à travers le temps. L'individu va adapter sa stratégie à ses besoins ou à de nouvelles contraintes. Par exemple, Bälter (1997) observe des changements dans la stratégie de gestion des courriels, tels que l'augmentation (« *pro-organizing* ») ou la diminution (« *anti-organizing* ») dans la tendance à classer, encouragés par la surcharge de documents à organiser. Ce comportement a été également observé dans l'organisation personnelle des favoris (Abrams, Baecker, & Chignell, 1998).

En résumé, les schémas de classification mis en place par les individus dans un contexte de travail sont qualifiés de « personnels » car ils se basent sur une représentation subjective du monde et visent à répondre aux besoins propres d'un individu (Jones & Maier, 2003; Jones, 2004). Si les individus semblent pouvoir se retrouver plus ou moins efficacement dans leur propre schéma de classification, qu'en est-il de la capacité d'une tierce personne à comprendre et interpréter cette structure classificatoire qualifiée d'« hermétique » (Sénécal, 1999)? Autrement dit, quelle est la performance des schémas de classification personnels sur le plan du repérage de documents effectué par autrui?

Dans la section suivante, nous expliquons les raisons pour lesquelles nous considérons l'efficacité du repérage comme moyen de comparer les schémas de classification.

2.5 Repérage des documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles

Le repérage des documents électroniques est présenté comme une « tâche » qui fait elle-même partie de l'organisation personnelle de l'information et s'inscrit, tout comme l'organisation documentaire, dans le concept plus large de gestion personnelle ou décentralisée de l'information (Barreau, 1995; Whittaker & Hirschberg, 2001; Jones, 2004; Raviasso, Schar & Krueger, 2004).

Dans la présente section, nous présentons brièvement la spécificité de la nature et des fonctions du repérage appliqué aux documents administratifs électroniques. Nous décrivons ensuite la nature de l'objet du repérage avant d'aborder de manière plus approfondie les facteurs déterminant l'efficacité du repérage et la relation théorique qui existe entre schéma de classification et repérage.

2.5.1 Nature et fonction du repérage des documents administratifs électroniques dans un contexte de gestion documentaire décentralisée

Cette section définit la nature du repérage des documents administratifs et dégage ses principales fonctions.

2.5.1.1 Le repérage

Nous entendons par repérage d'un document électronique l'accès par la localisation de l'emplacement d'un document (i.e., l'identification du répertoire) à l'intérieur d'une structure arborescente de répertoires électroniques.

2.5.1.2 Fonction du repérage des documents

Le repérage des documents en milieu de travail vise à permettre aux individus d'utiliser et de réutiliser les documents conservés sur leur poste de travail ou sur celui de leurs collègues. Dans un contexte de croissance de l'économie du savoir et de travail collaboratif, le repérage peut être considéré comme un processus important car l'impossibilité de repérer un document implique, par exemple, une diminution de la

productivité des employés et de la compétitivité de l'organisme, le ralentissement ou la cessation des activités, voire la perte de la mémoire institutionnelle (Douglas, 1995; Héon, 1995; Hertzum, 1999).

Dans un contexte électronique, on pourrait penser que les conséquences liées à l'impossibilité de repérer un document peuvent être moindres en raison de la valeur informelle et officieuse d'une grande partie de l'information administrative consignée sur support électronique. Il apparaît en effet que les documents électroniques sont le plus souvent utilisés comme outils de production (brouillons, annotations, versions préparatoires) ou de reproduction de documents imprimés sur papier pour des besoins ergonomiques (par exemple, une consultation plus agréable ou pour permettre des annotations), juridiques (nécessitant une signature manuscrite) ou de sécurité (Hertzum, 1993; Hackett, 1997; Hertzum, 1999; Karjalainen *et al.*, 2000; Dhérent & collaborateurs, 2002). Il est cependant reconnu que certains documents possèdent une valeur administrative, financière ou légale pour les organismes et sont essentiels pour la poursuite de leurs activités (Berlin *et al.*, 1993) et la défense de leurs droits (Saulnier, 1997-1998). Parmi ces documents, la plupart ont également une valeur historique et font alors partie de la mémoire corporative (Dhérent & collaborateurs, 2002). Cette assertion est appuyée par le fait que de plus en plus de documents existent uniquement sur support électronique et sont soumis à des délais de conservation, d'où l'importance de pouvoir les repérer pour l'accomplissement des activités quotidiennes et le témoignage des activités passées. Aujourd'hui, la reconnaissance légale de l'utilité institutionnelle ou historique des documents électroniques dans plusieurs pays rend l'enjeu du repérage de ces documents électroniques tout aussi important que celui des documents dans un environnement analogique (Québec, 2001).

2.5.2 Objet du repérage des documents

Dans notre recherche, nous avons mentionné que la notion de repérage des documents désigne l'opération qui consiste à localiser un document cible (ou un ensemble défini de documents) à partir de l'organisation plus ou moins systématique d'un ensemble documentaire.

Tant dans l'environnement analogique que dans l'environnement électronique, quand on utilise l'expression « repérage des documents administratifs » ce n'est pas tant les documents que l'on repère mais le dossier auquel ces documents sont rattachés et qui se rapporte à une activité, à un événement ou à une personne (Conférence des recteurs et des principaux des universités du Québec, 1994). La place du dossier dans l'organisation et le repérage des documents est primordiale puisque c'est le dossier qui correspond au premier niveau d'organisation des documents administratifs et qui peut servir de base au repérage de l'information (Mas & Gagnon-Arguin, 2003-2004; Jones *et al.*, 2005).

2.5.3 Facteurs qui influencent la qualité du repérage

Nous nous basons sur la modélisation du processus de représentation et de repérage de l'information proposée par Chu (2003) et Henderson (2003) pour décrire les principaux facteurs d'influence de la qualité du repérage des documents électroniques dans un contexte de gestion décentralisée des ressources informationnelles. Chu (2003) identifie trois entités interreliées qui agissent conjointement dans le processus de représentation et de repérage de l'information : (1) des individus, (2) des systèmes et (3) de l'information. Dans notre recherche, le repérage manuel des documents électroniques dans un système de classification hiérarchique est considéré comme le produit de l'interrelation de ces trois éléments puisque nous demandons à un individu ou agent humain (1) d'utiliser un mécanisme de repérage donné (une structure arborescente de répertoires considérée comme un schéma de classification) (2) pour repérer des documents dans un espace documentaire plus ou moins structuré (3).

La section suivante décrit chacune des composantes de ce modèle.

2.5.3.1 Agent humain

Les recherches sur les pratiques de repérage observées dans le domaine de la gestion personnelle de l'information consignée sur support électronique ont permis de distinguer deux situations distinctes : (1) le repérage est effectué par le créateur et organisateur même de l'espace documentaire, et (2) le repérage est effectué par une tierce personne étrangère au contenu et à l'organisation de cet espace documentaire.

Alors que des recherches empiriques ont évalué le repérage de documents électroniques par le producteur même de l'espace documentaire (Barreau & Nardi, 1995; Nardi & Barreau, 1997; Chapman, 1999; Boardman & Sasse, 2004; Ravasio, Schar & Krueger, 2004), aucune ne s'est intéressée au repérage réalisé par une tierce personne.

Dans la situation où le repérage est réalisé par le créateur et organisateur de l'espace documentaire, Lansdale (1988) et Berlin *et al.* (1993) rapportent que le repérage consiste à relocaliser un document déjà manipulé et qui a été personnellement nommé et classé à l'intérieur de la structure arborescente des répertoires. Dans ce cas, le repérage est basé sur des processus cognitifs tels que la *mémorisation* d'une ou de plusieurs caractéristiques du document recherché (Lansdale & Edmonds, 1992; Berlin *et al.*, 1993; Barreau & Nardi, 1995; Adar, Karger & Stein, 1999; Farrell *et al.*, 2002; Alvarado *et al.*, 2003; Ravasio, Schar & Krueger, 2004). Lansdale & Edmonds (1992, p. 97) rapportent qu'il peut s'agir de l'apparence du document, du contexte temporel (« *We were handling that just before Christmas* ») ou de sa relation avec l'information contextuelle (par exemple, le contexte de production ou de réception passé ou présent). D'autres individus, au contraire, privilégient l'effort cognitif lié à la *classification* pour réduire l'effort cognitif lié au repérage basé sur la capacité de la mémoire (Lansdale, 1988). Quelle que soit l'approche (mémorisation ou classification), Chapman (1999) et Boardman & Sasse (2004) observent que la situation dans laquelle un individu est incapable de repérer ses propres documents électroniques reste occasionnelle tout en étant perçue comme extrêmement frustrante. L'échec au niveau du repérage d'un document serait lié à la destruction, à l'archivage des documents, à une organisation documentaire désordonnée ou à une erreur de classement (Boardman & Sasse, 2004).

Le repérage effectué par une tierce personne implique de localiser un document qui n'a probablement jamais été entrevu ni manipulé mais qui se trouve classé là, « quelque part », dans le schéma de classification créé par un collègue ou un ancien employé, depuis plusieurs mois, ou plusieurs années, dans le cadre d'un contexte de production particulier (Hertzum, 1999). Le repérage d'un document s'appuie ici sur la capacité de cette personne à comprendre et à interpréter un schéma de classification

conçu par autrui (Hertzum, 1999). Dans ce cas, le processus de repérage peut être qualifié de subjectif puisqu'il existe plusieurs interprétations possibles d'un même schéma en raison du « caractère pluriel de toute lecture » (Sénécal, 1997-1998, p. 53). Les recherches sur la notion de genre permettent cependant d'anticiper que la qualité de l'interprétation d'un schéma de classification personnel dépend de l'appartenance de l'interprète à la même communauté de discours⁵ que le créateur et utilisateur principal de ce schéma (Orlikowski & Yates, 1994). Berlin *et al.* (1993) donnent l'exemple de documents relatifs au congrès INTERCHI (« *International Conference on Human Factors in Computing Systems* ») qui seraient en fait classés sous un répertoire intitulé CHI (« *Conference on Computer-Human Interaction* »). Une telle simplification ne serait significative que pour les individus (par exemple, les membres d'une équipe) qui connaîtraient déjà la relation entre INTERCHI et CHI.

2.5.3.2 Mécanismes de repérage des documents

La façon dont les documents peuvent être repérés est étroitement liée à la capacité des mécanismes de repérage proposés aux individus (Chu, 2003). Gagnon-Arguin & Bannouri (1998) notent que, dans le cas des documents électroniques, le repérage n'est pas aussi lié à la classification que pour les documents sur support papier, les systèmes informatiques proposant d'autres moyens pour effectuer cette tâche. Dans le même sens, Barreau & Nardi (1995) et Ravasio, Schar & Krueger (2004) identifient deux mécanismes fondamentalement différents conçus et proposés aux individus pour effectuer le repérage des documents électroniques entreposés sur le poste informatique d'un individu : (1) l'utilisation d'expressions logiques, c'est-à-dire l'utilisation d'un outil de recherche proposé par le système d'exploitation, ou (2) l'accès aux documents à partir de la structure arborescente de répertoires.

Dans le cas d'un repérage « logique » (i.e., utilise des expressions logiques), qui est aussi appelé « recherche » (Boardman & Sasse, 2004) ou « télé portage » (Alvarado

⁵ Une communauté de discours est un groupe de personnes liées par une occupation, ou des intérêts spéciaux et qui partagent un ensemble de connaissances, de possessions et de comportements (Bergquist & Ljungberg, 1999).

et al., 2003), un individu se « téléporte » directement au document recherché à partir de la saisie de mots-clés ou du nom du fichier dans un outil de recherche.

Le repérage « manuel » (ou sensoriel ⁶) des documents est aussi désigné par les termes « navigation » (Barreau & Nardi, 1995; Boardman & Sasse, 2004) ou « orientation » (Alvarado *et al.*, 2003). Ce mécanisme de repérage implique l'utilisation de l'information contextuelle pour se rapprocher du document cible recherché à travers l'arborescence de répertoires, souvent par une série d'étapes (O'Day & Jeffries, 1993; Alvarado *et al.*, 2003; Ravasio, Schar & Krueger, 2004). L'individu doit d'abord identifier le répertoire où le document est classé et doit ensuite afficher la liste des documents contenus dans ce répertoire. L'identification du document recherché s'effectue à partir de l'interprétation de la signification des noms de répertoires et de fichiers ou d'autres attributs donnés par le système (par exemple, la date de modification ou taille du fichier) (Jones, Bruce & Dumais, 2001; Boardman & Sasse, 2004).

Barreau & Nardi (1995), Alvarado *et al.* (2003) et Ravasio, Schar & Krueger (2004) rapportent que les utilisateurs sont généralement très réticents à faire le repérage au moyen des outils de recherche en raison de l'effort cognitif requis pour spécifier complètement le besoin documentaire et toutes ses métadonnées (par exemple, la fonctionnalité de recherche de Windows oblige le plus souvent à mémoriser les noms exacts des fichiers correspondants aux documents recherchés). Il apparaît dans les études empiriques que seuls les individus plus expérimentés dans l'utilisation d'un ordinateur personnel utilisent occasionnellement ces outils en dernier recours, c'est-à-dire lorsqu'ils n'ont aucune idée où le document peut être entreposé, lorsqu'ils ont besoin d'accéder à une information très précise ou pour repérer un document dans un espace de stockage partagé dont le contenu documentaire est mal connu (Nardi & Barreau, 1997; Alvarado *et al.*, 2003; Boardman & Sasse, 2004; Ravasio, Schar & Krueger, 2004).

⁶ Nous pensons que le repérage *manuel* pourrait également être qualifié de *sensoriel* puisque dans le cas de l'accès par navigation à un document particulier l'œil parcourt les niveaux avec plus ou moins de rapidité afin d'y identifier des signes discriminants qui permettent de s'arrêter sur le choix d'un répertoire plutôt qu'un autre.

Dans notre recherche, nous évaluons prioritairement l'efficacité du repérage manuel d'un ensemble donné de documents à partir de la navigation dans la structure arborescente de répertoires et sous-répertoires que nous considérons comme un outil d'organisation intellectuelle et physique des espaces documentaires personnels.

2.5.3.3 Espace documentaire personnel structuré vs. non structuré

Metzger & Lallich-Boidin (2004, p. 12) définissent l'espace documentaire comme un lieu « où s'organisent les collections » et un lieu « de mémoire, de structuration, de mise en relation de documents ». Selon Jacob (2004), l'organisation d'une collection de ressources informationnelles est un des éléments déterminants de l'efficacité du repérage.

Il est établi que le niveau de structuration de l'organisation de l'espace documentaire personnel est influencé par le format technologique et la dimension temporelle de l'information véhiculée par les documents constitutifs de cet espace (Barreau & Nardi, 1995; Boardman, 2001a; Boardman & Sasse, 2004; Ravasio, Schar & Krueger, 2004). La présente section décrit successivement ces deux facteurs d'influence sur le repérage.

Formats technologiques et systèmes d'organisation de l'espace documentaire personnel

Dans le cadre de leur travail, les individus utilisent différents formats technologiques (par exemple, document papier, courriels, pages Web, documents bureautiques) qui peuplent leur espace documentaire personnel (Gagnon-Arguin, 1999; Conférence des recteurs et des principaux des universités du Québec, 2004; Henderson, 2004). Ces formats technologiques sont le plus souvent organisés séparément selon différentes applications ou systèmes d'organisation (par exemple, système de gestion des répertoires et fichiers, système de gestion du courriel et système de gestion des favoris) (Dumais *et al.*, 2003; Henderson, 2003; Boardman & Sasse, 2004). L'espace documentaire personnel est donc un espace fragmentaire, morcelé parmi les différents systèmes, et obligeant la conception par un même individu de structures classificatoires

possédant des caractéristiques variées pour répondre à différents besoins personnels (Boardman, 2001a; Boardman & Sasse, 2004; Jones, 2007) : « A person may maintain several separate, roughly comparable but inevitably inconsistent, organizational schemes for electronic documents, paper documents, e-mail messages, and Web references » (Jones, 2007, p. 453).

Si des recherches récentes ont commencé à s'intéresser aux différentes structures classificatoires mises en place par un même individu pour organiser ses documents sous les différents systèmes qui lui sont proposés (Alvarado *et al.*, 2003; Boardman & Sasse, 2004), les recherches précédentes relatives à l'organisation et au repérage des documents se sont le plus souvent limitées à un seul format technologique en raison de la difficulté d'analyse causée par cette diversité d'outils d'organisation et de besoins (Jones & Maier, 2003). Par exemple, plusieurs études se sont intéressées uniquement à la façon dont les individus organisent et repèrent :

1. leurs documents papier (Cole, 1982; Malone, 1983; Suchman & Wynn, 1984; Lansdale, 1988; Kwasnik, 1989; Blomberg, Suchman & Trigg, 1994; Whittaker & Hirschberg, 2001);
2. leurs courriels (Mackay, 1988; Whittaker & Sidner, 1996; Bälter, 1997; Ducheneault & Belloti, 2001);
3. leurs favoris (Abrams, Baecker, & Chignell, 1998; Jones, Dumais & Bruce, 2002), ou
4. leurs documents électroniques (Carroll, 1982; Kidd, 1994; Barreau, 1995; Barreau & Nardi, 1995; Kaptelinin, 1996; Chapman, 1999; Boardman, 2001a; Alvarado *et al.*, 2003; Henderson, 2003; Henderson, 2004).

Dans le cadre de notre recherche, nous restreignons l'étude aux documents administratifs électroniques créés par des applications bureautiques et qui sont localisés dans le système de répertoires et de fichiers du poste informatique personnel d'un

individu. Ce choix est lié à des contraintes de temps et au fait que presque tous les documents de bureau sont depuis quelques années créés grâce à l'outil électronique et ne sont plus systématiquement imprimés sur papier ou enregistrés sur un serveur (Skupsky, 1999; Dhérent & collaborateurs, 2002; Patterson & Sprehe, 2002; Conférence des recteurs et des principaux des universités du Québec, 2004).

Dimension temporelle de l'espace documentaire personnel

Le niveau de structuration de l'organisation de l'espace documentaire personnel (et par conséquent le repérage à l'intérieur de cet espace) est également lié à la valeur informationnelle (l'utilité personnelle ou institutionnelle) des documents qui évolue à travers le temps. Les recherches de Cole (1982), de Hertzum (1993) et de Barreau & Nardi (1995) ont permis d'identifier trois types d'information qui s'apparentent aux trois âges du cycle de vie des documents d'archives (Pérotin, 1961) :

1. L'information action (« *action information* »), soit l'information d'intérêt à très court terme, et dont la durée de vie est de quelques heures ou de quelques jours, se manifeste par des documents tenus « à portée de main », le plus souvent empilés sur le bureau (virtuel ou non). Pour Hertzum (1999), le repérage de l'*action information* s'appuie essentiellement sur des repères spatiaux (les piles de documents) et nécessite de fréquentes interactions pour rester fonctionnel.

2. L'information de travail (« *working information* »), soit l'information pertinente pour une tâche, un projet ou une activité et qui est utilisée pendant plusieurs semaines, voire plusieurs mois, se manifeste par des documents vaguement organisés sur des étagères ou dans un système de classeurs ou de répertoires électroniques. Le repérage de ce type d'information s'appuie également sur la mémorisation de l'emplacement du document.

3. L'information archivée (« *archived information* »), soit l'information de référence qui peut être conservée pendant plusieurs mois ou années sur le poste de travail des employés. L'utilité de ce type d'information pour l'individu n'est pas clairement

établie. Bien que rarement utilisés, les documents qui véhiculent ce type d'information sont le plus souvent organisés et repérés à travers une vaste structure classificatoire (Berlin *et al.*, 1993; Kidd, 1994; Barreau & Nardi, 1995; Whittaker & Hirschberg, 2001; Henderson, 2003; Boardman & Sasse, 2004).

Pour Hertzum (1999), l'évolution du niveau de structuration de l'organisation de l'espace documentaire personnel, qui part d'une organisation spatiale vaguement structurée et qui s'appuie sur la capacité de la mémoire pour aller ensuite vers des schémas de classification plus élaborés, suit le cycle de vie des documents qui passent de l'état « actif » (« *action information* ») à celui d'« inactif » (« *archived information* »). La caractérisation de cette dimension temporelle est utile dans le cadre de notre recherche car elle rend également possible la catégorisation des modes de repérage mis en place par les individus allant de modes de repérage spatiaux ou qui s'appuient sur la capacité de la mémoire (« *piles* ») à des modes de repérage fortement structurés (« *files* ») (Henderson, 2003).

2.5.4 Schéma de classification et repérage

Cette section présente les présomptions théoriques qui justifient l'existence d'un lien entre le repérage et le schéma de classification utilisé ainsi que les propriétés ou qualités théoriques des schémas de classification qui visent à faciliter et améliorer ce repérage. La description de ces qualités est pertinente à notre recherche car elle permet de mieux comprendre la nature complexe de l'organisation et du repérage des documents. Par la suite, nous justifions le choix de l'efficacité du repérage comme moyen de comparer les schémas de classification des documents administratifs créés par des professionnels en gestion de l'information (schémas institutionnels) et par des employés (schémas personnels).

2.5.4.1 Lien théorique entre schéma de classification et repérage

Il existe des principes théoriques consensuels établissant l'existence d'un lien entre le schéma de classification et le repérage. Le postulat selon lequel l'organisation représentée par le schéma de classification facilite et améliore le repérage n'est pas

nouveau et a été défendu par Bates (1977), Svenonius (1983), Cochrane (1985), Marcella & Newton (1994) et Lin & Chan (1999), parmi d'autres.

Lin & Chan (1999) expliquent bien le rôle de la classification et du schéma de classification pour l'organisation et le repérage des documents. La structure logique et hiérarchique du schéma de classification montre explicitement les relations entre les différentes classes, ce qu'une liste alphabétique ou aléatoire de termes ne révèle pas. Elle est également un moyen de naviguer systématiquement de sujets généraux aux sujets spécifiques, ce qui permet d'augmenter la précision du repérage (Chan, 1995; Lin & Chan, 1999).

Même s'il fait l'objet de critiques en raison des limites liées à son modèle structurel (hiérarchique), le schéma de classification semble influencer l'efficacité du repérage de ressources documentaires par un ensemble d'individus. Ce postulat est appuyé par des réalisations concrètes, par exemple l'adoption ou l'adaptation des schémas de classification traditionnels tels la Classification décimale de Dewey pour normaliser l'organisation des ressources électroniques sur le Web en vue d'améliorer le repérage (McIlwaine & Williamson, 1999; Saàdani & Bertrand-Gastaldy, 2000). Dans le même sens, nous avons mentionné que de nombreux organismes utilisent ou recommandent l'utilisation des schémas de classification institutionnels pour construire de façon logique les répertoires et sous-répertoires pour l'organisation des documents électroniques (Direction des archives de France, 2002; Beaupré, 2004; Chamard, 2004; Conférence des recteurs et des principaux des universités du Québec, 2004). Nous avons également vu que dans un contexte de gestion décentralisée des ressources informationnelles, les individus préfèrent généralement repérer leurs documents organisés selon un schéma de classification personnel plutôt qu'utiliser les fonctionnalités de recherche proposées par les systèmes d'exploitation en vue d'accéder plus rapidement aux documents recherchés.

La recension des écrits suggère que tous les schémas de classification ne permettent pas un repérage également efficace des documents par un groupe d'individus.

Par exemple, les schémas de classification institutionnels, qui utilisent un langage formel, permettraient *a priori* de repérer plus facilement les documents que les schémas de classification personnels qualifiés « d’hermétiques » (Sénécal, 1999) ou « d’idiosyncratiques » (Berlin *et al.*, 1993; Adar, Karger & Stein, 1999; Boardman, 2001b; Saffady, 2002). La présomption sous-jacente est qu’un schéma de classification qui possède certaines propriétés théoriques permet d’optimiser le repérage et les besoins documentaires de plusieurs individus ou groupes d’individus.

La section suivante présente les propriétés théoriques des schémas de classification qui contribuent à optimiser l’organisation et le repérage.

2.5.4.2 Propriétés théoriques d’un schéma de classification et repérage

Certaines propriétés théoriques d’un schéma de classification qui ont une influence sur le repérage ont été identifiées par plusieurs chercheurs (Rousseau, 1980; Couture & Rousseau, 1982; Roberge, 1985; Molholt, 1995; Héon, 1999; Zins, 2002; Roberge, 2004). Il s’agit principalement des propriétés suivantes :

1. Simplicité,
2. Logique,
3. Hospitalité,
4. Autorité, et
5. Universalité.

Nous décrivons brièvement chacune de ces propriétés en les mettant en relation avec les caractéristiques structurelles, logiques et sémantiques d’un schéma de classification :

1. *Simplicité* : un schéma de classification doit être simple afin d’être compris rapidement. La simplicité d’un schéma de classification se manifeste par une structure peu profonde (de trois à quatre niveaux maximum) pour faciliter la navigation tout en maintenant un certain niveau de spécificité. Pour être facilement compris, les termes

utilisés pour nommer les répertoires (ou classes) doivent être concis et intelligibles à l'intérieur de la hiérarchie⁷ (Hertzum & Frøkjær, 1996; Lin & Chan, 1999). L'utilisation d'une terminologie contrôlée⁸ pour la dénomination des catégories (ou répertoires) vise à en diminuer l'ambiguïté⁹ (Hertzum & Frøkjær, 1996; Lin & Chan, 1999).

2. *Logique* : le schéma de classification doit permettre de couvrir le champ des activités ou sujet de manière exhaustive, exclusive et cohérente (Molholt, 1995; Zeller, 2004). Cette qualité se manifeste par l'utilisation d'une seule caractéristique de division logique (ou d'une caractéristique de division dominante) à chaque niveau de la hiérarchie. Un schéma de classification logique facilite théoriquement le repérage en permettant une recherche plus méthodique à l'intérieur d'une structure hiérarchique des répertoires (Maniez, 1987; Fallis & Mathiesen, 2000; Hunter, 2000).

3. *Hospitalité* : le schéma de classification doit être conçu de manière à refléter les développements continuels des activités de l'organisme tout en étant stable (Molholt, 1995; Hertzum, 2004). L'hospitalité d'un schéma de classification permet de s'assurer de l'intégration possible de l'ensemble des documents sans apporter de modifications majeures à la structure logique (au niveau de la macro-structure) (Rousseau, 1980). Dans un contexte administratif, cette qualité se manifeste par l'utilisation des fonctions et activités administratives comme premiers critères de division logique (Renaud, 1975; Héon, 1995; Héon, 1999; Sabourin, 2001; Roberge, 2002; Roberge, 2004).

4. *Autorité* : un schéma de classification doit être construit sur la base d'un consensus (Molholt, 1995). Cette qualité assure le niveau d'accessibilité du schéma de classification par tous les utilisateurs. Dans un contexte de gestion décentralisée des ressources informationnelles, cette qualité se manifeste notamment par l'adoption par l'individu du schéma de classification institutionnel (ou l'une de ses branches) pour

⁷ De nombreux termes sont significatifs seulement lorsqu'ils sont combinés avec des termes plus larges dans les niveaux plus élevés de la hiérarchie (Lin & Chan, 1999). Une fois pris hors de la hiérarchie, un terme peut ne pas toujours porter le contexte et, par conséquent, il peut ne pas être très efficace comme terme de repérage (Hertzum & Frøkjær, 1996).

⁸ Par exemple, le contrôle des synonymes et homonymes.

⁹ L'ambiguïté des classes ne peut être toutefois totalement évitée selon Hertzum (2004)

créer, nommer (à partir de l'énoncé des classes) et organiser les répertoires et sous-répertoires électroniques sur son poste de travail informatique (Beaupré, 2004; Chamard, 2004; Roberge, 2004).

5. *Universalité* : un schéma de classification doit pouvoir être utilisé par n'importe quelle application ou environnement (Molholt, 1995). Par exemple, dans un contexte de gestion centralisée des ressources informationnelles au sein d'un organisme, cette qualité peut se manifester par l'utilisation d'un schéma de classification unique pour l'organisation de l'ensemble des documents électroniques ou analogiques. Cette qualité favorise l'uniformité dans la classification des documents conservés sur plusieurs applications différentes ou sur différents supports et permet, par conséquent, la constitution de dossiers complets au sein d'un organisme (Roberge, 2004).

En résumé, les présomptions théoriques nous permettent de prédire qu'il existe un lien entre l'utilisation d'un schéma de classification institutionnel pour organiser les documents (qui posséderait *a priori* les propriétés théoriques d'un schéma de classification) et l'observation d'un repérage efficace de ces documents comparativement à l'utilisation d'un schéma de classification personnel (qui ne posséderait pas les propriétés théoriques). En effet, un schéma de classification, comme n'importe quel langage, est « porteur » d'information. Il nous communique à travers ses caractéristiques structurelles, logiques et sémantiques une représentation plus ou moins formelle d'une collection documentaire particulière. Ces différentes caractéristiques interagissent entre elles et jouent un rôle important dans l'interprétation qui est faite du schéma de classification pour des fins de repérage de documents. Notre recherche vise à vérifier non seulement l'influence du schéma de classification sur l'efficacité du repérage mais surtout l'importance relative de chacune de ses caractéristiques spécifiques au moment du repérage.

2.5.4.3 Le repérage des documents comme indicateur de mesure de la qualité des schémas de classification

Les schémas de classification produits par les employés pour organiser leurs documents électroniques ont pour fonctions principales, rappelons-le, de classer et de repérer ces documents. L'évaluation fonctionnelle ou opérationnelle de ces schémas de classification passe donc par l'évaluation de l'efficacité du repérage de documents ciblés réalisé à l'aide de ces schémas.

L'utilisation du repérage de documents comme moyen d'évaluer ou de mesurer la qualité (ou qualimétrie) des systèmes de classification (ou tout autre outil de repérage) est connue en informatique et en sciences de l'information (Harter & Hert, 1997; Zins, 2002). Plusieurs études ont utilisé le repérage pour évaluer la performance opérationnelle des systèmes de recherche automatique de l'information (Cleverdon, 1967; Lancaster, 1968; Salton, 1989; Borlund & Ingwersen, 1998), des catalogues et autres instruments de repérage en ligne (Krikelas, 1969; Lytle, 1980; Hertzum & Frøkjær, 1996; Ribeiro, 1996; Hildreth, 2000; Bertrand-Gastaldy & Deschatelets, 2002), ou des systèmes personnels de classification (Malone, 1983; Barreau & Nardi, 1995; Chapman, 1999; Alvarado *et al.*, 2003; Dumais *et al.*, 2003; Boardman & Sasse, 2004; Ravasio, Schar & Krueger, 2004). La précision (i.e., proportion de documents pertinents retrouvés parmi tous les documents retrouvés par le système) et le rappel (i.e., proportion de documents pertinents retrouvés parmi tous les documents pertinents dans la base) sont les deux métriques utilisées en IR (« *Information Retrieval* ») pour mesurer l'efficacité du repérage de l'information.

Dans notre recherche, qui vise à mesurer et comparer l'efficacité du repérage d'un même ensemble de documents préalablement ciblés à l'aide de différents schémas de classification et non pas l'efficacité d'un système de repérage en lui-même indépendamment des autres, les notions classiques de précision et de rappel ne s'appliquent pas (en raison de l'absence, dans notre recherche, d'une notion qui est centrale en IR, soit la notion de pertinence à une requête). Les mesures utilisées dans les travaux de Toms & Campbell (1999) (i.e., *nombre de documents correctement identifiés*

et temps requis pour l'identification), Ribeiro (1996) et Chapman (1999) (i.e., *nombre d'essais*) sont plus appropriées pour évaluer les résultats du repérage de documents déjà ciblés. Alors que le nombre de documents correctement identifiés s'apparente au calcul de la précision du repérage, le temps requis pour le repérage et le nombre de tentatives pour repérer un document sont des indicateurs de mesure de l'efficacité d'un système (i.e., analyse des coûts du repérage en termes de temps et d'effort) (Ribeiro, 1996).

Nous avons retenu les trois variables suivantes pour analyser et mesurer l'efficacité du repérage de documents dans un environnement contrôlé : (1) la proportion de documents cibles repérés, (2) le temps moyen requis pour repérer un document (mesuré en secondes) et (3) la proportion de documents repérés dès le premier essai. Ces trois variables qui mesurent le succès, la rapidité et la facilité du repérage de documents constituent des critères pertinents d'évaluation de l'efficacité du repérage réalisé à l'aide de différents schémas de classification.

2.6 Modèle d'analyse des schémas de classification personnels

La littérature scientifique a servi de cadre de référence pour la conception d'un modèle d'analyse des schémas de classification personnels. Ce modèle nous a permis d'examiner et de décrire de manière uniforme et systématique les caractéristiques des schémas de classification personnels.

Le modèle d'analyse des schémas de classification personnels a été conçu sur la base d'indicateurs de mesure utilisés principalement dans trois domaines de recherche ou disciplines : la gestion personnelle de l'information, les sciences et technologies de l'information et la théorie de la classification. Nous avons vu dans la section 2.4.3.5 Pratiques d'organisation personnelle que les caractéristiques des schémas de classification personnels des documents électroniques ont été étudiées par les chercheurs selon trois aspects principaux : (1) la structure arborescente des répertoires et sous-répertoires électroniques (Boardman, 2001a; Henderson, 2003; Gonçalves & Jorge, 2003; Boardman & Sasse, 2004), (2) les critères de division et de regroupement logique

des répertoires (Boardman & Sasse, 2004; Henderson, 2005; Khoo *et al.*, 2007), et (3) la dénomination des répertoires et des fichiers (Carroll, 1982; Chapman, 1999; Gonçalves & Jorge, 2003). Nous avons regroupé dans notre modèle d'analyse les différents indicateurs de mesure des caractéristiques des schémas de classification selon trois dimensions : la *dimension structurelle*, la *dimension logique* et la *dimension sémantique*. Cette section présente les indicateurs de mesure retenus pour chacune de ces trois dimensions.

2.6.1 Indicateurs de mesure de la dimension structurelle des schémas de classification

La théorie de la classification (Maniez, 1987; Richmond, 1990; Langridge, 1992; Iyer, 1995; Bowker & Star, 2000) confirme la dimension *structurelle* des schémas de classification puisqu'il y est fait mention de profondeur (par exemple, le *nombre de niveaux*) et d'étendue (par exemple, le *nombre de classes principales*) de la structure. En sciences de l'information, les caractéristiques structurelles des schémas de classification conçus pour organiser les ressources Web ont ainsi été étudiées en regard des règles théoriques de classification (Van der Walt, 1998; Hudon, 2000; Zins, 2002). Les principaux indicateurs utilisés par ces chercheurs sont essentiellement le *nombre de classes* et le *nombre de niveaux hiérarchiques*.

Dans le domaine de la gestion personnelle des ressources informationnelles (« *Personal information management* »), les schémas de classification personnels ont été caractérisés à partir d'indicateurs de mesure de leur *structure hiérarchique ou arborescente* (Gonçalves & Jorge, 2003; Boardman & Sasse, 2004; Henderson, 2004; Khoo *et al.*, 2007). Les principaux indicateurs de mesure utilisés sont relatifs à l'étendue (i.e., *nombre de classes principales* et *nombre moyen de classes à chaque niveau*) et à la profondeur (par exemple, *nombre maximum* et *nombre moyen de niveaux hiérarchiques*) (Voir Figure 2.2). D'autres indicateurs sont relatifs à la complexité (i.e., *nombre total de classes*) et à l'équilibre de la structure arborescente (i.e., *écart-type de la moyenne du nombre de classes à chaque niveau*) (Gonçalves & Jorge, 2003; Henderson, 2004).

documents électroniques : (1) l'étendue de la structure (nombre de classes principales, nombre moyen de classes à chaque niveau), (2) la profondeur de la structure (nombre maximum de niveaux hiérarchiques, nombre moyen de niveaux hiérarchiques), (3) la complexité de la structure (nombre total de classes), et (4) l'équilibre de la structure (écart-type de la moyenne du nombre de classes à chaque niveau). L'utilisation de ces indicateurs visait à établir si un modèle structurel particulier était privilégié parmi les participants à l'étude.

2.6.2 Indicateurs de mesure de la dimension logique des schémas de classification

Un schéma de classification doit théoriquement respecter le principe de division logique qui consiste à utiliser *une seule caractéristique* à chaque étape ou niveau de division ce qui permet d'assurer l'exclusivité mutuelle des classes (Maniez, 1987; Van Der Walt, 1998; Bowker & Star, 2000). Selon ce principe, les différentes notions sont réparties dans un ordre logique à l'intérieur de classes qui s'emboîtent du général au plus précis ce qui permet de répondre généralement aux attentes des utilisateurs, de fournir un élément de cohérence important et de permettre une exploration méthodique (Maniez, 1987; Maniez, 1991; Langridge, 1992; Canonne, 1993; Molholt, 1995; Jacob, 2004).

Boardman & Sasse (2004), Henderson (2005) et Khoo *et al.* (2007) ont analysé les concepts représentés dans la dénomination des classes de schémas de classification personnels (Voir section 2.4.3.5 Pratiques d'organisation personnelle). Les résultats des travaux de Boardman & Sasse (2004), de Henderson (2005) et de Khoo *et al.* (2007) nous ont permis d'identifier 18 critères de division et de regroupement logique différents (Voir Tableau 2.1 Principaux critères de division logique recensés dans les écrits, section 2.4.3.5 Pratiques d'organisation personnelle).

Boardman & Sasse (2004) et Henderson (2005) ne disent pas si ces critères de division s'appliquent à la macro-structure (classes principales) ou à la micro-structure (classes subordonnées) des schémas de classification personnels des documents électroniques. Seuls Khoo *et al.* (2007) précisent que les noms des classes au premier

niveau de la hiérarchie représentent le plus souvent des *tâches*, des *projets* ou des *types de documents*.

Notre modèle prévoit une analyse des concepts représentés par les énoncés des classes utilisées pour classifier les documents. Cette analyse fait la distinction entre les concepts représentés par les classes du premier et deuxième niveau hiérarchique en vue d'identifier les critères de division logique qui sont privilégiés par les employés et de vérifier les résultats de Boardman & Sasse (2004), Henderson (2005) et Khoo *et al.* (2007).

Le modèle d'analyse de la dimension logique des schémas de classification examine l'ordre de succession des classes. Cet ordre de succession des classes peut être sans importance puisque cela n'affecte en rien la logique du schéma de classification. Toutefois, la similitude de l'ordre de succession des classes principales d'un schéma de classification à un autre permet une adaptation et donc une utilisation plus rapide d'un schéma de classification dont la structure est familière. Nous avons examiné dans notre recherche l'ordre de succession des classes privilégié dans les schémas de classification personnels observés.

Nous avons retenu les indicateurs de mesure suivants pour décrire la dimension logique des schémas de classification utilisés par les employés pour organiser et classer leurs documents électroniques : (1) critère de division logique au premier et deuxième niveaux hiérarchiques et (2) ordre de succession des classes.

2.6.3 Indicateurs de mesure de la dimension sémantique des schémas de classification

Un schéma de classification est un langage ou un instrument de communication qui comprend des concepts et des classes porteurs d'un contenu *sémantique*. Les termes (simples ou composés) utilisés pour représenter les concepts doivent permettre la compréhension du contenu des classes et des documents qui y sont rattachés (Dumais & Landauer, 1983). La conception et l'application de règles théoriques ou de conventions

relatives à la dénomination des répertoires visent à améliorer la capacité à comprendre la signification de la structure hiérarchique et logique et à repérer des documents.

Alors que plusieurs chercheurs se sont intéressés aux différentes approches utilisées par les individus pour nommer leurs documents électroniques (Carroll, 1982; Chapman, 1999), très peu de recherches ont porté sur la dénomination même des répertoires électroniques (Gonçalves & Jorge, 2003). Nous avons examiné les variables utilisées par les chercheurs qui se sont intéressés aux différentes approches employées par les individus pour nommer leurs documents électroniques (Carroll, 1982; Chapman, 1999). L'examen des résultats de leurs travaux a permis d'identifier trois types de données que nous avons jugé pertinentes pour caractériser les schémas de classification sur le plan sémantique : (1) la longueur moyenne (en nombre de caractères) des noms de classes, (2) les stratégies d'abréviations utilisées et (3) le degré de redondance structurelle et conceptuelle.

Le premier type de données relatif à la dimension sémantique d'un schéma de classification est la *longueur moyenne (en nombre de caractères) des noms de classes* (Gonçalves & Jorge, 2003). Dans certains contextes, les noms des classes continuent d'être limités à huit caractères pour assurer la compatibilité entre les différents systèmes d'exploitation ou les versions de logiciels utilisés (Chamard, 2004). L'usage d'énoncés courts ne fait toutefois pas toujours l'unanimité car il laisse place à une mauvaise interprétation, c'est-à-dire une interprétation trop étroite ou trop large des documents classés sous cette catégorie (Dumais & Landauer, 1983; Roberge, 2004). L'étude du nombre moyen de caractères constitutifs des noms des classes permettra d'établir avec plus de précision quelles sont les pratiques observées dans la dénomination des classes.

Un deuxième type de données est relatif aux *stratégies d'abréviation* privilégiées par les employés pour nommer les classes. Les recherches de Carroll (1982) et Chapman (1999) ont montré que les individus font preuve de beaucoup de créativité dans la dénomination des fichiers dans l'environnement électronique et dans l'utilisation des abréviations. Nous avons conçu une grille de codage (Voir Tableau 2.3) qui contient les

principales stratégies d'abréviation identifiées dans la littérature (Hodge & Pennington, 1973; Carroll, 1982; Chapman, 1999). Dans notre recherche, le codage a été fait manuellement à partir de l'analyse de la forme lexicale du nom des classes et sous-classes et de la description donnée par les employés lors de l'entrevue.

Tableau 2.3 Principales stratégies d'abréviation recensées dans les écrits et observées

Stratégies d'abréviation	Définition
1. Troncature	Décrit la pratique de supprimer des caractères à partir de la droite ou de la gauche (par exemple, « Prof » est une troncature de « Professeur » et « 05 » est une troncature de « 2005 »).
2. Contraction	Décrit la pratique de supprimer les caractères à l'intérieur d'un mot (par exemple, « Px » est une contraction de « Prix »).
3. Acronyme	Décrit la pratique d'utiliser de multiples troncatures dans lesquelles chaque élément lexical est réduit à son premier caractère (par exemple, « CV » est un acronyme de « Curriculum Vitae »).
4. Juxtaposition	Décrit la pratique de juxtaposer deux ou plusieurs chaînes de caractères sans les espacer ou sans placer de symbole entre elles (par exemple, « AssembléeDépartementale » est une juxtaposition de « Assemblée départementale »).
5. Code	Décrit la pratique d'utiliser un système conventionnel de symboles (par exemple, « SPEF 1010 » représente un numéro de formulaire bien connu des techniciennes de l'Université de Montréal) ou personnel (par exemple, « XXXX-03 » signifie « jusqu'en 2003 » pour un répondant interrogé).
6. Mélange	Décrit la pratique d'utiliser simultanément plusieurs stratégies d'abréviation (par exemple, « H2005 » qui signifie « Hiver 2005 » qui est la combinaison d'une troncature et d'une juxtaposition).

Le troisième type de données est relatif au *degré de redondance structurelle* (répétition d'une sous-chaîne de caractères) *et conceptuelle* (répétition d'un même concept) (par exemple, « Long Island Railroad » et « Long Island Expressway ») (Carroll, 1982). Ces redondances considèrent chaque nom de classe dans le contexte des autres noms de classes dans la hiérarchie. Bien que non souhaitable au premier niveau en raison du principe théorique d'exclusivité des classes, cette redondance présente aux niveaux inférieurs aiderait à s'assurer que les noms soient mémorisés (Carroll, 1982, p. 409).

En résumé, les indicateurs de mesure suivants seront utilisés pour décrire les dimensions structurelles, logiques et sémantiques des schémas de classification utilisés par les employés pour organiser et classer leurs documents électroniques :

Dimension structurelle :

1. Étendue de la structure : nombre de classes principales; nombre moyen de classes à chaque niveau;
2. Profondeur de la structure : nombre maximum de niveaux hiérarchiques; nombre moyen de niveaux hiérarchiques;
3. Complexité de la structure : nombre total de classes;
4. Équilibre de la structure : écart-type de la moyenne du nombre de classes à chaque niveau.

Dimension logique :

1. Critère de division logique : critère de division au premier niveau hiérarchique; critère de division au deuxième niveau hiérarchique;
2. Ordre de succession des classes : par exemple, ordre alphabétique, ordre numérique, ordre chronologique.

Dimension sémantique :

1. Longueur moyenne (en nombre de caractères) des noms de classes;
2. Stratégies d'abréviation utilisées;
3. Degré de redondance structurelle et conceptuelle.

2.7 Conclusion

L'analyse des écrits a permis de présenter les assises théoriques de la classification documentaire, les principes de classification des documents administratifs en archivistique, le contexte général relatif à l'organisation personnelle des documents électroniques et la notion de repérage des documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles. Un modèle d'analyse des schémas de classification personnels a été développé, fortement inspiré des indicateurs de mesure mentionnés dans la littérature scientifique et utilisés par plusieurs chercheurs.

Le besoin demeure de poursuivre la recherche dans le domaine de l'organisation personnelle des documents électroniques en milieu de travail. Peu de recherches empiriques ont décrit les caractéristiques multidimensionnelles des schémas de classification conçus et utilisés par des employés pour organiser et classer leurs documents électroniques. De plus, l'influence de ces langages classificatoires sur l'efficacité du repérage des documents réalisé par une tierce personne n'a pas été vérifiée empiriquement.

La recension des écrits a servi de cadre de référence pour sélectionner, décrire, manipuler et évaluer des schémas de classification distincts aux plans structurel, logique et sémantique. Nous présentons dans la section suivante la méthodologie utilisée pour réaliser notre recherche.

3 Méthodologie

3.1 Introduction

La recension des écrits dans le domaine de l'archivistique et de la gestion personnelle de l'information suggère la propension des individus à utiliser des schémas de classification personnels pour l'organisation et le repérage des documents administratifs électroniques qui sont sous leur contrôle direct. Alors que les individus semblent plus ou moins se retrouver dans l'utilisation de leur propre schéma, on peut s'interroger sur la capacité de tierces personnes à repérer des documents à l'aide de schémas de classification personnels avec lesquels ils ne sont pas familiers.

Le but de notre recherche était d'étudier l'efficacité du repérage de documents électroniques réalisé à l'aide de différents schémas de classification dans un contexte de gestion décentralisée des ressources informationnelles. Les résultats de notre recherche devaient nous permettre d'identifier les caractéristiques des schémas qui favorisent un repérage efficace d'un ensemble donné de documents électroniques par des individus autres que le créateur du schéma de classification lui-même.

Les objectifs spécifiques de notre recherche étaient, rappelons-le, de :

OS1. : Décrire les caractéristiques de différents schémas de classification utilisés par des employés pour organiser et classer les documents administratifs électroniques qui sont sous leur contrôle direct.

OS2. : Vérifier les différences sur le plan de l'efficacité du repérage de documents électroniques dans un environnement contrôlé qui sont fonction du schéma de classification utilisé.

Un protocole de recherche en trois étapes fut réalisé pour atteindre ces deux objectifs. La première étape du projet impliquait l'identification et la description des caractéristiques de 21 schémas de classification personnels utilisés par des employés

pour organiser les documents administratifs électroniques qui sont sous leur contrôle direct. Cette caractérisation de schémas de classification personnels fut effectuée à l'aide d'un modèle d'analyse (Voir la description du processus d'élaboration de ce modèle d'analyse à la section 2.6 Modèle d'analyse des schémas de classification personnels) spécialement conçu pour les besoins de notre recherche. La section 3.2 présente le processus de caractérisation d'un échantillon de schémas de classification personnels.

Dans la deuxième étape de notre projet, nous avons procédé à plusieurs manipulations des caractéristiques structurelles, logiques et sémantiques d'un schéma de classification personnel. L'objectif de cette opération était de constituer un corpus de cinq schémas de classification présentant des caractéristiques variées aux plans structurel, logique et sémantique en vue de tester nos hypothèses. La section 3.3 décrit le protocole de manipulation et de création de ces schémas de classification artificiels. La section 3.4 présente les hypothèses de recherche que nous souhaitons vérifier.

Dans la troisième étape de notre projet, une simulation a été réalisée avec 70 répondants pour collecter des données sur la performance de ces cinq schémas de classification sur le plan de l'efficacité du repérage d'un ensemble de 20 documents électroniques. La section 3.5 présente le protocole de simulation du repérage.

À toutes les étapes de notre recherche des mesures ont été prises pour assurer la qualité de la recherche en termes de validité et de fiabilité. La section 3.6 présente ces mesures.

3.2 Caractérisation des schémas de classification

Le but de cette étape de l'étude était de répondre au premier objectif spécifique de notre recherche, soit de décrire les caractéristiques de différents schémas de classification utilisés par des employés pour organiser et classer les documents administratifs électroniques qui sont sous leur contrôle direct. La procédure d'échantillonnage des schémas de classification personnels utilisés pour l'organisation

des documents électroniques et la taille de l'échantillon sont présentées dans la section 3.2.1.

La collecte des données a été réalisée sur une période de plusieurs mois en suivant une procédure rigoureuse. La section 3.2.2 décrit le contexte de cette collecte des données ainsi que la procédure utilisée.

En l'absence d'un instrument d'analyse spécifique des schémas de classification personnels, nous avons conçu un modèle d'analyse qui tient compte des caractéristiques multidimensionnelles de ces langages classificatoires particuliers. Les éléments constitutifs de ce modèle d'analyse ont été identifiés à la suite de la recension des écrits et ont été décrits dans la section 2.6 Modèle d'analyse des schémas de classification personnels. La section 3.2.3 présente la méthode d'analyse de données et rappelle les principaux éléments constitutifs du modèle d'analyse qui a été utilisé à cette étape de la recherche.

Enfin, la section 3.2.4 décrit les caractéristiques d'un schéma de classification personnel qui reflètent quelques-unes des caractéristiques observées au sein de l'échantillon. Ce schéma de classification personnel fut retenu pour des fins de manipulation et d'évaluation dans le cadre de la simulation du repérage d'un ensemble de documents ciblés.

3.2.1 Procédure d'échantillonnage et taille de l'échantillon

Nous avons utilisé la technique d'échantillonnage par choix raisonné en vue d'augmenter l'utilité de l'information tout en restreignant la taille de l'échantillon (Contandriopoulos *et al.*, 1990). En raison des contraintes de temps, nous avons volontairement limité l'étude à 21 schémas de classification. La taille de l'échantillon et la technique d'échantillonnage n'autorisent pas la généralisation statistique des résultats.

3.2.1.1 Sélection du site d'observation

Le choix de l'Université de Montréal comme site d'observation a été basé sur plusieurs considérations relatives (1) à la possibilité d'observer divers schémas de classification personnels, (2) à l'interface informatique utilisée au sein de cet organisme, (3) à la faisabilité de la recherche et (4) à son accessibilité. Ces considérations sont développées dans la présente section.

Nous devons nous assurer d'avoir un site (organisme ou unité administrative) doté d'un schéma de classification institutionnel et qui présentait des situations variées en matière d'organisation des documents électroniques. Le choix de ce contraste au sein du site visait à permettre l'observation de différences dans les caractéristiques des schémas de classification utilisés par les employés et à éviter autant que possible la manipulation (coûteuse en temps) de ces schémas de classification. L'Université de Montréal possède un schéma de classification institutionnel (i.e., *Système officiel de classification*) pour l'organisation des documents au sein de ses différentes unités (Chouinard, Rousseau & Roy, 1990; Université de Montréal, 1995). Le nombre élevé d'employés¹⁰ dans cette institution offrait également la possibilité d'observer une grande diversité de méthodes et de pratiques classificatoires, et d'effectuer des comparaisons à l'intérieur de groupes d'individus ayant le même rôle administratif ou professionnel.

Afin de neutraliser l'impact de l'interface informatique sur le schéma de classification mis en place par un individu et de pouvoir effectuer des comparaisons entre différents schémas de classification des documents, nous avons sélectionné un site d'observation qui utilise l'interface conventionnelle hiérarchique proposée par le système d'exploitation Windows, soit l'Explorateur Windows, en raison de son utilisation prédominante par les employés dans les corporations, agences gouvernementales et autres organismes.

¹⁰ Selon le Rapport annuel de l'Université de Montréal de 2005, plus de 10 570 personnes étaient employées par l'Université de Montréal et ses écoles affiliées. Parmi ces employés, 40 % occupaient des fonctions dans l'administration, le soutien à l'enseignement et à la recherche.

Pour permettre d'effectuer des comparaisons, le repérage devait porter sur des documents ne faisant pas l'objet de restrictions d'accès. Pour ce faire, nous devons privilégier le choix d'un organisme public (i.e., organisme gouvernemental) ou parapublic (par exemple, une université) puisqu'ils sont tenus par la *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels et d'autres dispositions législatives* (L.R.Q., ch. A-2.1) de rendre accessible leurs documents quel que soit le support (Québec, 2001). La réglementation de l'Université de Montréal relative à la gestion des archives, son obligation légale de classifier et de rendre accessibles ses documents quel qu'en soit le support ainsi que la volonté de cette université d'améliorer ses méthodes de gestion des documents administratifs électroniques (Université de Montréal, 1995) constituaient également des facteurs propices à notre recherche au sein de cette institution.

Enfin, le site d'observation choisi était raisonnablement accessible pour la chercheuse en termes de proximité physique mais aussi d'ouverture à la recherche. La chercheuse s'est présentée aux employés, qui pouvaient être *a priori* réticents à autoriser l'accès au contenu de leur ordinateur, comme étudiante au doctorat et chargée de cours à l'École de bibliothéconomie et des sciences de l'information ainsi que comme membre de la Commission des archives de l'Université de Montréal. Cette triple appartenance à l'Université de Montréal a instauré un sentiment de confiance de la part d'employés habitués à travailler eux-mêmes dans cette institution d'enseignement et de recherche.

3.2.1.2 Sélection des employés utilisateurs de schémas de classification personnels

La réalisation de notre deuxième objectif spécifique de recherche supposait la possibilité de disposer d'un échantillon de schémas de classification présentant des caractéristiques variées aux plans structurel, logique et sémantique pour fins de comparaison.

Pour s'assurer d'inclure une variété de caractéristiques dans les schémas de classification des documents électroniques sélectionnés, nous nous sommes basées sur les présomptions émises par plusieurs chercheurs selon lesquelles différents facteurs

peuvent potentiellement affecter le niveau de formalisation de l'organisation des documents électroniques réalisée par les individus. Même si l'influence de ces variables n'a pas été statistiquement vérifiée nous avons tenu compte de ces différents facteurs dans la constitution de notre échantillon. Ces facteurs d'influence du niveau de formalisation de l'organisation des documents électroniques sont principalement : (1) le style de comportement organisationnel de l'individu (« *neat* » or « *messy* ») (Barreau & Nardi, 1995), (2) l'expérience de l'individu dans l'utilisation d'un ordinateur (Chapman, 1999), (3) les tâches professionnelles de l'individu (Whittaker & Snider, 1996; Henderson, 2004), et (4) la présence ou non d'un schéma de classification institutionnel pour l'organisation des documents électroniques (Conférence des recteurs et des principaux des universités du Québec, 2004) (Voir Annexe 19 Modélisation de la relation entre les variables Schéma de classification et Repérage).

Notre recherche a attaché une importance particulière au rôle professionnel des employés en vue d'explorer la variété des schémas de classification d'une catégorie d'employés à l'autre. Plusieurs chercheurs ont observé des différences dans la gestion des documents personnels basées sur les tâches des employés (« *job content* ») (Kidd, 1994; Whittaker & Hirschberg, 2001). Kidd (1994) a défini trois types de travailleurs de bureau : les travailleurs du savoir (« *knowledge workers* »), les travailleurs de la communication (« *communication workers* ») et les travailleurs cléricaux (« *clerical workers* »). Selon Kidd (1994), les secrétaires ou les assistants personnels, par exemple, sont des travailleurs cléricaux dans le sens où ils gèrent les documents de « quelqu'un d'autre »; la gestion des documents constitue une part importante de leurs fonctions. Les travailleurs non-cléricaux, quant à eux, gèrent leurs propres documents. Pour ces derniers, la gestion des documents ne constitue pas une activité importante. Dans le cadre de notre recherche, le groupe d'employés cible était composé à la fois de travailleurs cléricaux (des techniciens à la coordination de travail de bureau) et non-cléricaux (des coordonnateurs de stages) pour permettre l'observation de différences ou de ressemblances dans les caractéristiques des schémas de classification personnels utilisés dans le cadre de leur fonction. En raison de leurs activités communes, on

s'attendait à une certaine homogénéité dans la manière d'organiser les documents à l'intérieur d'une même catégorie d'employé.

Trois raisons sous-tendent l'inclusion dans notre recherche de ces deux catégories particulières d'employés :

1. le rapport quasi-quotidien que ces deux catégories d'employés entretiennent avec des documents administratifs électroniques dans le cadre de leurs activités respectives, ce qui permet de supposer,

2. l'application régulière d'une classification pour organiser un corpus sans doute volumineux de documents administratifs électroniques, et d'envisager à la fois,

3. l'existence de différences et de ressemblances dans les caractéristiques des schémas de classification utilisés en regard de la variété et de l'homogénéité des activités auxquelles participent chaque technicien et coordonnateur de stages.

3.2.2 Collecte des données

Conformément aux exigences relatives aux aspects éthiques de la recherche à l'Université de Montréal, nous avons demandé et obtenu l'autorisation du Comité d'éthique de la Faculté des arts et des sciences de l'Université de Montréal de réaliser notre collecte de données. Le certificat d'éthique ainsi que le formulaire de consentement que chaque employé participant à la recherche devait lire et signer avant d'entamer la procédure de collecte des données sont reproduits en annexe (Voir Annexe 1 Certificat d'éthique et formulaires de consentement).

3.2.2.1 Phase de la collecte des données

Un appel à participation a été lancé en janvier 2006. La collecte des données a été réalisée entre la fin du mois de janvier et le début du mois d'avril 2006. La chercheuse a utilisé le bottin disponible sur le site Web de l'Université de Montréal pour la sélection des techniciens à la coordination de travail de bureau et des coordonnateurs de stages. La sélection des techniciens en coordination de travail de bureau s'est également faite à

partir de la liste communiquée par la Division des archives de l'Université de Montréal qui identifiait les employés ayant reçu une formation sur l'utilisation du schéma de classification institutionnel, soit le *Système officiel de classification* de l'Université de Montréal.

3.2.2.2 Information préliminaire donnée aux employés participants à la recherche

Vingt-cinq coordonnateurs de stages et 235 techniciennes à la coordination de bureau dont le nom figurait dans le bottin accessible à partir du site Web de l'Université de Montréal ont reçu une lettre par messagerie électronique leur présentant brièvement le projet et les tâches à réaliser et sollicitant leur participation pour une rencontre individuelle d'une durée d'une heure et demie à deux heures (Voir Annexe 2 Lettre d'appel à participation adressée aux employés). Pour pouvoir participer, les employés devaient déjà utiliser une structure arborescente de répertoires pour organiser les documents administratifs électroniques qui sont sous leur contrôle direct. Les employés étaient avisés que, dans le cadre de cette rencontre, un logiciel serait utilisé pour collecter des statistiques sur la structure arborescente de leurs répertoires, qu'ils auraient à décrire oralement cette structure et à répondre à diverses questions relatives à l'organisation et au repérage de leurs documents électroniques. Les employés ainsi contactés étaient également invités à remplir un questionnaire et à le remettre à la chercheuse au début de l'entrevue. Ce questionnaire spécifique pour chaque catégorie d'employés comportait, outre des questions socio-démographiques, des questions relatives, par exemple, à leur niveau d'expertise dans l'utilisation d'un ordinateur, à leurs tâches et fonctions professionnelles, à la formation reçue en matière de gestion des documents, à leur perception de la qualité de leur organisation personnelle des documents électroniques et à leur connaissance et utilisation du schéma de classification institutionnel (Voir Annexe 3 Questionnaire adressé aux techniciennes à la coordination de travail de bureau et Annexe 4 Questionnaire adressé aux coordonnateurs de stages). Les employés étaient invités à manifester leur intérêt à participer à la recherche en communiquant avec la chercheuse par courriel ou par téléphone.

Le questionnaire a été essentiellement conçu pour s'assurer que les employés qui se portaient volontaires pour participer à notre recherche remplissaient bien les critères d'inclusion, c'est-à-dire qu'ils occupaient un poste de technicien à la coordination de travail de bureau ou de coordonnateur de stages (certains employés contactés avaient changé de poste ou étaient à la retraite) et qu'ils utilisaient un outil de classification pour organiser les documents administratifs électroniques qui sont sous leur contrôle direct. La candidature des employés volontaires dont la nature du poste avait changé ou qui n'avaient pas développé de schéma de classification pour organiser leurs documents électroniques n'a pas été retenue.

3.2.2.3 Profil des employés

Pour les besoins de notre recherche, nous avons obtenu la participation de 19 employés soit 11 techniciens à la coordination de travail de bureau et huit coordonnateurs de stages qui ont volontairement accepté de remplir le questionnaire et de participer à l'entrevue. Le profil démographique et professionnel de ces deux catégories d'employés est présenté dans les deux sections suivantes.

Profil des techniciens à la coordination de travail de bureau

Les 11 techniciens interrogés étaient tous des femmes. Les réponses au questionnaire révèlent que cinq d'entre elles (45.46 %) avaient entre 35 et 49 ans, et les six autres (54.54 %), 50 ans ou plus.

La grande majorité des techniciennes se perçoivent comme des utilisatrices expérimentées en informatique. Cinq d'entre elles (45.46 %) se définissent comme ayant un niveau « intermédiaire » de compétence en informatique (i.e., exécutent des tâches simples avec un nombre limité de logiciels) et six d'entre elles (54.54 %) se définissent comme « expertes » (i.e., utilisent d'une façon compétente une grande variété de technologies informatiques).

Dans un contexte de travail, chaque employé produit, reçoit et organise des documents électroniques dans le cadre des activités qui découlent de son mandat. Les

techniciennes à la coordination de travail de bureau réalisent ou participent à différentes activités administratives reliées à leur fonction. Parmi les principales activités énoncées, ces techniciennes sont amenées à préparer et faire le suivi des réunions des assemblées et comités. Elles peuvent être invitées à collaborer à la gestion des dossiers d'engagement et de promotion des professeurs. Elles participent souvent à l'organisation des cérémonies de bourses et de prix remis aux étudiants. Elles peuvent être amenées à aider les professeurs dans la rédaction de documents ou dans la préparation des demandes de subventions de recherche.

Les techniciennes entretiennent un rapport quasi-quotidien avec les documents dans le cadre de leurs activités. Elles doivent gérer de nombreux formulaires, modèles de lettres, listes et tableaux. Elles sont aussi amenées à produire, à corriger et à diffuser de nombreux documents de réunion (par exemple des avis de convocation et des procès-verbaux), de la correspondance et des publications (par exemple, des rapports annuels).

Alors que toutes les techniciennes ont déclaré dans le questionnaire être responsables du système de classement des documents, seulement sept d'entre elles (63.63 %) ont suivi une formation en gestion de documents dans le cadre d'un programme d'étude de niveau post-secondaire, collégial ou universitaire. Parmi elles, six (54.54 %) ont déclaré avoir suivi une formation, le plus souvent d'une durée de trois heures, organisée par la Division des archives de l'Université de Montréal; une seule (9.09 %) se souvient d'avoir suivi une formation sur la classification des documents électroniques.

Malgré la formation reçue en gestion de documents par sept techniciennes, seulement quatre d'entre elles (36.36 %) utilisent certaines classes principales (Voir Annexe 5 Classes du *Système officiel de classification* de l'Université de Montréal utilisées par quatre techniciennes) du *Système officiel de classification* de l'Université de Montréal pour l'organisation des documents administratifs électroniques qui sont sous leur contrôle direct.

Les techniciennes interrogées évaluent positivement l'organisation de leurs documents électroniques puisque huit d'entre elles (72.72 %) perçoivent leurs documents électroniques comme étant « très bien organisés » ou « extrêmement bien organisés ». La perception de la qualité du repérage des documents électroniques par les techniciennes va dans le même sens puisque sept d'entre elles (63.63 %) perçoivent le repérage de leurs documents électroniques comme étant « très facile », voire « extrêmement facile ».

Profil des coordonnateurs de stages

Parmi les huit coordonnateurs de stages interrogés, nous avons rencontré sept femmes et un homme. Les réponses au questionnaire révèlent que quatre d'entre eux (50 %) avaient entre 26 et 34 ans, deux (25 %) avaient entre 35 et 49 ans et les deux autres (25 %), 50 ans ou plus.

Les coordonnateurs de stages sont des utilisateurs le plus souvent expérimentés en informatique. Trois d'entre eux (37.5 %) se définissent comme de niveau « intermédiaire » (i.e., exécutent des tâches simples avec un nombre limité de logiciels) et cinq (62.5 %) se définissent comme « experts » (i.e., utilisent d'une façon compétente une grande variété de technologies informatiques).

Les réponses au questionnaire et les entrevues nous éclairent sur les différentes activités des coordonnateurs de stages. Ces derniers peuvent être amenés à participer à différentes activités de nature administrative, professionnelle ou pédagogique reliées à leur fonction. Les coordonnateurs de stages sont non seulement responsables des activités relatives à la coordination générale des stages (par exemple, l'organisation de rencontres avec les étudiants, la planification des stages, l'assignation des étudiants en stage et parfois les visites dans les milieux), ils jouent également le rôle d'intermédiaire ou de personne-ressource entre les milieux de stage, les employeurs et les stagiaires. Les coordonnateurs de stages peuvent être appelés à participer à la formation des partenaires qui supervisent les étudiants en stage, à la formation des étudiants qui ont de la difficulté en stage, à l'élaboration et à la mise à jour d'outils pédagogiques (par exemple, grilles d'évaluation et guides de l'étudiant). Enfin, les coordonnateurs de stages peuvent être

impliqués dans la planification d'événements spéciaux (par exemple, la Journée de l'emploi) ou invités à participer à diverses activités de prospection ou de relations publiques (par exemple, le Salon des études).

Dans le cadre de ses activités, le coordonnateur de stages est donc amené à gérer d'année en année beaucoup de documentation et de correspondance, d'offres de stages, d'offres d'emploi, de banques de données ainsi que de nombreux rapports produits ou reçus.

Malgré la relation quasi-quotidienne que les coordonnateurs de stages entretiennent avec les documents, seulement deux (25 %) ont déclaré avoir suivi une formation en gestion de documents dans le cadre d'un programme d'étude ou d'une formation de trois heures donnée par la Division des archives de l'Université de Montréal. Aucun coordonnateur de stages n'utilise le *Système officiel de classification* de l'Université de Montréal pour l'organisation des documents administratifs électroniques qui sont sous son contrôle direct.

Trois coordonnateurs de stages (37.5 %) perçoivent leurs documents électroniques comme étant « moyennement bien organisés » et quatre (50 %) perçoivent leurs documents électroniques comme étant « très bien organisés ». La perception de la qualité du repérage des documents électroniques par les coordonnateurs de stages va dans le même sens puisque trois coordonnateurs de stages (37.5 %) perçoivent le repérage de leurs documents électroniques comme étant « moyennement facile » et quatre coordonnateurs (50 %) perçoivent le repérage de leurs documents électroniques comme étant « très facile ».

Les réponses au questionnaire confirment la variété des activités auxquelles participent chaque technicienne et coordonnateur de stages, ce qui nous garantit un corpus de documents administratifs électroniques suffisamment volumineux et varié pour inciter ces deux catégories d'employés à concevoir un schéma de classification pour l'organisation de leurs documents électroniques.

Les réponses au questionnaire illustrent également la variété des pratiques d'organisation et de repérage des documents privilégiées par les employés interrogés tout en révélant leur satisfaction générale quant à la performance de leur schéma de classification personnel des documents électroniques.

3.2.2.4 Procédure de collecte des données

La chercheuse a rencontré individuellement chaque employé participant à la recherche sur son lieu de travail, devant son poste informatique. Les rencontres individuelles ont duré entre une heure et demie et deux heures. L'objectif de ces rencontres était de collecter, en présence de l'employé, des données de nature quantitative (par exemple, le nombre total de classes) et qualitative (i.e., description des classes) sur le schéma de classification personnel utilisé pour l'organisation des documents électroniques. Cette section présente le déroulement de ces rencontres individuelles ainsi que les outils de collecte utilisés pour colliger les données relatives aux caractéristiques des schémas de classification étudiés.

Un protocole d'entrevue (Voir Annexe 6 Protocole d'entrevue) a été conçu en vue d'énumérer l'ensemble de tâches que la chercheuse et l'employé devaient systématiquement réaliser. Les entrevues ont été enregistrées, avec le consentement écrit des employés.

Comme certaines études font état des gens qui « modifient, consciemment ou non, le traitement des données qu'ils produisent ou manipulent lorsqu'ils savent que ces dernières servent à des fins de recherche » (Contandriopoulos *et al.*, 1990, p. 70), chaque employé participant à l'entrevue a eu pour consigne de ne pas réorganiser ses répertoires et ses documents électroniques avant l'entrevue. Tous les employés ont déclaré avoir respecté cette consigne. Dans l'éventualité où des modifications auraient malgré tout été réalisées, ces petits changements n'invalident pas les résultats de la recherche puisque l'unité d'analyse demeure l'ensemble des caractéristiques du schéma de classification personnel.

Dans un premier temps, chaque employé a été invité à décrire brièvement ses fonctions et tâches administratives ou professionnelles. Certains employés ont remis à la chercheuse une description écrite de leurs tâches, le cas échéant. La connaissance des activités dans lesquelles est impliqué un individu est importante car il s'agit, en principe, du premier critère de division logique à la base de la conception d'un schéma de classification des documents administratifs (Voir section 2.3.3.5 Principes et méthodes d'élaboration du schéma de classification des documents administratifs). Rappelons que les répondants de la recherche menée par Boardman & Sasse (2004) organisaient intuitivement une partie de leurs documents selon leurs activités à court ou à plus long terme.

Nous avons ensuite réalisé, toujours avec l'accord et en présence de l'employé, des saisies d'écran (Voir Annexe 7 Exemple de saisie d'écran) de l'arborescence des classes situées dans le répertoire racine « Mes Documents » (ou son équivalent). L'objectif de cette opération était d'obtenir une « photo » de la structure arborescente des classes et sous-classes créée par l'employé pour des fins d'analyse. Préalablement à cette opération, l'employé avait été invité à indiquer les classes ou sous-classes (généralement une ou deux classes par schéma) dont il ne désirait pas faire la présentation ou la description pour des raisons de confidentialité ou de respect de la vie privée. Ces classes n'ont pas été incluses dans l'analyse des schémas de classification personnels.

Pour accélérer la collecte des données quantitatives relatives aux caractéristiques structurelles des schémas de classification, nous avons utilisé le programme informatique d'analyse de l'espace documentaire personnel PDS (« *Personal Document Space* ») conçu par Gonçalves & Jorge (2003). Ce logiciel permet une collecte et une analyse statistique automatisées d'une partie des données sur les caractéristiques structurelles des schémas de classification personnels (i.e., nombre total de classes et de sous-classes, nombre total de fichiers, nombre moyen de classes à chaque niveau hiérarchique et écart-type de la moyenne du nombre de classes à chaque niveau hiérarchique). Le rapport produit par le logiciel qui fait état de l'analyse statistique des données a été montré aux employés pour les rassurer sur la nature des données ainsi collectées, et sur le fait

qu'aucune information sur le contenu même des documents n'avait été colligée (Voir Annexe 8 Extrait d'un rapport produit par le logiciel PDS). L'autre partie des données relatives aux caractéristiques structurelles (i.e., nombre de classes principales, nombre maximum de niveaux hiérarchiques, nombre minimum de niveaux hiérarchiques et nombre moyen de niveaux hiérarchiques par classes principales), logiques et sémantiques fut manuellement colligée par la chercheuse après l'entrevue à partir de l'observation des saisies d'écran de la structure arborescente de répertoires et sous-répertoires.

Suite aux saisies d'écran et à la collecte automatique des données relatives aux caractéristiques structurelles, chaque employé était invité à décrire oralement et brièvement la signification et le contenu de ses classes et sous-classes. La durée de cette opération pouvait varier entre une heure et une heure et demie en raison du nombre plus ou moins élevé de classes et de sous-classes utilisées par chaque employé. L'objectif de cette opération était de collecter des données qualitatives relatives aux critères de division et de regroupement logique utilisés par les employés. En raison de la polysémie de certains termes et de l'usage d'abréviations dans la dénomination des classes, il s'agissait de valider la nature des concepts représentés par les classes.

Suite à la description et à la définition des classes utilisées pour organiser les documents électroniques, chaque employé était invité à décrire 20 documents sélectionnés par la chercheuse à l'aide d'une table aléatoire de nombres à l'intérieur de la structure arborescente de répertoires. L'objectif de cette dernière tâche était d'obtenir un échantillon de descriptions de documents existant réellement et classés à l'aide de chaque schéma de classification personnel observé. Ce corpus devait servir à la simulation du repérage des documents dans le cas où le schéma de classification de l'employé interrogé était retenu pour le test des hypothèses.

Avant de conclure la rencontre, les employés étaient invités à formuler des commentaires généraux sur leur schéma de classification personnel des documents électroniques. Si les employés ont déclaré dans le questionnaire être généralement

satisfaits du schéma de classification utilisé pour l'organisation personnelle de leurs documents électroniques, le tiers d'entre eux ont dit durant l'entrevue qu'ils rencontraient cependant quelques difficultés dans son utilisation et ont manifesté leur désir de l'améliorer. Durant les entrevues, six employés (quatre coordonnateurs de stages et deux techniciennes) (31.57 %) ont identifié certaines « faiblesses » au sein de leur schéma de classification. Par exemple, ils ont mentionné la présence de classes redondantes à l'intérieur d'un même niveau hiérarchique, l'usage fréquent de la sous-classe « Divers » et le besoin de mettre à jour le schéma utilisé en « faisant le ménage ».

En conclusion, les entrevues menées auprès de ces employés ont permis de collecter des données sur les caractéristiques structurelles, logiques et sémantiques de 21 schémas de classification personnels : onze schémas utilisés par des techniciennes et dix¹¹ schémas utilisés par des coordonnateurs de stages. La section suivante présente la méthode utilisée pour analyser ces données.

3.2.3 Méthode d'analyse des données

3.2.3.1 Préparation des données pour l'analyse

Les schémas de classification des employés ont été reproduits et « mis à plat » dans un fichier Excel pour des fins d'analyse. L'utilisation de ce tableur a permis de faciliter l'obtention de statistiques descriptives (i.e., pourcentages, moyennes et écarts-type) de chacune des caractéristiques observées.

3.2.3.2 Analyse des données

Les schémas de classification personnels ont été analysés à partir d'un modèle qui comporte des mesures continues (i.e., données quantitatives) et des mesures catégorielles (i.e., données qualitatives) regroupées selon trois dimensions : la dimension structurelle, la dimension logique et la dimension sémantique. Le processus de conception de cette grille ainsi que les indicateurs de mesure retenus pour caractériser les schémas de

¹¹ Deux employés parmi les huit coordonnateurs de stages interrogés utilisent deux schémas de classification personnels : un schéma de classification partagé avec d'autres collègues de travail et un schéma de classification pour leurs besoins personnels.

classification personnels ont été décrits dans la section 2.6 Modèle d'analyse des schémas de classification personnels. Nous rappelons que la dimension structurelle est représentée dans ce modèle d'analyse par des données quantitatives relatives à l'étendue, la profondeur, la complexité et l'équilibre de la structure. Les données qualitatives relatives aux critères de division logique et à l'ordre de succession des classes ont été utilisées pour décrire la dimension logique. La dimension sémantique est représentée à la fois par les données quantitatives et qualitatives relatives à la longueur moyenne (en nombre de caractères) des noms de classes, aux stratégies d'abréviation utilisées et au degré de redondance structurelle et conceptuelle.

Le Chapitre 4 présente les résultats de l'analyse des caractéristiques de 21 schémas de classification personnels. Les caractéristiques structurelles, logiques et sémantiques relativement semblables des schémas de classification personnels observés n'ont pas permis de neutraliser l'effet des autres caractéristiques lors de la vérification de l'influence d'une caractéristique donnée. Afin de pouvoir réaliser notre deuxième objectif de recherche, soit de vérifier les différences sur le plan de l'efficacité du repérage de documents électroniques dans un environnement contrôlé qui sont fonction du schéma de classification utilisé, nous avons complété l'analyse de schémas existants par la création de schémas de classification artificiels. Le but de cette manipulation était d'assurer la représentativité de différents schémas de classification et de tester, dans une deuxième phase de collecte de données, leur influence sur l'efficacité du repérage de documents électroniques ciblés.

Pour ce faire, nous avons sélectionné au sein de notre échantillon le schéma de classification personnel d'un employé pour servir de base à la création de quatre schémas de classification artificiels présentant des caractéristiques variées aux plans structurel, logique et sémantique. Les caractéristiques du schéma de classification personnel de cet employé et les raisons qui ont motivé le choix de ce schéma pour des fins de manipulation sont décrites dans la section suivante.

3.2.4 Caractéristiques du schéma de classification personnel retenu pour des fins de manipulation (Schéma A)

Le schéma de classification personnel retenu (Schéma A) pour des fins de manipulation et d'évaluation dans le cadre d'une expérimentation quasi-contrôlée a été conçu par un coordonnateur de stages pour l'organisation des documents électroniques qui sont sous son contrôle direct.

Les raisons de ce choix sont présentées en trois points : les justifications empiriques, les justifications théoriques et les justifications méthodologiques.

En ce qui concerne les justifications empiriques, le schéma A a été choisi en raison de ses caractéristiques communes avec quelques-unes des caractéristiques observées dans l'échantillon de schémas de classification personnels (Voir Annexe 20 Synthèse des caractéristiques des schémas de classification personnels observés et du schéma A). Examinés globalement, les schémas de classification personnels diffèrent sensiblement les uns des autres aux plans structurel, logique et sémantique. Cependant, la description détaillée et la comparaison systématique de chacun des indicateurs de mesure des caractéristiques ont permis d'identifier plusieurs traits structurels, logiques et sémantiques communs à la plupart des schémas. Nous considérons à partir des définitions données par Gonçalves & Jorge (2003), Boardman & Sasse (2004) et Khoo *et al.* (2007) relatives à l'étendue, la profondeur, la complexité et l'équilibre d'une structure, qu'un plus grand nombre de schémas observés ont une macro-structure étendue (entre 15 et 50 classes principales), une micro-structure peu étendue (une moyenne de 2.73 classes à chaque niveau), une structure peu profonde (moins de trois niveaux hiérarchiques), complexe (entre 100 et 200 classes) et peu équilibrée (écart-type de la moyenne du nombre de classes à chaque niveau qui varie entre 3 et 5). Au plan logique, la thématique est le plus souvent utilisée comme premier critère de division et les classes sont ordonnées alphabétiquement. Sur le plan sémantique, on observe une longueur moyenne des noms de classes de 16.12 caractères, un usage privilégié d'acronymes et un degré de redondance structurelle et conceptuelle élevé (entre 20 % et 40 % des classes sont redondantes).

Le Tableau 3.1 présente les caractéristiques du Schéma A mises en relation avec les caractéristiques observées dans les schémas de classification personnels étudiés.

Tableau 3.1 Caractéristiques des schémas de classification personnels étudiés et caractéristiques du schéma de classification personnel retenu pour des fins de manipulation (Schéma A)

Caractéristiques observées dans l'échantillon de 21 schémas de classification personnels	Caractéristiques observées dans le schéma de classification personnel retenu (Schéma A)
Dimension structurelle	
Macro-structures étendues (28.14 classes principales en moyenne)	→ Macro-structure étendue (42 classes principales)
Micro-structures peu étendues (2.73 classes en moyenne à chaque niveau)	→ Micro-structure peu étendue (2.61 classes en moyenne à chaque niveau)
Structures peu profondes (1.67 niveau en moyenne)	→ Structure peu profonde (1.56 niveau en moyenne)
Structures complexes (355.05 classes en moyenne)	→ Structure complexe (192 classes)
Structures peu équilibrées (Écart-type de 6.61)	→ Structure peu équilibrée (Écart-type de 6.91)
Dimension logique	
Aucun critère de division logique ne prédomine bien qu'un regroupement par « thème » et « mélange » soit privilégié au niveau de la macro-structure (classes principales)	→ Critère de division logique qui est privilégié au niveau de la macro-structure : « thème » (40 % des classes principales)
Ordre alphabétique de succession des classes	→ Ordre alphabétique de succession des classes
Dimension sémantique	
Nom des classes ayant un nombre moyen de 16.12 caractères	→ Nom des classes ayant un nombre moyen de 19.12 caractères
Utilisation privilégiée d'acronymes	→ Utilisation privilégiée d'acronymes
Degré de redondance structurelle et conceptuelle élevé (23.02 % de classes redondantes)	→ Degré de redondance structurelle et conceptuelle élevé (35.05 % de classes redondantes)

Le schéma A présente une macro-structure (classes principales) que l'on peut qualifier d'étendue (42 classes principales) et une micro-structure (classes subordonnées) peu étendue (une moyenne de 2.61 classes à chaque niveau). Ce schéma de classification est peu profond (moyenne de 1.56 niveaux par classe principale), complexe (total de 192 classes) et peu équilibré (écart-type du nombre moyen de classes à chaque niveau

relativement élevé). Le schéma A présente un regroupement thématique et un ordonnancement alphabétique des classes principales, un usage privilégié d'acronymes et un degré de redondance structurelle et conceptuelle élevé.

D'un point de vue théorique, la sélection du schéma A est pertinente car quelques-unes de ses caractéristiques logiques et sémantiques démontrent le non respect des principes théoriques de conception des schémas de classification. Par exemple, plusieurs critères de division sont utilisés à chaque niveau ce qui suggère une faible cohérence (Maniez, 1987; Fallis & Mathiesen, 2000; Hunter, 2000). De plus, un usage fréquent d'abréviations dans la dénomination des classes pourrait rendre ces dernières peu explicites (Carroll, 1982; Saffady, 2002). Le schéma A apparaît *a priori* loin d'être optimal sur le plan de l'efficacité du repérage et représente, par conséquent, un modèle réel particulièrement intéressant dans le cadre d'une simulation du repérage.

Il existe peu de références dans la littérature scientifique sur les attributs possibles d'un modèle de schéma de classification personnel pour justifier au plan méthodologique la sélection du schéma A. Cependant, l'approche expérimentale retenue dans notre recherche justifie l'importance des observations empiriques et des bases théoriques pour appuyer notre choix. D'après Brewer (1985), la construction d'une expérimentation est un processus en deux étapes :

« Il s'agit d'abord de dégager les éléments essentiels au sein d'une situation complexe, puis de les reconstruire dans une situation différente et plus réduite de telle manière à préserver l'isomorphisme entre les deux situations. Pour chaque composante de la situation réelle qui apparaît théoriquement importante, une composante correspondante doit se retrouver en laboratoire » (Brewer, 1985 cité dans Jones, 2000, p. 309).

N'importe lequel des schémas de classification personnels observés au sein de notre échantillon qui n'applique pas les principes théoriques de conception des schémas de classification aurait pu être sélectionné pour des fins de manipulation. La réalisation du pré-test a toutefois révélé que le choix d'un schéma de classification personnel complexe et non conforme aux principes théoriques de classification (par exemple, un

schéma qui n'applique pas le principe de division logique) facilitait la manipulation. Le pré-test a notamment révélé qu'il est plus facile de manipuler une macro-structure étendue (par exemple, en réduisant le nombre de classes principales) qu'une macro-structure peu étendue. Par conséquent, le schéma A qui possédait ces deux conditions (complexité et non-conformité aux principes classificatoires) fut sélectionné.

3.3 Manipulation des schémas de classification

Rappelons que cette deuxième étape de notre recherche avait pour objectif de créer un échantillon de cinq schémas de classification présentant des caractéristiques variées aux plans structurel, logique et sémantique en vue d'étudier, dans le cadre d'une simulation, l'influence de ces différentes caractéristiques sur le repérage des documents électroniques.

D'après Walliser (1977), la manipulation de variables est une opération inhérente à toute technique de simulation :

« La simulation consiste ainsi à manipuler les différents groupes de variables d'un modèle (variables de commande, variables internes ou paramètres structuraux du modèle, variables d'environnement), soit pour comprendre le fonctionnement d'un système à l'étude (fonction cognitive), soit pour prévoir comment évoluerait le système selon différentes valeurs des variables de commande (fonction prévisionnelle), soit pour déterminer quelles valeurs devraient prendre les variables de commande pour atteindre des objectifs (fonction décisionnelle), soit enfin pour représenter les relations souhaitables entre les variables d'entrée et les variables de sorties (fonction normative) » (Walliser, 1977 cité dans Contandriopoulos *et al.*, 1990, p. 39).

Dans notre recherche, la manipulation des caractéristiques du schéma A visait à neutraliser l'influence d'une caractéristique donnée et à évaluer dans le cadre d'une simulation l'influence des deux autres caractéristiques sur le repérage des documents électroniques. À partir d'un schéma de classification personnel existant (Schéma A) la chercheuse a artificiellement créé quatre autres schémas de classification (Schéma B, C, D et E) dotés de caractéristiques structurelles, logiques et sémantiques variées.

Nous présentons dans cette section la nature et le produit (résultat) des différentes manipulations réalisées aux plans structurel, logique et sémantique.

3.3.1 Nature des manipulations

Les quatre manipulations effectuées à partir du schéma A sont les suivantes :

1. Manipulation de la sémantique (i.e., schéma de classification personnel de l'employé A augmenté d'une définition des classes) (Schéma B);

2. Manipulation de la structure et de la sémantique (i.e., suppression de la structure arborescente et modification des dimensions sémantique et logique) (Schéma C);

3. Manipulation de la structure et de la logique (i.e., macro-structure moins étendue, structure plus profonde, division basée sur les activités de l'employé plutôt que sur des thèmes) (Schéma D);

4. Manipulation de la structure, de la logique et de la sémantique (i.e., macro-structure moins étendue, structure plus profonde, division basée sur les activités de l'employé plutôt que sur des thèmes, diminution de la proportion de classes abrégées) (Schéma E).

Nous présentons dans les sections suivantes les produits de ces manipulations, soit les caractéristiques structurelles, logiques et sémantiques des schémas de classification créés artificiellement (Schémas B, C, D et E).

3.3.2 Produits de la manipulation

Il n'existe pas de référence méthodologique ni de travaux similaires pour justifier les manipulations réalisées sur le schéma de classification personnel de l'employé A. Ces manipulations s'appuient essentiellement sur des bases théoriques et empiriques. La présente section décrit et justifie les caractéristiques de chacun des schémas de classification artificiellement créés.

3.3.2.1 Caractéristiques du schéma de classification manipulé dans sa dimension sémantique (Schéma B)

L'objectif de cette manipulation était de modifier la dimension sémantique tout en maintenant la structure et la logique du schéma de classification de l'employé (Schéma A). L'approche retenue a été d'intégrer à la structure la définition de chacune des classes telle que communiquée par l'employé en entrevue et transcrite par la chercheure (Voir Tableau 3.2).

Tableau 3.2 Schéma B : Manipulation de la sémantique

Schéma original (Schéma A)	Schéma manipulé au plan sémantique (Schéma B)
Accueil des étudiants	Accueil des étudiants Documents et information pour les nouveaux étudiants de première année participant à la journée d'accueil
sept. 2005	sept. 2005 Septembre 2005
Attestation SS	Attestation SS Documents relatifs à l'attestation des superviseurs de stage : listes d'envoi, lettres de félicitations, etc.
Documents officiels	Documents officiels Documents finaux
Envoi avril 2004	Envoi avril 2004 Listes des superviseurs de stages qui ont reçu leur attestation en avril 2004
Envoi janvier 2003	Envoi janvier 2003 Listes des superviseurs de stages qui ont reçu leur attestation en janvier 2003
Envoi Octobre 2004	Envoi Octobre 2004 Listes des superviseurs de stages qui ont reçu leur attestation en octobre 2004
Obtention et renouvellement du titre - A05	Obtention et renouvellement du titre - A05 Obtention et renouvellement du titre de superviseur de stage à l'automne 2005
Base de données	Base de données Listes qui ont été produites en utilisant la base de données
Comité de gestion	Comité de gestion Documents relatifs au Comité de gestion (par exemple, les comptes-rendus de réunions)
Comité des stages	Comité des stages Documents relatifs au Comité des stages
[Sigle de cours 1]	[Sigle de cours 1] Tout ce qui concerne le 1er stage (1 ^{ère} année)
[Sigle de cours 2]	[Sigle de cours 2] Tout ce qui concerne le 2ème stage (2 ^{ème} année)
Etc.	Etc.

Cette manipulation se justifie empiriquement par l'utilisation comparable de notes d'application et de définitions de classes dans les schémas de classification documentaires et institutionnels. Par exemple, le *Système officiel de classification* proposé par la Division des archives de l'Université de Montréal, décrit la classe « 11. Fournitures » de la manière suivante :

« 11. FOURNITURES

Documents relatifs aux fournitures. En particulier aux fournitures scolaires et aux articles de bureau. Il peut s'agir d'inventaire, de catalogues, de listes de prix, de modes d'emploi ou de correspondance avec des fournisseurs. »

(En ligne : <http://www.archiv.umontreal.ca/descript.html>)

D'un point de vue théorique, l'ajout de ces définitions contribuerait à diminuer l'ambiguïté des termes (Dahlberg, 1988).

La création artificielle du schéma de classification B visait à permettre l'étude de l'influence d'une sémantique personnelle rendue explicite (par une définition des classes) sur l'efficacité du repérage de documents ciblés.

3.3.2.2 Caractéristiques du schéma de classification manipulé dans ses dimensions structurelle et sémantique (Schéma C)

L'objectif de cette manipulation était de modifier la structure arborescente et la sémantique du schéma de classification de l'employé (Schéma A). L'approche retenue a consisté à supprimer tous les niveaux hiérarchiques (Voir Tableau 3.3) ce qui a eu aussi pour effet de modifier la dimension sémantique en augmentant : (1) la longueur moyenne (en nombre de caractères) des intitulés des classes (i.e., longueur moyenne de 37.9 caractères), (2) le niveau de redondance structurelle et conceptuelle (i.e., 90 % des classes sont redondantes) et (3) la proportion de classes abrégées (48 %).

Tableau 3.3 Schéma C - Manipulation de la structure et de la sémantique

Schéma original (Schéma A)	Schéma manipulé aux plans structurel et sémantique (Schéma C)
Accueil des étudiants	Accueil des étudiants
sept. 2005	Accueil des étudiants - sept. 2005
Attestation SS	Attestation SS
Documents officiels	Attestation SS - Documents officiels
Envoi avril 2004	Attestation SS - Envoi Avril 2004
Envoi janvier 2003	Attestation SS - Envoi janvier 2003
Envoi Octobre 2004	Attestation SS - Envoi Octobre 2004
Obtention et renouvellement du titre - A05	Attestation SS - Obtention et renouvellement du titre - A05
Base de données	Base de données
Comité de gestion	Comité de gestion
Comité des stages	Comité des stages
[Sigle de cours 1]	Comité des stages - [Sigle de cours 1]
[Sigle de cours 2]	Comité des stages - [Sigle de cours 2]
Etc.	Etc.

Cette manipulation s'appuie empiriquement sur les caractéristiques structurelles des schémas de classification personnels observés et sur les témoignages de plusieurs employés. Les schémas de classification personnels observés ont des structures peu profondes avec une moyenne de 1.67 niveaux sous chaque classe principale (écart-type = 0.46). Plusieurs employés ont cependant déclaré en entrevue avoir « trop de sous-dossiers avant d'arriver au document » et envisageaient de limiter le nombre de niveaux. Un coordonnateur de stages a délibérément conçu un schéma de classification à un seul niveau (sans classes subordonnées), en vue d'éviter « trop de manipulation [dans l'ouverture des répertoires] ». Il a également exprimé le désir de pouvoir « d'un clic arriver tout de suite à l'endroit où il aimerait aller » et d'éviter d'« aller à une place, à une place, à une place ». L'importance d'accéder rapidement à l'information en contexte de travail a été relevée par Jones (2007).

D'un point de vue théorique, l'utilisation d'une liste alphabétique de classes pour organiser et classer les documents ne faciliterait pas le repérage selon Lin & Chan (1999). En effet, une structure hiérarchique ou arborescente a l'avantage de montrer explicitement les relations entre les différentes classes, ce qu'une liste alphabétique ou aléatoire de termes ne révèle pas. Une structure arborescente offre aussi la possibilité de

naviguer systématiquement de sujets généraux à sujets spécifiques, ce qui permet d'augmenter la précision du repérage (Chan, 1995; Lin & Chan, 1999). Par ailleurs, la manipulation de la structure qui a eu pour conséquence d'augmenter considérablement la longueur des intitulés de classes suggère une augmentation de la fonction « informative » de ces intitulés (Saffady, 2002). L'utilisation *a priori* contraignante d'un schéma de classification présenté sous la forme d'une liste alphabétique de classes combinée à l'usage d'intitulés de classes plus précis suggère la possibilité d'obtenir des résultats imprévisibles dans le cadre de la simulation du repérage.

La création artificielle du schéma de classification C visait à permettre l'étude de l'influence de l'absence d'une structure arborescente et l'influence de la sémantique sur le repérage des documents.

3.3.2.3 Caractéristiques du schéma de classification manipulé dans ses dimensions logique et structurelle (Schéma D)

L'objectif de cette manipulation était de modifier la structure et la logique du schéma A tout en conservant la sémantique initiale. Cette manipulation relativement simple à réaliser consistait à utiliser le schéma de classification préalablement manipulé aux plans logique et structurel (i.e., schéma E) tout en maintenant la sémantique personnelle de l'employé A. Un extrait de la manipulation réalisée est présenté dans le Tableau 3.4.

Tableau 3.4 Schéma D – Manipulation de la structure et de la logique

Schéma original (Schéma A)	Schéma manipulé aux plans logique et structurel (Schéma D)
Accueil des étudiants sept. 2005	Accueil des étudiants sept. 2005
Attestation SS	01 Attestation SS (...)
Documents officiels	03 Documents officiels
Envoi avril 2004	04 Envoi
Envoi janvier 2003	01 Envoi janvier 2003
Envoi Octobre 2004	02 Envoi Avril 2004
Obtention et renouvellement du titre - A05	03 Envoi Octobre 2004
Base de données	05 Obtention et renouvellement du titre (...)
Comité de gestion	04 Comités
Comité des stages	01 Comité des stages
[Sigle de cours 1]	[Sigle de cours 1] (...)
[Sigle de cours 2]	02 Comité de gestion
Etc.	Etc.

La manipulation des caractéristiques logiques et structurelles du schéma A s'appuie à la fois sur des observations empiriques et des bases théoriques.

L'observation empirique des schémas de classification personnels révèle qu'aucun critère de division logique ne prédomine (mixité des concepts représentés) même si on note un usage privilégié de classes thématiques. D'un point de vue théorique, l'utilisation de plusieurs critères ou caractéristiques de division à chaque niveau de la hiérarchie diminuerait la cohérence et ne permettrait pas une recherche méthodique à l'intérieur de la structure (Maniez, 1987; Fallis & Mathiesen, 2000; Hunter, 2000). La création artificielle d'un schéma de classification qui privilégie un regroupement par classes d'activités, tel que préconisé par les archivistes pour l'organisation des documents administratifs (Couture & Rousseau, 1982; Roberge, 1985; Héon, 1995), visait à vérifier l'influence de la dimension logique sur l'efficacité du repérage de documents ciblés.

L'observation empirique des caractéristiques structurelles des schémas de classification personnels montre qu'il existe de grandes différences dans l'étendue de la macro-structure (classes principales) des schémas de classification personnels, laquelle peut varier de 6 à 105 classes principales (Voir Annexe 9 Caractéristiques des schémas

de classification personnels observés). L'analyse des résultats (Voir Section 4.2.1.3 Synthèse) révèle des structures peu étendues plus précisément au niveau de la micro-structure (classes subordonnées) et non pas au niveau de la macro-structure (classes principales). La manipulation des caractéristiques structurelles visait à combler cette lacune et à répondre au besoin d'examiner empiriquement l'influence d'une macro-structure peu étendue (par exemple, huit classes principales) sur l'efficacité du repérage de documents ciblés.

La création artificielle du schéma de classification D visait à permettre l'étude de l'influence d'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique personnelle est celle de l'employé A sur l'efficacité du repérage de documents ciblés.

3.3.2.4 Caractéristiques du schéma de classification manipulé dans ses dimensions logique, structurelle et sémantique (Schéma E)

L'objectif de cette manipulation était de modifier simultanément l'ensemble des caractéristiques logique, structurelle et sémantique du schéma A pour produire un schéma de classification davantage conforme aux principes théoriques de conception d'un schéma de classification des documents administratifs (par exemple, critère de division basé sur les activités de l'employé, peu d'abréviations). Un extrait de la manipulation réalisée par la chercheuse est présenté dans le Tableau 3.5.

Tableau 3.5 Schéma E – Manipulation de la structure, de la logique et de la sémantique

Schéma original (Schéma A)	Schéma manipulé aux plans structurel, logique et sémantique (Schéma E)
Accueil des étudiants sept. 2005	Accueil des étudiants à la rentrée Automne 2005
Attestation SS	01 Attestation des superviseurs de stage (...)
Documents officiels	04 Attestations reçues
Envoi avril 2004	01 Liste de janvier 2003
Envoi janvier 2003	02 Liste d'avril 2004
Envoi Octobre 2004	03 Liste d'octobre 2004
Obtention et renouvellement du titre - A05	05 Obtention et renouvellement de l'attestation
Base de données	Automne 2005
Comité de gestion	(...)
Comité des stages	04 Comités
[Sigle de cours 1]	01 Comité des stages
[Sigle de cours 2]	1er stage (1ère année) (...)
Etc.	Etc.

La manipulation réalisée aux plans structurel et logique ayant abouti à la création du schéma E s'appuie sur les mêmes observations empiriques et justifications théoriques présentées dans la section 3.3.2.3 Caractéristiques du schéma de classification manipulé dans ses dimensions logique et structurelle (Schéma D). Autrement dit, nous voulions étudier l'influence des caractéristiques d'un schéma de classification qui privilégie une macro-structure peu étendue et un critère de division dominant, soit un regroupement par classes d'activités.

Cependant, le schéma E diffère du schéma D au plan sémantique. Nous avons substitué une sémantique institutionnelle plus explicite à la sémantique personnelle initiale de l'employé A en privilégiant des titres intelligibles (par exemple, peu abrégés) (Hertzum, 1996; Lin & Chan, 1999). Le schéma de classification ainsi conçu a été validé auprès d'un archiviste de la Division des archives de l'Université de Montréal.

La création artificielle d'un schéma de classification identique au schéma D sauf au plan de la sémantique visait à permettre l'étude de l'influence d'un schéma de

classification dont la macro-structure est peu étendue, la logique est partiellement basée sur des classes d'activités et la sémantique utilise peu d'abréviations sur l'efficacité du repérage de documents ciblés.

3.3.2.5 Synthèse

Les quatre schémas de classification développés par manipulations du schéma A présentent des caractéristiques structurelles, logiques et sémantiques suffisamment variées et discriminantes pour pouvoir être évalués. Le Tableau 3.6 résume les principales caractéristiques structurelles, logiques et sémantiques du schéma de classification personnel retenu pour des fins de manipulation et de simulation (Schéma A) ainsi que les caractéristiques des quatre schémas de classification artificiellement créés (Schémas B, C, D et E).

Tableau 3.6 Caractéristiques structurelles, logiques et sémantiques des schémas de classification A, B, C, D et E

	Schéma personnel d'un employé	Schémas de classification artificiellement créés			
	Schéma A	Schéma B	Schéma C	Schéma D	Schéma E
Structure					
Étendue	Macro-structure étendue (42 classes)	Macro-structure étendue (42 classes)	Macro-structure très étendue (192 classes)	Macro-structure peu étendue (8 classes)	Macro-structure peu étendue (8 classes)
Profondeur	Peu profonde (1.56 niveaux)	Peu profonde (1.56 niveaux)	n/a	Moyennement profonde (2.62 niveaux)	Moyennement profonde (2.62 niveaux)
Complexité	192 classes	192 classes	192 classes	202 classes	202 classes
Équilibre	Peu équilibrée (Écart-type=6.90)	Peu équilibrée (Écart-type=6.90)	n/a	Équilibrée (Écart-type=2.32)	Équilibrée (Écart-type=2.32)
Logique					
Premier critère de division logique	Thème	Thème	Mélange	Activité	Activité
Deuxième critère de division logique	Mélange	Mélange	n/a	Type de contenu	Type de contenu
Ordre de succession des classes	Alphabétique	Alphabétique	Alphabétique	Numérique	Numérique
Sémantique					
Nombre moyen de caractères dans l'intitulé des classes	19.12	19.12	37.69	19.33	20.47
Proportion des classes abrégées	32%	32%	48%	28%	1%
Principales stratégies d'abréviation	Acronyme (10% des classes)	Acronyme (10% des classes)	Mélange (17% des classes)	Code (7%) et Mélange (7%)	Acronyme (1% des classes)
Proportion des classes redondantes	35.05%	35.05%	90%	29.70%	10.89%
Définition des classes	Non	Oui	Non	Non	Non

Caractéristiques structurelles

Comme l'indique le Tableau 3.6, le groupe de schémas de classification formé par les schémas A et B possède des caractéristiques structurelles identiques soit une structure étendue à la base, complexe mais peu profonde et peu équilibrée.

Le groupe de schémas de classification formé par les schémas D et E possède également des caractéristiques structurelles identiques soit une structure moins étendue à la base mais plus profonde et plus équilibrée que les schémas A et B tout en ayant un

nombre total de classes similaire. Quant au schéma de classification C, il se caractérise essentiellement sur le plan structurel par une longue liste alphabétique de classes.

Caractéristiques logiques

Les cinq schémas de classification présentent des critères de division différents au premier niveau hiérarchique. Alors que le groupe formé des schémas de classification A et B privilégie un regroupement par *thème*, le second groupe formé des schémas de classification D et E privilégie un regroupement par *activité*. Enfin, le schéma C privilégie des critères de division qui combinent ou *mélangent* plusieurs concepts (par exemple, des *activités* et des *types de contenu*) pour la très grande majorité de ses classes.

Les cinq schémas de classification présentent un ordre de succession des classes principales différent. Alors que les schémas de classification A, B et C privilégient un ordre exclusivement alphabétique, le groupe formé des schémas de classification D et E privilégie un ordre exclusivement numérique.

Caractéristiques sémantiques

Les cinq schémas de classification présentent des noms de classes d'une longueur moyenne (en nombre de caractères) variant de 19 (Schémas A, B, D et E) à un peu plus de 37 caractères (Schéma C).

Comme l'indique le Tableau 3.6, les schémas A, B, C et D, basés sur la sémantique personnelle de l'employé A, privilégient l'usage des *acronymes* comme principale stratégie d'abréviation. Par ailleurs, l'usage des acronymes est très faible dans le schéma E.

La proportion moyenne des classes redondantes au niveau structurel et conceptuel est différente selon le schéma de classification. La redondance structurelle et conceptuelle est plus élevée dans le schéma C. Cette redondance est essentiellement provoquée par l'utilisation d'une dénomination très spécifique des classes dès le premier niveau de la hiérarchie. Au contraire, la redondance structurelle et conceptuelle est

délibérément plus faible, voire inexistante, pour les classes de premier niveau dans les schémas D et E.

Le protocole de manipulation a permis la création artificielle de schémas de classification ayant des caractéristiques structurelles, logiques et sémantiques variées afin de répondre à la question générale suivante : « Peut-on repérer avec la même efficacité un document électronique à l'aide de schémas de classification ayant des caractéristiques variées? ». Les questions spécifiques de recherche étaient les suivantes :

Existe-t-il des différences sur le plan de la capacité à repérer tous les documents ciblés à l'aide de schémas de classification ayant des caractéristiques variées?

Existe-t-il des différences sur le plan de la capacité à repérer rapidement les documents ciblés à l'aide de schémas de classification ayant des caractéristiques variées?

Existe-t-il des différences sur le plan de la capacité à repérer les documents ciblés dès le premier essai à l'aide de schémas de classification ayant des caractéristiques variées?

La section suivante présente nos hypothèses pour chacune de ces questions de recherche.

3.4 Hypothèses de recherche

L'hypothèse générale (HG) que nous souhaitons vérifier dans le cadre de notre recherche a été formulée ainsi :

HG. : Il existe des différences sur le plan de l'efficacité du repérage des documents électroniques réalisé à l'aide de schémas de classification distincts aux plans structurel, logique et sémantique.

Les hypothèses spécifiques (HS) sont les suivantes :

HS1. Un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus de documents.

HS2. Un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus rapidement les documents.

HS3. Un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus de documents dès le premier essai.

Ces hypothèses ont été testées dans le cadre d'un protocole de simulation conçu pour mesurer l'efficacité du repérage de documents réalisé à l'aide de schémas de classification ayant des caractéristiques variées. La section suivante présente cette dernière étape de notre protocole de recherche.

3.5 Simulation du repérage

Nous avons retenu une approche quantitative exploratoire pour répondre aux questions de recherche et vérifier les hypothèses. Les méthodes quantitatives sont régulièrement utilisées en sciences de l'information pour mesurer et comparer la performance des systèmes de repérage de l'information (Cleverdon, 1967; Krikelas, 1969; Lytle, 1980; Hertzum & Frøkjær, 1996; Ribeiro, 1996; Hildreth, 2000). L'approche quantitative exploratoire représente un choix pertinent pour cette recherche qui étudie l'efficacité du repérage de documents électroniques à l'aide de différents schémas de classification. Cette approche est adéquate car elle nous a permis d'introduire un « élément d'objectivité », de décrire avec une plus grande précision l'efficacité du repérage réalisé à partir de différents schémas de classification des documents électroniques, et d'effectuer des comparaisons par une mesure uniforme du phénomène (Trudel & Antonius, 1991).

La technique de simulation est particulièrement appropriée dans le cadre de notre recherche. Cette technique présente l'avantage de permettre de déceler des résultats imprévisibles et met l'accent sur des découvertes (Jones, 2000). Les répondants sont amenés à se comporter comme si la situation qui leur est présentée est bien réelle et à montrer comment ils y feraient face (Jones, 2000; Gonçalves, 2003). La simulation permet également de préserver la confidentialité des documents et de protéger la vie privée des individus en évitant toute consultation des documents qui sont réellement classés dans l'environnement documentaire.

L'utilisation d'une approche par groupes équivalents de répondants et groupes équivalents de documents est apparue appropriée et pertinente dans le cadre de notre recherche. Cette approche nous a permis de neutraliser les facteurs humains (par exemple, habileté à repérer plus ou moins rapidement un document, effet de fatigue, effet d'apprentissage) et documentaires (par exemple, le niveau de difficulté de repérage d'un document particulier). L'approche par groupe est fréquemment utilisée dans les recherches qui font appel à des agents humains pour mesurer et comparer la performance de différents systèmes de repérage et de recherche d'information (Lytle, 1980; Hertzum & Frøkjær, 1996; Ribeiro, 1996; Hildreth, 2000).

La section suivante présente la procédure d'échantillonnage et la taille de l'échantillon ainsi que les méthodes de collecte et d'analyse de données utilisées dans le cadre de cette dernière étape de la recherche.

3.5.1 Procédure d'échantillonnage et taille de l'échantillon

Dans le cadre de la simulation, cinq schémas de classification (i.e., schémas A, B, C, D et E) ont été utilisés par 70 répondants pour repérer 20 documents. Tous les répondants étaient des étudiants ou des diplômés de l'Université de Montréal, principalement dans le domaine des sciences de l'information. Les sections suivantes décrivent le processus de sélection des documents et justifient le choix de cette catégorie particulière de répondants.

3.5.1.1 Sélection des documents

Les 20 documents ont été sélectionnés aléatoirement lors de la rencontre individuelle avec l'employé propriétaire du schéma de classification retenu comme base pour la construction de schémas artificiels (Schéma A). Le caractère aléatoire de cette sélection a permis de réduire l'influence de variables étrangères (par exemple, la difficulté à repérer un document en particulier).

Le choix du nombre de documents est basé sur une justification empirique plutôt que théorique. Ainsi, le choix de cinq documents à repérer a été empiriquement jugé insuffisant par Chapman (1999) pour obtenir des résultats significatifs. Lytle (1980), quant à lui, a utilisé 15 questions (ou requêtes) pour évaluer deux systèmes de repérage de documents d'archives. Un nombre de plusieurs dizaines de documents est habituellement jugé significatif dans le domaine de la recherche d'information (RI) pour tester les systèmes. Dans le cadre de notre recherche, nous avons retenu le nombre de 20 documents, soit quatre documents à repérer par schéma de classification. Le pré-test de la simulation a permis de constater qu'un nombre de documents supérieur à 20 aurait eu comme corollaire d'augmenter la durée de la simulation et le niveau de fatigue des répondants.

3.5.1.2 Sélection des répondants

Le choix de l'École de bibliothéconomie et des sciences de l'information comme principal bassin de recrutement des répondants a été guidé par la nécessité de contrôler les variables étrangères (par exemple, l'expérience professionnelle d'un individu ou sa capacité à repérer plus rapidement les documents) en visant l'homogénéité du groupe de répondants. La sélection des répondants a été réalisée sur la base de quatre critères : (1) la connaissance du milieu universitaire, (2) la motivation à participer à la recherche, (3) l'habileté à naviguer dans une structure arborescente de répertoires dans l'environnement Windows, et (4) l'habileté à travailler avec des descriptions de documents. Ces considérations sont développées dans la présente section.

Idéalement, il aurait été préférable de demander à des employés (par exemple, les coordonnateurs de stages eux-mêmes) de participer à la simulation pour se rapprocher des conditions réelles de repérage. Il nous a semblé *a priori* difficile de recruter un nombre aussi important de volontaires parmi les employés en raison de contraintes de temps exprimées par les employés eux-mêmes mais aussi parce que le nombre de coordonnateurs de stages identifiés et volontaires (i.e., 25) aurait été insuffisant. Notre choix s'est plutôt porté sur des étudiants ou diplômés de l'Université de Montréal en raison de la possibilité de recruter un plus grand nombre de volontaires, de leur connaissance du milieu universitaire et de leur rapport quasi-quotidien avec l'institution dans le cadre de leurs études.

Le pré-test auprès de deux volontaires a montré qu'une séance de simulation du repérage de 20 documents pouvait varier entre une heure et une heure et demie. Étant dans l'impossibilité de rémunérer les répondants, il était nécessaire de recruter des volontaires motivés à participer bénévolement à notre expérimentation. Les étudiants et diplômés de l'École de bibliothéconomie et des sciences de l'information se sont montrés particulièrement motivés à participer à notre recherche qui touche leur domaine de formation, mais aussi par simple altruisme, en vue d'aider une collègue étudiante dans le cadre de sa recherche doctorale dans le même département que le leur.

Le choix de répondants en provenance du domaine des sciences de l'information avait aussi l'avantage de neutraliser l'impact du niveau d'habileté des répondants à naviguer à travers l'Explorateur Windows pour repérer des documents. Les travaux de Lytle (1980) ont montré qu'il existe une relation entre la performance du repérage (« *retrieval performance scores* ») et l'expérience ou familiarité du répondant (« *searcher* ») avec la méthode de repérage utilisée. Le système d'exploitation Windows, utilisé par les employés ayant participé à la recherche pour organiser leurs documents électroniques, est également un environnement informatique familier pour les étudiants et diplômés de l'École de bibliothéconomie et des sciences de l'information en raison de son utilisation prédominante dans la réalisation de travaux pratiques en laboratoires et d'autres travaux académiques.

Le choix de répondants en provenance du domaine des sciences de l'information permettait enfin de s'assurer de contrôler l'impact de l'habileté des répondants à travailler avec des descriptions de documents. Les étudiants et diplômés en sciences de l'information sont formés pour gérer l'objet « document » et pour travailler à partir des descriptions de documents dans le cadre d'activités d'indexation, de condensation, de description ou de classification, par exemple. Le pré-test a permis de constater avec quelle facilité ce type de répondant particulier pouvait effectuer la tâche *a priori* complexe de lecture et d'analyse de la description d'un document, puis d'identification de l'emplacement du document en utilisant un schéma de classification.

3.5.2 Collecte des données

Conformément aux exigences de l'Université de Montréal relatives aux aspects éthiques de la recherche, nous avons demandé et obtenu l'autorisation de réaliser cette collecte de données auprès du Comité d'éthique de la Faculté des arts et des sciences de l'Université de Montréal. Le certificat d'éthique ainsi que le formulaire de consentement que chaque répondant devait lire et signer avant que la chercheure débute la simulation sont reproduits en annexe (Voir Annexe 1 Certificat d'éthique et formulaires de consentement).

3.5.2.1 Phase de la simulation

Un appel à participation a été lancé en avril 2006 par le biais des listes de courriel des étudiants et des listes professionnelles. La collecte des données a été réalisée entre la fin du mois d'avril et le début du mois de juillet 2006. Les répondants ont été invités à indiquer à quelle période du jour, du soir et de la semaine ils étaient disponibles. La grande majorité des répondants (80 %) (n=56) ont réalisé la simulation le jour, entre 9h et 17h. Certains répondants (20 %) (n=14) ont réalisé la simulation le soir (i.e., entre 17h et 22h), à la sortie de leur travail ou de leur stage.

3.5.2.2 Information préliminaire donnée aux répondants

Les étudiants et les diplômés ont reçu une lettre envoyée par messagerie électronique présentant brièvement le projet, les tâches à réaliser durant la simulation et sollicitant leur participation pour une rencontre individuelle d'une durée d'une heure et demie à deux heures (Voir Annexe 10 Lettre d'appel à participation – Étudiants et diplômés). Pour pouvoir participer, les étudiants et diplômés devaient être capables d'utiliser une souris et de naviguer à travers une structure arborescente de répertoires pour repérer des documents électroniques dans un système d'exploitation Windows. Les étudiants et diplômés étaient avisés que la simulation serait enregistrée pour collecter leurs commentaires généraux sur les schémas de classification utilisés. Les étudiants et diplômés ainsi contactés étaient également invités à répondre à des questions d'ordre socio-démographiques (i.e., âge, niveau d'étude, formation académique suivie). Les étudiants et diplômés étaient invités à manifester leur intérêt à participer à la recherche en communiquant avec la chercheuse par courriel ou par téléphone.

3.5.2.3 Profil des répondants

Le groupe de 70 répondants ayant participé volontairement à cette expérimentation était composé de 20 hommes (28.5 %) et de 50 femmes (71.5 %). L'âge moyen était de 32 ans (écart-type = 7.38). Six répondants (8.5 %) étaient des étudiants de 1^{er} cycle, 27 répondants (38.5 %) étaient des étudiants de 2^{ème} cycle et 10 répondants (14.5 %) étaient des étudiants de 3^{ème} cycle de l'École de bibliothéconomie et des sciences de l'information. Vingt-cinq répondants (35.5 %) étaient des diplômés de l'École de bibliothéconomie et des sciences de l'information. En vue d'obtenir un nombre égal de répondants par groupe (i.e., cinq répondants pour 14 groupes), nous avons sélectionné un étudiant en mathématiques et un diplômé en psychologie de l'Université de Montréal qui répondaient aux critères d'inclusion (2.85 % des répondants).

3.5.2.4 Procédure de simulation

Préparation de la simulation

Nous avons reproduit sur un portable HP doté du système d'exploitation Windows XP professionnel cinq structures arborescentes des répertoires correspondant au schéma de classification de l'employé (Schéma A) et aux quatre schémas de classification artificiellement créés (Schémas B, C, D, E). Ces structures de répertoires directement placées sur le Bureau étaient intitulées respectivement : « 111 » (Schéma A), « 113 » (Schéma B), « 311 » (Schéma C), « 331 » (Schéma D) et « 333 » (Schéma E).

Nous avons conçu 14 grilles d'observation ou feuilles de réponse (Voir Annexe 11 Exemple de grille d'observation), soit une grille pour chaque groupe de répondants. Ces grilles furent utilisées pour consigner, pour chacun des documents à repérer, le nombre d'essais nécessaires pour repérer le document, le temps pris pour chacun des essais, et le score du repérage (succès, échec ou abandon, le cas échéant). Le nom du répertoire où était réellement classé le document dans le schéma original conçu par l'employé était indiqué en caractères de couleur pâle sur la feuille de réponse afin que seule la chercheuse puisse en faire la lecture et évaluer la qualité de la réponse du répondant (i.e., identification du répertoire adéquat ou non).

Nous avons retranscrit les descriptions des 20 documents telles que communiquées par l'employé utilisateur du schéma de classification personnel A (Voir Annexe 13 Description des documents à repérer). L'ensemble des descriptions a été regroupé dans un cartable, chaque description étant retranscrite sur une seule page.

Composition des groupes

Chaque individu qui répondait aux critères d'inclusion a été assigné aléatoirement¹² à l'un des 14 groupes afin de diminuer l'influence que des variables étrangères à la recherche peuvent avoir sur les variables étudiées (Easterby-Smith, Thorpe & Lowe, 1993; Fortin, Côté & Filion, 2006). Le Tableau 3.7 présente les

¹² Au fur et à mesure que les répondants se présentaient à la simulation, ils furent assignés séquentiellement à l'un des 14 groupes.

caractéristiques socio-démographiques des répondants pour chacun des groupes expérimentaux, considérés comme équivalents.

Tableau 3.7 Caractéristiques socio-démographiques des répondants pour chaque groupe expérimental

	Âge moyen	Femmes (n=50)	Hommes (n=20)	Étudiants (n=43)	Diplômés (n=27)	Formation en Sciences de l'information (SI) (n=68)
Groupe 1	32	5	-	4	1	5
Groupe 2	30	1	4	5	-	5
Groupe 3	30	4	1	4	1	5
Groupe 4	33	3	2	3	2	5
Groupe 5	35	3	2	4	1	5
Groupe 6	26	4	1	5	-	5
Groupe 7	34	4	1	4	1	5
Groupe 8	35	4	1	3	2	5
Groupe 9	36	4	1	-	5	4
Groupe 10	31	4	1	1	4	5
Groupe 11	32	3	2	2	3	5
Groupe 12	31	5	-	3	2	5
Groupe 13	38	3	2	1	4	4
Groupe 14	34	3	2	4	1	5

Tous les répondants ont repéré les 20 documents dans le même ordre en alternant l'utilisation des différents schémas de classification. Pour ce faire, les 20 documents ont été regroupés dans cinq blocs, chaque bloc comprenant quatre documents. Le premier bloc de quatre documents a été repéré avec l'un des cinq schémas de classification (par exemple, le schéma A), le deuxième bloc avec l'un des quatre autres schémas de classification (par exemple, le schéma C), le troisième bloc avec l'un de trois autres schémas de classification (par exemple, le schéma B), etc.

L'utilisation des schémas de classification dans un ordre qui variait systématiquement selon le groupe d'appartenance a permis de neutraliser les effets d'apprentissage et de fatigue (Voir Tableau 3.8) (Lytle, 1980; Hertzum & Frøkjær, 1996). Par exemple, tous les répondants du groupe 1 ont utilisé systématiquement d'abord le schéma A suivi du schéma C, du schéma D, du schéma B et du schéma E;

tous les répondants du groupe 2 ont utilisé d'abord le schéma C, suivi du schéma A, du schéma E, du schéma D pour terminer avec le schéma B.

Tableau 3.8 Protocole de simulation basé sur le protocole expérimental de Hertzum & Frøkjær (1996, p. 146)

70 répondants	14 groupes	Bloc 1 : d1 ... d4	Bloc 2 : d5 ... d8	Bloc 3 : d9 ... d12	Bloc 4 : d13 ... d16	Bloc 5 : d17 ... d20
Répondant 1	Groupe 1	Schéma A	Schéma C	Schéma D	Schéma B	Schéma E
Répondant 2						
Répondant 3						
Répondant 4						
Répondant 5						
Répondant 6	Groupe 2	Schéma C	Schéma A	Schéma E	Schéma D	Schéma B
Répondant 7						
Répondant 8						
Répondant 9						
Répondant 10						
:	:			:		
:	:			:		
Répondant 66	Groupe 14	Schéma E	Schéma C	Schéma D	Schéma A	Schéma B
Répondant 67						
Répondant 68						
Répondant 69						
Répondant 70						

Ce protocole qui permet de contrôler le facteur humain est basé sur le devis expérimental de Hertzum & Frøkjær (1996) visant à étudier de manière comparable l'efficacité du repérage réalisé à l'aide de trois modes différents d'organisation de l'information. Notre protocole a permis de s'assurer que les mesures objectives de performance reflètent les différences entre les schémas de classification seulement (et non les individus). La seule condition que ce protocole devait respecter était une permutation relativement homogène de l'ordre d'utilisation des schémas de classification (par exemple, ne pas commencer toujours avec le même schéma) tout en ne permettant pas d'utiliser le schéma B (dont la sémantique est dévoilée) avant le schéma A (schéma identique excepté sur le plan sémantique).

Description d'une séance de simulation

En vue de minimiser les distractions (Hildreth, 2000), toutes les séances de simulation ont eu lieu dans le même environnement, soit un local isolé de l'École de bibliothéconomie et des sciences de l'information de l'Université de Montréal, dans des conditions identiques de bruit et de température. Seule la lumière différait selon l'heure de la journée. Étant donné le grand nombre de répondants et leur répartition à l'intérieur de différents groupes, l'influence des facteurs environnementaux sur la simulation n'est pas significative.

Chaque répondant a été invité à lire les instructions écrites qui expliquaient le déroulement de la séance de simulation (Voir Annexe 12 Directives communiquées aux répondants au début de la séance de simulation). La chercheuse a utilisé sa propre structure arborescente de répertoires dans l'Explorateur Windows pour illustrer les tâches à réaliser. Les seules habiletés exigées pour les besoins de la simulation étaient la capacité de chaque répondant à utiliser une souris et à naviguer dans la structure arborescente de l'Explorateur Windows (i.e., utilisation de la barre de défilement avec la souris, ouverture et fermeture de l'arborescence).

Chaque répondant a été averti qu'il s'agissait d'une simulation. Autrement dit, il devait « faire semblant » d'être en situation de recherche de document. Pour chacun des documents électroniques à repérer, chaque répondant devait indiquer à la chercheuse le répertoire dans lequel serait localisé, selon lui, le document cible. Chaque répondant a été informé que parmi les cinq schémas de classification, un seul était réellement utilisé par un employé de l'Université de Montréal et que les quatre autres schémas étaient des variantes artificiellement créées pour tester différentes hypothèses. Chaque répondant a été avisé que cette mise en situation était artificielle pour deux raisons. Premièrement, le répondant n'avait pas accès au contenu des documents eux-mêmes mais devait plutôt se baser sur une description écrite et formulée par l'employé lui-même. La raison était que certains documents sont nominatifs ou renferment de l'information confidentielle et ne peuvent être rendus publics. Deuxièmement, les cinq schémas qui ont été utilisés sont tous « vides »; aucun fichier ou document décrit n'étant physiquement organisé à

l'intérieur des cinq structures arborescentes de répertoires. Chaque répondant devait se baser uniquement sur les noms des classes et sur l'organisation des classes pour prendre sa décision.

Les répondants ont été avertis que certains schémas de classification étaient plus ou moins compréhensibles que d'autres et que c'était normal qu'ils aient plus ou moins de difficultés dans l'utilisation ou la compréhension d'un schéma particulier. Il en était de même pour les documents : le pré-test a montré que certains documents pouvaient être plus ou moins faciles à repérer que d'autres. Les répondants ont été bien informés qu'il ne s'agissait pas d'un test d'intelligence ni d'une évaluation de leur habileté personnelle à repérer un document parmi d'autres; l'évaluation portait uniquement sur l'intelligibilité des différents schémas de classification utilisés pour le repérage des documents ciblés.

Les répondants ont été informés que les documents et l'un des cinq schémas de classification étaient utilisés par un employé de l'Université de Montréal qui occupe les fonctions de coordonnateurs de stages. Pour des raisons de confidentialité, les noms réels de l'employé et de l'unité dans laquelle il travaillait n'ont pas été dévoilés.

Comme les individus organisent plus ou moins intuitivement leurs documents selon leurs activités, les répondants ont lu la description des tâches du coordonnateur de stages. Les répondants ont été avertis de l'importance de bien comprendre ce contexte car certaines descriptions de documents et les intitulés de certains répertoires reflètent les fonctions et les tâches de l'employé. Certaines informations additionnelles relatives aux tâches d'un coordonnateur de stages ont été apportées par la chercheuse comme la définition de superviseur de stage (i.e., professionnel distinct du coordonnateur de stages qui est chargé de l'encadrement de l'étudiant durant son stage) ou du concept d'assignation des stages (i.e., processus qui consiste à assigner des étudiants en stage).

Avant chaque première utilisation d'un schéma de classification et avant même la lecture de la première description du document à repérer avec ce schéma, chaque répondant a été invité à examiner la macro-structure (classes principales), sans possibilité de dévoiler la hiérarchie. Le but était de permettre au répondant d'avoir une première

vue d'ensemble du schéma de classification avant même de commencer le processus de repérage des documents.

Chaque répondant a été informé des données qui seraient colligées durant la simulation, soit (1) le temps requis pour le repérage, (2) l'abandon, le succès ou l'échec du repérage, (3) le nombre d'essais complétés.

Avant le repérage de chaque document, une description sommaire du document à repérer a été communiquée par écrit à chacun des répondants (Voir Annexe 13 Description des documents à repérer). Les répondants ont eu la possibilité de consulter à tout moment cette description durant le processus de repérage et pouvaient poser des questions de clarification à la chercheuse. Les répondants ont été avertis que durant toute la séance de simulation la chercheuse ne pouvait répondre à aucune de leurs questions concernant les schémas de classification (par exemple, la signification des classes).

Au moment où le répondant cessait de lire la description et commençait à regarder l'écran, le chronomètre était lancé. Ce dernier était arrêté au moment où le répondant cliquait sur un répertoire ou un sous-répertoire et déclarait à vive voix : « C'est lui! » ou « Il [le document] est là-dedans! ». Les répondants disposaient d'un temps illimité pour repérer un document.

Lorsque le répertoire qui avait été identifié par le répondant était le « bon répertoire » autrement dit celui qui était réellement utilisé pour classer le document, la chercheuse notait sur une grille d'observation le temps en secondes qui avait été mesuré depuis le début de la recherche du répertoire jusqu'à l'identification verbale et on répétait la procédure pour le document suivant. Si le répertoire identifié par le répondant n'était « pas le bon », le répondant pouvait réaliser deux autres essais ou décider d'abandonner. Les répondants étaient encouragés à utiliser les trois essais plutôt que d'abandonner. Toutefois, ils avaient été informés qu'il était toujours préférable d'abandonner plutôt que de désigner un répertoire au hasard ce qui invaliderait les données.

Pendant que la chercheuse consignait les données entre chaque essai (i.e., temps pris), le répondant était invité à appliquer une technique utilisée en psychologie cognitive qui consiste à compter à rebours à haute voix des chiffres pairs ou impairs (Lemaire, 1999). Le répondant pouvait également choisir de lire un journal à haute voix. Chaque répondant avait été bien informé qu'il ne s'agissait pas d'une « punition » pour ne pas avoir repéré le document dès le premier essai mais bien d'une technique visant à l'empêcher de réfléchir sur la tâche qui lui était demandée en détournant son attention vers une autre tâche cognitive plus ou moins complexe. L'utilisation de cette technique a permis de s'assurer que le répondant ne continuait pas à chercher mentalement alors que la mesure du temps n'était plus possible durant la prise de note par la chercheuse. Après avoir colligé la mesure du temps, le chronomètre était remis à zéro.

Chaque répondant a été avisé qu'il avait 20 documents à repérer en utilisant successivement cinq schémas de classification différents. Pour ce faire, tous les quatre documents, le répondant a utilisé un schéma différent désigné par la chercheuse. Conformément au protocole, l'ordre d'utilisation des schémas de classification variait selon le groupe auquel était assigné le répondant. L'ordre de présentation des descriptions de document était le même pour tous les groupes.

À la fin de l'utilisation de chacun des schémas de classification, autrement dit tous les quatre documents, chaque répondant était invité à donner son appréciation générale du schéma. Le repérage des documents reprenait ensuite avec le schéma de classification suivant. Toujours avec l'accord des répondants, chaque séance de simulation fut enregistrée afin de s'assurer de conserver l'intégralité de leurs propos.

Le pré-test de la séance de simulation du repérage réalisé auprès de deux volontaires nous a permis de raffiner le protocole de simulation en introduisant des éléments de contrôle (par exemple, le comptage à rebours) et de précision (par exemple, indiquer de vive voix le choix du répertoire afin de stopper le chronomètre). Nous avons aussi augmenté le contrôle de la lecture de la description des documents à repérer en ne communiquant aux répondants qu'une description à la fois. Cette méthode a permis

d'éviter que les répondants puissent lire la description suivante sur la liste et commencent à repérer deux documents à la fois.

Durée d'une séance de simulation

Comme l'indique le Tableau 3.9, la durée moyenne d'une séance de simulation (comprenant l'explication de la procédure et le repérage des documents proprement dit) a été d'une heure et vingt minutes.

Tableau 3.9 Durée moyenne de la simulation (en minutes)

Groupes	Durée de la séance de simulation (en minutes)
1	78
2	74
3	83
4	83
5	73
6	76
7	92
8	92
9	70
10	82
11	85
12	79
13	81
14	78
Moyenne	80.43
Écart-type	6.47

La durée d'une séance de simulation variait en fonction du nombre d'essais requis pour repérer un document (jusqu'à trois essais maximum), du temps pris lors de chaque essai (de quelques secondes à plusieurs minutes) et de la propension ou non des répondants à abandonner le repérage d'un document.

3.5.3 Méthode d'analyse des données

3.5.3.1 Préparation des données pour l'analyse

Les données inscrites par la chercheuse sur les feuilles de réponses ont été retranscrites dans un fichier *Excel* pour des fins d'analyse par l'équipe du Service de consultation en méthodes quantitatives (SCMEQ) de l'Université de Montréal. La préparation et la présentation des données ont été réalisées selon les directives du SCMEQ (Voir Annexe 14 Préparation des données pour le SCMEQ).

3.5.3.2 Analyse statistique

Pour tester les hypothèses de recherche, des statistiques descriptives ont été calculées et l'analyse de variance (ANOVA) à mesures répétées, communément utilisée dans la comparaison de systèmes de recherche d'information (Hertzum & Frøkjær, 1996; Prom, 2004), a été retenue et validée par le SCMEQ. Les statistiques descriptives et l'analyse de variance sont appropriées dans le cadre de notre recherche quantitative exploratoire car elles nous permettent d'analyser et de comparer, pour chaque mesure, les moyennes obtenues entre les schémas de classification et de déterminer si les différences observées sont significatives (Fortin, 2006).

L'analyse statistique a été réalisée en septembre 2006 par deux étudiants des cycles supérieurs suivant un cours de consultation en statistique sous la supervision d'un professeur titulaire en statistiques et d'un consultant professionnel en statistiques. Le tableur *Excel* et le logiciel d'analyse statistique en sciences sociales SPSS (« *Statistical package for social sciences* ») ont été utilisés comme outils d'analyse et de présentation des résultats. Les résultats de l'analyse statistique sont présentés au Chapitre 4 et discutés au Chapitre 5.

3.6 Qualité de la recherche

Plusieurs mesures ont été prises afin d'assurer la rigueur et la possibilité de réappropriation de notre recherche par d'autres chercheurs. La présente section évalue la qualité de ce projet par rapport à deux critères : la validité et la fiabilité de la recherche.

3.6.1 Validité de la recherche

La validité est une composante essentielle de la qualité d'une recherche (Fortin, Côté & Filion, 2006). Elle permet de s'assurer de la pertinence et de la rigueur des résultats (validité du construit, validité interne) ainsi que du niveau de généralisation potentielle des résultats (validité externe) (Thiétart & collaborateurs, 2003). Trois types de validité ont été considérés dans le cadre de cette méthodologie : la validité du construit, la validité interne et la validité externe des résultats de la recherche.

3.6.1.1 Validité du construit

La validité du construit est le « degré auquel une opérationnalisation permet de mesurer le concept qu'elle est supposée représenter » (Thiétart & collaborateurs, 2003, p. 259). Dans le cadre de notre recherche qui visait à caractériser des schémas de classification personnels, plusieurs mesures ont été prises pour proposer une définition opérationnelle la plus complète et précise possible du concept de schéma de classification.

Pour s'assurer de la validité du construit, un modèle d'analyse multidimensionnelle a été développé sur la base des indicateurs de mesure utilisés dans les travaux de Boardman (2001a), Boardman & Sasse (2004), Gonçalves & Jorge (2003) et Henderson (2003) portant sur l'analyse de structures arborescentes de répertoires électroniques. Les règles théoriques de construction d'un schéma de classification documentaire ont également servi de cadre de référence lors de la conception de ce modèle.

Le modèle d'analyse prend en compte non seulement les caractéristiques structurelles (par exemple, le nombre de classes) des schémas de classification mais

aussi leurs caractéristiques logiques (par exemple, le critère de division) et sémantiques (par exemple, la dénomination des classes). Ce faisant, cet outil permet une analyse quantitative et qualitative des schémas de classification.

Ce modèle d'analyse des schémas de classification a été validé dans le cadre d'un pré-test.

3.6.1.2 Validité interne

La validité interne désigne le caractère d'une recherche « dans laquelle il apparaît hors de tout doute que la variable indépendante est la seule cause du changement survenu dans la variable dépendante » (Fortin, Côté & Fillion, 2006, p. 181). La validité interne d'une méthode expérimentale implique quatre conditions : l'assurance que les effets observés sur la variable dépendante (dans notre projet, il s'agit du repérage) sont effectivement dûs à la variable indépendante (dans notre projet, les schémas de classification), la présence du concept de randomisation et de contrôle, l'absence de biais, et la présence du concept de reproductibilité du protocole expérimental (Busha & Harter, 1980). Les facteurs d'invalidité interne susceptibles d'affecter les résultats de cette expérimentation sont la sélection des répondants, l'effet de maturation, l'accoutumance au test et les fluctuations de l'instrument de mesure.

La sélection de répondants ayant un profil socio-démographique relativement homogène (i.e., uniquement des étudiants et diplômés de l'Université de Montréal) et leur répartition aléatoire dans plusieurs groupes sont des mesures qui ont eu pour effet de neutraliser les différences préexistantes entre les individus (par exemple, l'habilité à repérer des documents) et de rendre les groupes équivalents (Easterby-Smith, Thorpe et Lowe, 1993; Fortin, Côté & Fillion, 2006).

Les biais liés à l'effet de fatigue (ou effet de maturation) et à l'effet d'apprentissage (ou accoutumance au test) ont été évités grâce à l'alternance des schémas de classification testés et au nombre significatif de répondants ayant à réaliser les mêmes tâches (Lytle, 1980; Hertzum & Frøkjær, 1996). Ainsi, à mesure qu'un

répondant peu familier avec un schéma de classification prend de l'expérience, cette expérience est distribuée également à travers les schémas (Lytle, 1980).

Dans notre recherche, les fluctuations de l'instrument de mesure pouvaient être liées à la façon dont les données sont collectées. Afin de contrôler ce biais, la collecte des données a été réalisée par une seule personne, soit la chercheuse elle-même, selon un protocole écrit et décrivant précisément la procédure (Fortin, Côté & Filion, 2006). La procédure de simulation du repérage de documents a été pré-testée et a permis à la chercheuse de se familiariser avec cette technique de collecte de données.

3.6.1.3 Validité externe

La validité externe « réfère à la possibilité de généraliser les résultats d'une étude à d'autres personnes que les seuls participants et à d'autres contextes que ceux qui sont considérés dans l'étude » (Fortin, Côté & Filion, 2006, p. 182).

Notre méthodologie et le modèle d'analyse des schémas de classification personnels peuvent être appliqués à d'autres schémas de classification et à d'autres contextes. Le modèle d'analyse peut être utilisé dans toute discipline reliée aux sciences de l'information ou dans toute recherche qui s'intéresse aux schémas de classification. Ce modèle a d'ailleurs été utilisé dans le cadre d'une recherche visant à caractériser et à comparer différents répertoires des ressources du Web spécialisés dans le domaine de l'éducation (Hudon, Mas & Gazo, 2005).

Notre recherche exploratoire ne permet pas de généralisation statistique des observations relatives aux caractéristiques des schémas de classification personnels en raison de la petite taille de l'échantillon (i.e., 21 schémas de classification) et de la procédure d'échantillonnage (i.e., biais lié à un échantillon de volontaires).

La validité externe d'une recherche est fortement dépendante de sa validité interne (Cook & Campbell, 1979). L'augmentation de la validité interne entraîne le plus souvent une diminution de la validité externe du fait de l'exclusion des répondants ayant des caractéristiques différentes ou appartenant à des contextes différents (Fortin, Côté &

Filion, 2006). Dans le cadre de la simulation du repérage (i.e., approche quasi-expérimentale), nous avons opté pour une forte validité interne afin de contrôler les variables étrangères et de s'assurer que les résultats obtenus sont bien dus à la variable indépendante (i.e., schémas de classification) et non à des variables étrangères (par exemple, l'habileté de certains répondants à repérer plus facilement les documents que d'autres). Cette priorité donnée à la validité interne d'une recherche est également jugée préférable selon plusieurs chercheurs (Polit & Beck, 2004; Norwood, 2000 cités dans Fortin, Côté & Filion, 2006) car si les résultats sont invalides, il ne servirait à rien de vouloir les généraliser à d'autres contextes ou à d'autres populations (Fortin, Côté & Filion, 2006).

3.6.2 Fiabilité de la recherche

La fiabilité d'une étude « cherche à démontrer que les opérations de la recherche pourraient être répétées par un autre chercheur ou à un autre moment avec le(s) même(s) résultat(s) » (Thiétart & collaborateurs, 2003, p. 258). Cette notion de fiabilité concerne donc, tout comme la validité, différents niveaux : la fidélité des instruments de mesure et la fiabilité des résultats. Pour assurer la fiabilité de cette recherche, la chercheuse a pris différentes mesures tant au niveau de la préparation du projet que de sa réalisation.

3.6.2.1 Fidélité des instruments de mesure

La fidélité d'un instrument de mesure réfère à la capacité de l'instrument de mesurer d'une fois à l'autre un même objet de façon précise et constante (Fortin, Côté & Filion, 2006, p. 282). Deux instruments de mesure ont été utilisés dans le cadre de notre recherche : (1) un logiciel de collecte et d'analyse statistique des données relatives aux caractéristiques structurelles des schémas de classification et (2) un modèle d'analyse des caractéristiques structurelles, logiques et sémantiques des schémas de classification. Deux critères d'évaluation de la fidélité des instruments de mesure ont été utilisés dans notre recherche soit la stabilité et la fidélité interjuges.

« Un instrument est considéré comme stable quand les prises de mesure effectuées dans les mêmes conditions et auprès des mêmes individus donnent des

résultats identiques » (Fortin, Côté & Filion, 2006, p. 283). Le protocole d'utilisation du logiciel d'analyse de l'espace documentaire personnel (PDS) conçu et utilisé par Gonçalves & Jorge (2003) dans le cadre de leurs travaux a été décrit avec précision. La stabilité de ce logiciel a été pré-testée sur le poste de travail informatique de cinq employés de l'Université de Montréal pour permettre la collecte automatisée de données statistiques relatives aux caractéristiques structurelles des schémas de classification.

Le codage réalisé par la chercheuse fournit quelques-unes des données de notre recherche (i.e., critères de division logique et stratégies d'abréviation utilisées). « La fidélité interjuges réfère à la concordance des résultats enregistrés par deux observateurs indépendants ou plus qui ont utilisé les mêmes instruments de mesure auprès des participants ou observé le même événement. Elle peut aussi consister dans [sic] la concordance des observations faites par une seule personne à différents moments » (Fortin, Côté & Filion, 2006, p. 284). Dans le cadre de notre recherche, les caractéristiques logiques et sémantiques des schémas de classification ont été analysées plusieurs fois par la chercheuse à au moins trois semaines d'intervalle pour contrôler le caractère faillible de l'observation humaine (test intra codeur).

Les différents outils de collecte de données tels que les saisies d'écran, le modèle d'analyse et le logiciel ont été pré-testés auprès de cinq employés volontaires afin de s'assurer qu'ils permettaient d'obtenir de l'information de qualité.

3.6.2.2 Fiabilité des résultats

Le processus d'échantillonnage (i.e., nature et nombre de répondants, répartition aléatoire des répondants à l'intérieur des groupes), le repérage effectué par plusieurs répondants différents à l'intérieur de plusieurs groupes et la technique statistique utilisée (i.e., analyse de variance à mesures répétées) ont permis d'obtenir des résultats statistiquement significatifs et d'en augmenter la fiabilité. Autrement dit, l'influence d'un schéma de classification sur l'efficacité du repérage est bien réelle et n'est pas due au hasard ou à des fluctuations de l'échantillonnage. On obtiendrait des résultats identiques dans le cadre d'une autre simulation réalisée dans les mêmes conditions.

Afin d'augmenter la fiabilité des résultats, l'analyse des données quantitatives et la présentation des résultats statistiques ont été réalisées avec la collaboration d'une équipe d'experts en statistiques du Service de consultation en méthodes quantitatives (SCMEQ) de l'Université de Montréal.

3.7 Conclusion

Le but de notre recherche était d'étudier l'efficacité du repérage d'un ensemble de documents électroniques à l'aide de différents schémas de classification distincts aux plans structurel, logique et sémantique. Dans un premier temps, notre recherche visait à décrire les caractéristiques de différents schémas de classification utilisés par des employés pour organiser et classer leurs documents électroniques. Dans un deuxième temps, notre recherche voulait vérifier l'existence de différences sur le plan de l'efficacité du repérage de documents électroniques qui sont fonction du schéma de classification utilisé.

Pour ce faire, notre recherche a caractérisé un échantillon de 21 schémas de classification produits et utilisés par des employés pour organiser et repérer les documents administratifs électroniques qui sont sous leur contrôle direct. L'objectif de cette étape préliminaire de la recherche était d'approfondir notre connaissance des schémas de classification personnels non seulement sur le plan structurel mais aussi aux plans logique et sémantique. Par la suite nous avons comparé, dans un environnement contrôlé, la capacité de 70 individus à repérer 20 documents électroniques à l'aide de cinq schémas de classification ayant des caractéristiques structurelles, logiques et sémantiques variées. Trois ensembles de données ont été utilisés dans l'analyse en vue d'examiner de quelle façon les différents schémas de classification des documents électroniques influençaient l'efficacité du repérage des documents électroniques : la proportion de documents repérés, le temps moyen requis (mesuré en secondes) pour repérer les documents cibles et la proportion de documents repérés dès le premier essai.

Différentes mesures ont été prises pour assurer la qualité de la recherche en termes de validité et de fidélité. Pour assurer la validité de la recherche, nous avons tenu compte des critères de validité du construit (grille d'analyse multidimensionnelle pré-testée basée sur les résultats des recherches empiriques et sur les écrits théoriques), de validité interne (groupes de répondants équivalents, protocole expérimental qui neutralise les variables étrangères, collecte des données réalisée par une seule personne selon une procédure écrite et pré-testée) et de validité externe (possibilité d'utiliser la méthodologie et le modèle d'analyse dans un autre contexte). En ce qui concerne la fiabilité de la recherche, nous avons pris différents moyens pour assurer la fidélité des instruments de mesure (protocole d'utilisation du logiciel écrit et pré-testé, analyse intra codeur, etc.) et la fiabilité des résultats (protocole expérimental qui neutralise les variables étrangères, analyse statistique effectuée par des experts).

L'analyse de variance à mesures répétées est l'approche d'analyse statistique des données qui a été retenue pour confirmer ou infirmer les hypothèses de notre recherche.

Les résultats de l'analyse des données sont présentés dans le chapitre suivant.

4 Résultats

4.1 Introduction

La recherche dont nous rapportons ici les résultats avait deux objectifs spécifiques. Le premier objectif consistait à décrire les caractéristiques de différents schémas de classification utilisés par des employés pour organiser et classer les documents administratifs électroniques qui sont sous leur contrôle direct. Le deuxième objectif consistait à vérifier, dans un environnement contrôlé, les différences sur le plan de l'efficacité du repérage de documents électroniques ciblés qui sont fonction du schéma de classification utilisé. La réalisation de ces deux objectifs visait à augmenter nos connaissances sur les schémas de classification personnels en contexte de travail et à vérifier s'il était possible de repérer un document avec la même efficacité, quel que soit le schéma de classification utilisé pour ce faire. Ce faisant, nous souhaitons identifier les caractéristiques d'un schéma de classification qui permettraient d'augmenter la probabilité de repérer avec succès un document électronique, tout en minimisant le temps requis pour compléter la tâche ainsi que le risque d'erreur.

Une collecte et une analyse des données en deux étapes furent réalisées pour atteindre ces deux objectifs. La première étape de la collecte de données a été menée auprès de 19 employés. Elle a été suivie d'une analyse des caractéristiques structurelles, logiques et sémantiques de 21 schémas de classification utilisés par ces employés pour organiser les documents administratifs électroniques qui sont sous leur contrôle direct. La section 4.2 présente les résultats descriptifs de la caractérisation de ces schémas de classification personnels.

Dans la deuxième étape de la collecte des données, des séances de simulation du repérage d'un ensemble de documents ciblés ont été réalisées auprès de 70 volontaires. L'objectif de ces séances était de collecter des données en vue d'étudier l'efficacité du repérage de documents électroniques réalisé à l'aide de cinq schémas de classification

présentant des caractéristiques variées. La section 4.3 présente les résultats de l'analyse des données collectées lors des séances de simulation.

Le présent chapitre est consacré à la présentation des résultats obtenus et des observations empiriques réalisées. L'explication et l'interprétation des résultats de cette recherche sont présentées dans le Chapitre 5 Discussion.

4.2 Résultats de la caractérisation des schémas de classification personnels

La première étape de collecte et d'analyse de données visait à répondre à la question spécifique de recherche relative aux schémas de classification personnels utilisés dans un milieu de travail pour organiser et classer les documents administratifs électroniques : « Quelles sont les caractéristiques des schémas de classification utilisés par les employés pour organiser et classer les documents administratifs électroniques qui sont sous leur contrôle direct? ».

Vingt-et-un schémas de classification personnels utilisés par deux catégories d'employés (i.e., des techniciennes à la coordination de travail de bureau et des coordonnateurs de stages) ont été systématiquement examinés à partir d'un modèle d'analyse (Voir section 2.6 Modèle d'analyse des schémas de classification personnels) qui a permis de colliger des mesures continues (i.e., données quantitatives) et des mesures catégorielles (i.e., données qualitatives) regroupées selon trois dimensions :

1. la dimension structurelle désigne les données quantitatives relatives à l'étendue, la profondeur, la complexité et l'équilibre de la structure;
2. la dimension logique désigne les données qualitatives relatives aux critères de division utilisés et à l'ordre de succession des classes;
3. la dimension sémantique désigne à la fois les données quantitatives et qualitatives relatives à la longueur moyenne (en nombre de caractères) des noms de classes, aux stratégies d'abréviation utilisées et au degré de

redondance structurelle et conceptuelle (i.e., proportion de classes redondantes).

La présentation des résultats de l'analyse des caractéristiques des schémas de classification personnels est organisée en trois sections faisant référence aux trois dimensions analysées : les caractéristiques structurelles (section 4.2.1), les caractéristiques logiques (section 4.2.2) et les caractéristiques sémantiques (section 4.2.3) des schémas de classification observés. Puisque la recension des écrits en sciences de l'information et en gestion personnelle de l'information suggère une influence possible du rôle professionnel sur les caractéristiques des schémas de classification conçus et utilisés par les employés, nous présentons systématiquement les résultats de l'analyse de chacune de ces caractéristiques par catégorie d'employé, pour fins de comparaison.

Dans ce chapitre, nous avons attribué un numéro à tous les employés ayant participé à notre recherche afin de pouvoir les référencer individuellement tout en préservant une complète confidentialité. Les numéros précédés des lettres « TCTB » réfèrent à la fois à une technicienne à la coordination de travail de bureau et à son schéma de classification. Les numéros précédés des lettres « CS » font référence à un coordonnateur de stages et à son schéma de classification.

4.2.1 Caractéristiques structurelles des schémas de classification observés

L'analyse de la dimension structurelle des schémas de classification personnels s'intéresse aux indicateurs quantitatifs qui permettent de mesurer l'étendue, la profondeur, la complexité et l'équilibre de la structure de ces schémas. Nous présentons d'abord les caractéristiques structurelles des schémas de classification des techniciennes à la coordination de travail de bureau, suivies des caractéristiques structurelles des schémas de classification des coordonnateurs de stages. En terminant, nous présentons une synthèse des observations relatives aux caractéristiques structurelles des schémas de classification personnels quelle que soit la nature du poste occupé.

4.2.1.1 Caractéristiques structurelles des schémas de classification des techniciennes

Étendue de la structure

Comme l'indique la Figure 4.1 (Voir aussi Tableau 9.1 en Annexe 9 Caractéristiques des schémas de classification personnels observés), on observe de grandes variations dans l'étendue de la macro-structure (i.e., nombre de classes principales) des schémas de classification des techniciennes à la coordination de travail de bureau.

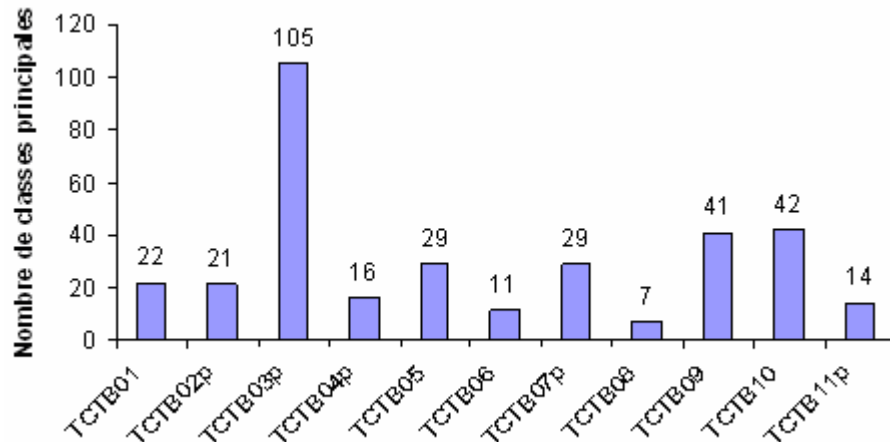


Figure 4.1 Nombre de classes principales observé dans les schémas de classification des techniciennes à la coordination de travail de bureau.

Le nombre de classes principales (i.e., classes de premier niveau) varie d'un minimum de 7 à un maximum de 105. Trois schémas de classification (*TCTB03*, *TCTB09* et *TCTB10*) (27.27 % des schémas de classification des techniciennes) dépassent la moyenne de 30.64 classes (écart-type = 27.16), une moyenne gonflée par le schéma de classification d'une technicienne (*TCTB03*) qui comporte 105 classes principales. Le schéma de classification (*TCTB08*) qui possède seulement 7 classes principales organise principalement les documents par type de contenu (par exemple, « Modèle » ou « Power Point »).

La Figure 4.2 présente le facteur de ramification (« *Branching Factor* ») ou nombre moyen de classes à chaque niveau hiérarchique. Cette variable a été calculée automatiquement par le logiciel *PDS* (Gonçalves & Jorge, 2003) pour chacun des schémas de classification des techniciennes. Elle est considérée comme un indicateur de mesure de l'étendue de la micro-structure (classes subordonnées) des schémas de classification personnels. Le facteur de ramification correspond au nombre b tel que $N = b + b^2 + \dots + b^p$ (N correspond au nombre total de classes et p à la profondeur de la structure) (Gonçalves & Jorge, 2003).

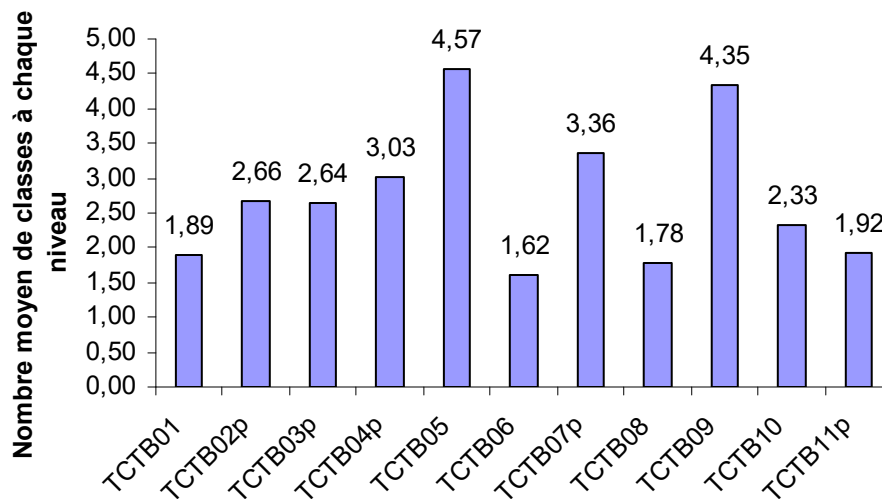


Figure 4.2 Nombre moyen de classes à chaque niveau hiérarchique observé dans les schémas de classification des techniciennes à la coordination de travail de bureau.

Le nombre moyen de classes à chaque niveau varie d'un minimum de 1.62 à un maximum de 4.57 avec un facteur de ramification moyen de 2.74 (écart-type = 1.01) ce qui correspond à des micro-structures peu étendues.

Profondeur de la structure

Six schémas de classification (54.54 % des schémas de classification des techniciennes) ont un nombre maximum de cinq ou six niveaux hiérarchiques (Voir Figure 4.3).

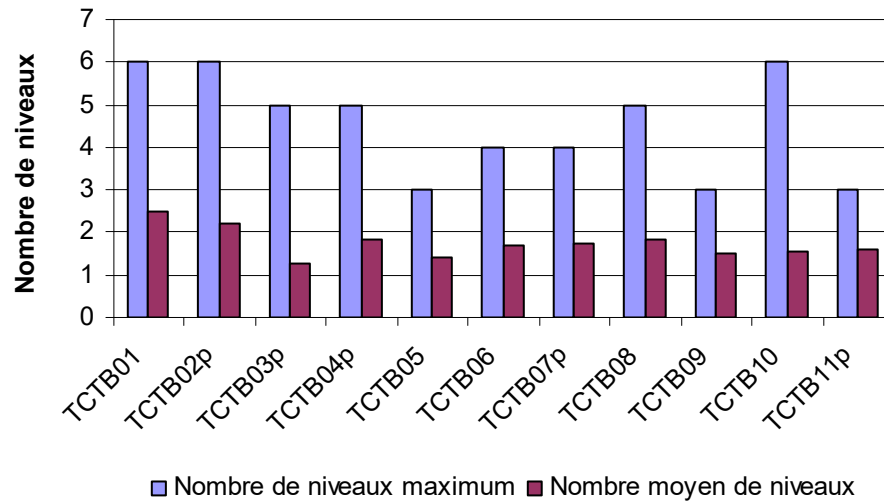


Figure 4.3 Nombre de niveaux maximum et nombre moyen de niveaux observés dans les schémas de classification des techniciennes à la coordination de travail de bureau.

L'observation d'un nombre moyen de niveaux généralement peu élevé (i.e., 1.73) (écart-type = 0.36) nous permet d'affirmer que les schémas de classification étudiés sont généralement peu profonds.

Complexité de la structure

La Figure 4.4 indique de grandes différences dans le nombre total de classes, lequel varie d'un minimum de 106 à un maximum de 3541 classes.

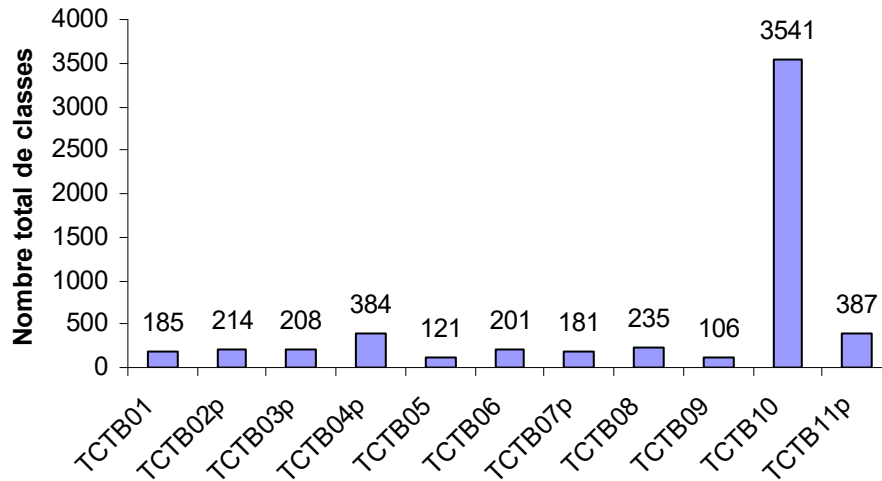


Figure 4.4 Nombre total de classes observé dans les schémas de classification des techniciennes à la coordination de travail de bureau.

Dix schémas de classification (90.9 % des schémas de classification des techniciennes) tombent sous la moyenne de 523.91 classes (écart-type = 1004.70), une moyenne gonflée par le schéma de classification d'une technicienne (*TCTB10*) qui duplique et réutilise le même schéma de classification d'année en année pour faciliter les sauvegardes.

Équilibre de la structure

La Figure 4.5 indique l'écart-type de la moyenne du nombre de classes à chaque niveau. Cette variable est un indicateur du degré d'équilibre des schémas de classification sur le plan de l'étendue de la structure. Autrement dit, cette variable nous permet de déterminer si un schéma de classification a un nombre moyen de sous-classes à chaque niveau qui est similaire d'une classe à l'autre, les valeurs les plus faibles correspondant aux schémas de classification les plus équilibrés (Gonçalves & Jorge, 2003).

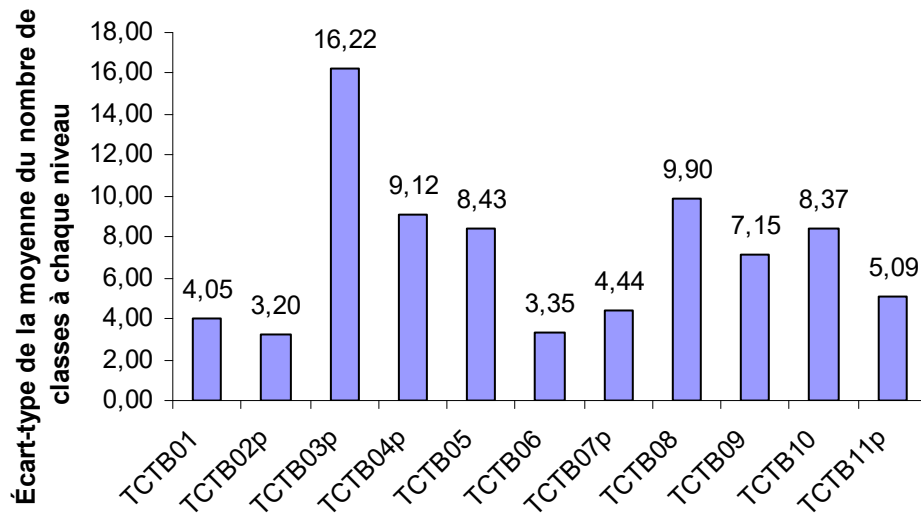


Figure 4.5 Écart-type de la moyenne du nombre de classes à chaque niveau pour chacun des schémas des techniciennes à la coordination de travail de bureau.

Six schémas de classification ont un écart-type de la moyenne du nombre de classes à chaque niveau qui tombe sous la moyenne de 7.21 (écart-type = 3.85), une moyenne gonflée par la valeur de l'écart-type du schéma de classification *TCTB03* (i.e., 16.22). Les résultats indiquent que quatre schémas de classification (36.36 %) présentent une structure relativement plus équilibrée que les sept autres schémas, avec un écart-type inférieur à 5.

4.2.1.2 Caractéristiques structurelles des schémas de classification des coordonnateurs de stages

Étendue de la structure

La Figure 4.6 présente le nombre de classes principales pour chacun des schémas de classification des coordonnateurs de stages.

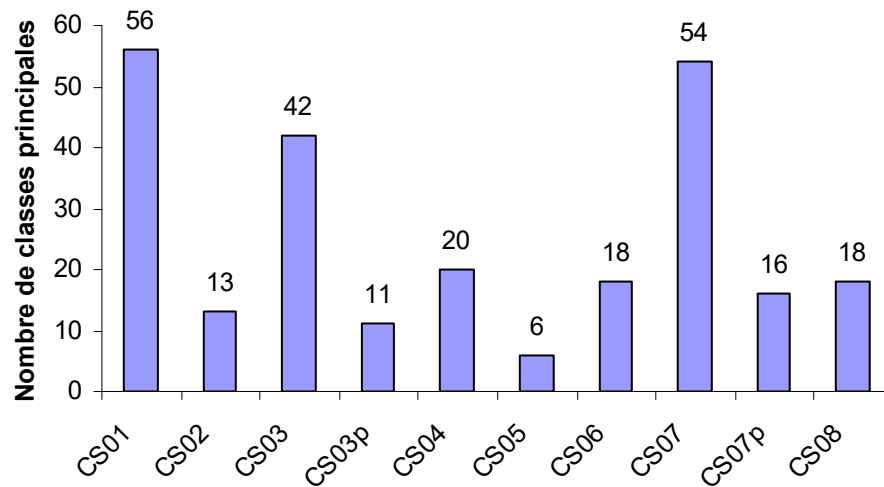


Figure 4.6 Nombre de classes principales observé dans les schémas de classification des coordonnateurs de stages.

Le nombre de classes principales varie d'un minimum de 6 à un maximum de 56. Sept schémas de classification (70 % des schémas de classification des coordonnateurs de stages) tombent sous la moyenne de 25.40 classes principales. L'écart-type de 18.24 confirme cette variété dans l'étendue de la macro-structure (classes principales) des schémas de classification des coordonnateurs de stages. Le schéma de classification (CS05) qui possède le plus petit nombre de classes principales (i.e., six classes principales) est le seul à organiser les documents administratifs électroniques principalement par activité (par exemple, « Administration » ou « Communications »).

La Figure 4.7 présente le nombre moyen de classes à chaque niveau hiérarchique (i.e., facteur de ramification) pour chacun des schémas de classification.

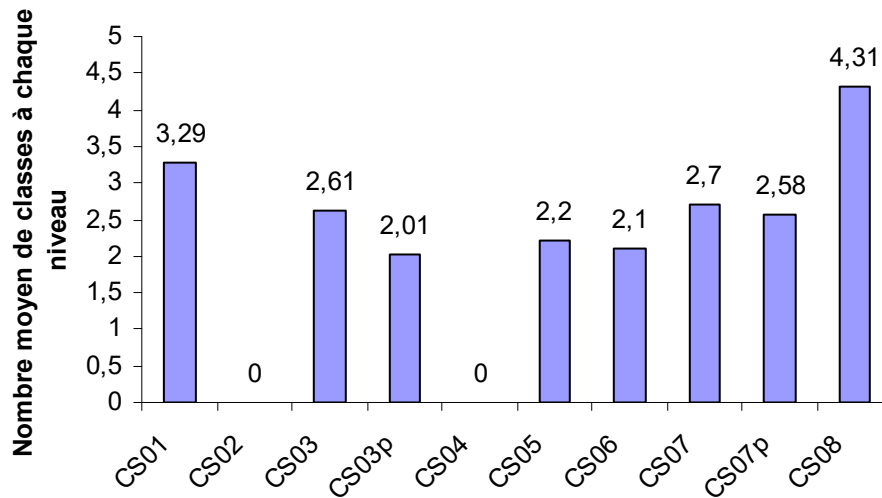


Figure 4.7 Nombre moyen de classes à chaque niveau hiérarchique observé dans les schémas de classification des coordonnateurs de stages.

Le facteur de ramification n'a pas été calculé pour deux schémas de classification (*CS02* et *CS04*) qui ne possèdent qu'un seul niveau hiérarchique. La micro-structure des huit autres schémas de classification peut être qualifiée de faiblement étendue en raison d'un nombre moyen de classes à chaque niveau peu élevé (i.e., 2.73) (écart-type = 0.76) (Gonçalves & Jorge, 2003).

Profondeur de la structure

Comme l'indique la Figure 4.8, cinq schémas de classification (50 % des schémas de classification des coordonnateurs de stages) ont un nombre maximum de cinq ou six niveaux hiérarchiques.

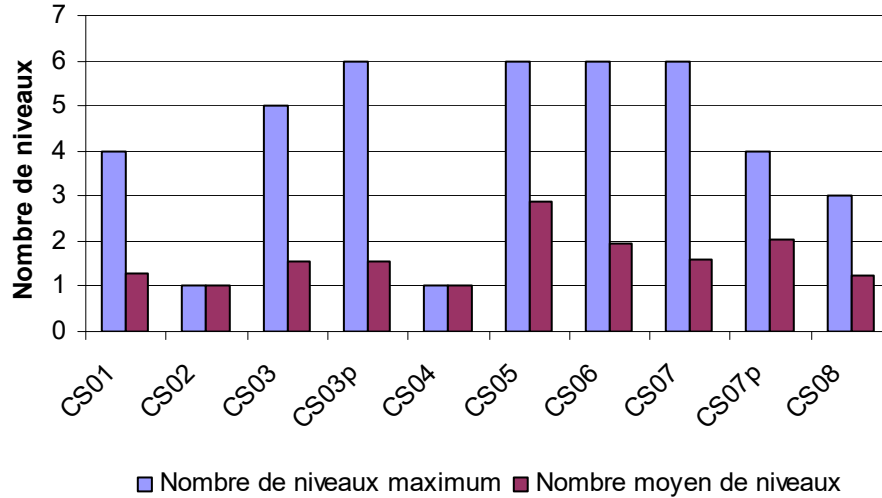


Figure 4.8 Nombre de niveaux maximum et nombre moyen de niveaux observés dans les schémas de classification des techniciennes à la coordination de travail de bureau.

Le nombre moyen de niveaux sous chaque classe principale est très faible (i.e., 1.61) (écart-type = 0.56), ce qui correspond à des structures peu profondes.

Complexité de la structure

On peut observer aussi dans les schémas de classification des coordonnateurs de stages de grandes différences dans le nombre total de classes, lequel varie d'un minimum de 13 à un maximum de 618 classes (Voir Figure 4.9).

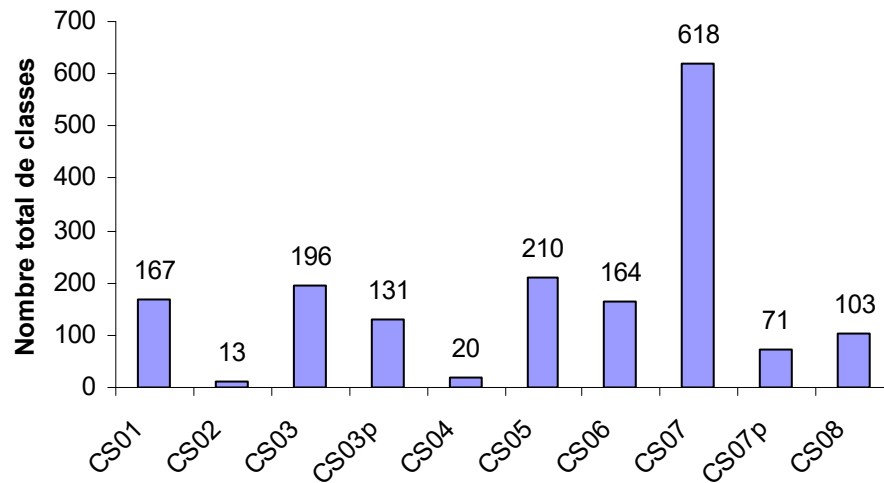


Figure 4.9 Nombre total de classes observé dans les schémas de classification des coordonnateurs de stages.

Sept schémas de classification tombent sous un total moyen de 169.30 classes (écart-type = 171.91), une moyenne gonflée par un schéma de classification comportant 618 classes (CS07) et qui intègre à la fois les classes créées par le coordonnateur de stages CS07 et celles créées par un autre employé ayant déjà occupé le même poste.

Équilibre de la structure

La Figure 4.10 indique l'écart-type de la moyenne du nombre de classes à chaque niveau pour chacun des schémas de classification des coordonnateurs de stages.

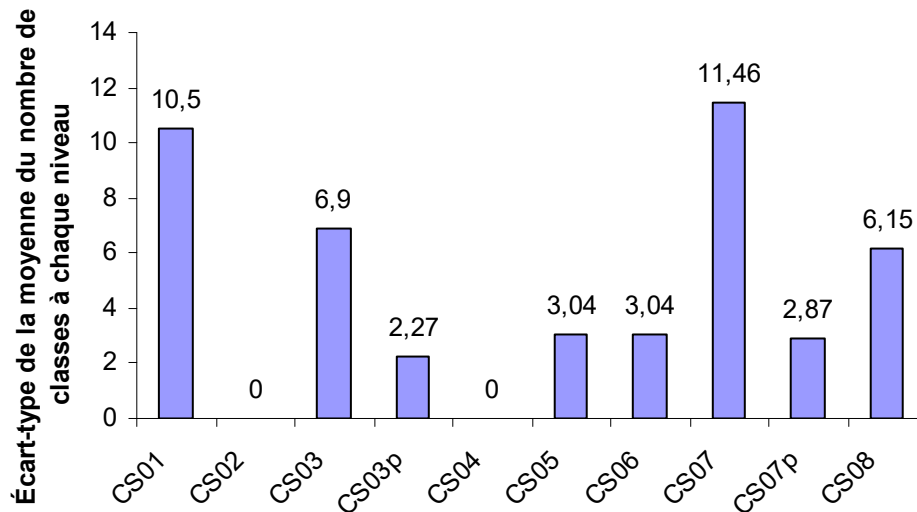


Figure 4.10 Écart-type de la moyenne du nombre de classes à chaque niveau pour chacun des schémas de classification des coordonnateurs de stages.

L'écart-type de la moyenne du nombre de classes à chaque niveau n'a pas été calculé pour les deux schémas de classification (CS02 et CS04) ne possédant qu'un seul niveau hiérarchique. Les résultats du calcul de cette variable montrent que seulement quatre schémas de classification de coordonnateurs de stages (40 %) présentent une structure relativement plus équilibrée que les autres, avec un écart-type inférieur à cinq. CS01 et CS07 sont les schémas de classification les moins équilibrés de tous avec un écart-type supérieur à 10.

4.2.1.3 Synthèse

Nous présentons dans cette section une synthèse des résultats obtenus pour chaque indicateur de mesure de la dimension structurelle des schémas de classification observés, soit l'étendue, la profondeur, la complexité et l'équilibre de la structure, quel que soit le rôle professionnel de l'employé.

Macro-structures étendues, micro-structures peu étendues : Si l'on considère l'ensemble des schémas de classification observés, le nombre de classes principales varie d'un minimum de 6 à 105 avec une moyenne de 28.14 classes principales (écart-type = 22.93). Le nombre moyen de classes à chaque niveau varie entre 1.61 et 4.57 avec une moyenne de 2.73 classes à chaque niveau (écart-type = 0.89). Ces résultats correspondent à des schémas de classification dont la macro-structure est généralement étendue (nombre élevé de classes principales) et la micro-structure est peu étendue (nombre peu élevé de classes à chaque niveau).

Structures peu profondes : Malgré un maximum de cinq à six niveaux hiérarchiques observé dans l'absolu, les schémas de classification présentent des structures peu profondes avec un nombre moyen de 1.67 niveaux (écart-type = 0.46). On note que deux schémas de classification utilisés par des coordonnateurs de stages ne comportent qu'un seul niveau hiérarchique.

Structures complexes : Le nombre total de classes varie considérablement d'un schéma de classification à l'autre allant d'un minimum de 13 à un maximum de 3541 classes avec une moyenne de 355.05 classes (i.e., écart-type = 742.25). Ces résultats correspondent à des structures généralement complexes (Boardman & Sasse, 2004; Henderson, 2004).

Structures peu équilibrées : L'écart-type de la moyenne du nombre de classes à chaque niveau (i.e., 6.61) montre que les schémas de classification présentent une structure généralement peu équilibrée. Autrement dit, l'étendue de la structure n'est pas uniforme, certaines classes se subdivisent en une seule sous-classe (fréquent) alors que d'autres classes peuvent contenir plus d'une vingtaine de sous-classes (plus rare). Seuls les deux schémas de classification partagés des coordonnateurs de stages présentent un très faible écart-type moyen de 2.57 correspondant à une structure plus équilibrée (Gonçalves & Jorge, 2003).

En résumé, les schémas de classification personnels observés se caractérisent au plan structurel par une macro-structure généralement étendue et une micro-structure peu étendue, une structure peu profonde, complexe et généralement déséquilibrée. Les résultats obtenus confirment les travaux de Khoo *et al.* (2007) qui ont aussi observé des schémas de classification personnels dont la macro-structure est généralement étendue (« *broad* ») (i.e., minimum de 19 classes principales) et peu profonde (« *shallow* ») (i.e., maximum de trois niveaux hiérarchiques). Les résultats de l'analyse des données relatives à l'étendue de la micro-structure (i.e., nombre moyen de classes à chaque niveau) confirment les résultats de Gonçalves & Jorge (2003) qui avaient observé un facteur de ramification peu élevé (i.e., 1.84; écart-type = 0.187), soit une micro-structure peu étendue ou étroite (« *narrow* »). De plus, les variations observées dans le nombre total de classes ont été également relevées par Boardman & Sasse (2004) (i.e., entre 5 et 218 classes) et Henderson (2004) (i.e., entre 211 et 7200 classes).

Les résultats infirment les observations de Gonçalves & Jorge (2003) relatives à l'équilibre et à la profondeur des structures. L'écart-type de la moyenne du nombre de classes à chaque niveau observé au sein de notre échantillon est relativement plus élevé (i.e., 6.61; écart-type = 3.72) que celui obtenu par Gonçalves & Jorge (2003) (i.e., 3.61; écart-type = 0.67). Ces résultats indiquent que les schémas de classification des techniciennes et des coordonnateurs de stages sont généralement peu équilibrés. De plus, les schémas de classification observés sont peu profonds avec un nombre moyen de 1.67 niveaux (écart-type = 0.46). Ces résultats sont loin de la moyenne de 8.45 niveaux hiérarchiques (écart-type = 2.9) observée par Gonçalves & Jorge (2003).

Les résultats révèlent que les différences les plus marquées au plan structurel entre les schémas de classification des techniciennes et des coordonnateurs de stages sont observées aux plans de la complexité (i.e., nombre total de classes). Dans notre échantillon, les employés qui conservent le plus de documents électroniques et qui utilisent le plus grand nombre de classes sont des techniciennes. Trois coordonnateurs de stages ont déclaré en entrevue ne pas conserver les « anciens » documents sauf comme « preuve de communication » ou « preuve d'envoi ».

La discussion des résultats relatifs aux caractéristiques structurelles des schémas de classification personnels est présentée dans le Chapitre 5, section 5.2.1 Caractéristiques structurelles des schémas de classification observés.

4.2.2 Caractéristiques logiques des schémas de classification observés

L'analyse de la logique des schémas de classification s'intéresse aux critères qualitatifs de division et de regroupement ainsi qu'à l'ordre de succession des classes aux premier et deuxième niveaux de la structure classificatoire.

Comme l'indique le Tableau 4.1 (Voir aussi Tableaux 9.4 et 9.5 en Annexe 9 Caractéristiques des schémas de classification personnels observés), l'analyse des 21 schémas de classification a permis de révéler huit principaux critères de division logique utilisés par les employés pour organiser les documents administratifs électroniques qui sont sous leur contrôle direct. Il s'agit de : (1) Activité, (2) Thème, (3) Type de contenu, (4) Projet, (5) Temps, (6) Organisme, (7) Contact, et (8) Mélange (i.e., combinaison de plusieurs critères de division).

Tableau 4.1 Critères de division logique observés aux 1er et 2ème niveaux hiérarchiques

Critères de division	Définition
1. Activité	Division basée sur une activité générique importante ou régulière du participant qui est réalisée sur une base annuelle (par exemple, la gestion des opérations comptables, la gestion du courrier et des télécommunications, la gestion des documents d'archives, l'organisation de réunions, etc.).
2. Thème	Division basée sur le contenu thématique des documents (par exemple « plagiat », « bourses ») ou sur l'absence d'un contenu thématique uniforme (par exemple « divers » ou « généralité »).
3. Type de contenu	Division basée sur les propriétés communes sur le plan de la forme et de la fonction du contenu d'un ensemble de documents (par exemple « formulaires », « rapport » ou « statistiques »).
4. Projet	Division basée sur une activité spécifique à court terme telle que l'organisation ou la gestion d'un événement (par exemple « Journée d'accueil »), les travaux réalisés dans le cadre d'un cours suivi (par exemple « Cours de français ») ou un projet en cours de développement (par exemple « Révision programme »).
5. Temps	Division basée sur une année ou une session universitaire (par exemple « 2003-2004 » ou « Aut2003 »).
6. Organisme	Division basée sur un organisme décisionnel ou consultatif (conseils, assemblées, comités) de l'Université de Montréal, d'une de ses facultés ou d'un de ses départements (par exemple « Comité de gestion »).
7. Contact	Division basée sur des noms de personnes ou groupes de personnes physiques ou morales (organismes publics, parapublics ou privés) avec lesquelles le répondant échange des documents sur une base régulière ou occasionnelle (par exemple « Julie »).
8. Mélange	Division basée sur un critère de division logique complexe ne permettent pas d'identifier clairement le critère de division logique qui est prépondérant. Par exemple, peuvent être classés ici les types de document ayant une spécification pour la période couverte (par exemple « Correspondance 2006 ») ou de sa provenance (par exemple « PV comite MScPhD »).

Les autres critères de division identifiés dans la littérature (Voir Tableau 2.1 Principaux critères de division logique recensés dans les écrits) sont utilisés de manière plus sporadique par les employés interrogés.

L'ordre de succession des classes proposé par défaut par le système informatique est l'ordre alphabétique. Toutefois, il peut arriver que les employés utilisent des chiffres, des lettres ou des caractères spéciaux pour ordonner différemment leurs classes selon une logique personnelle ou institutionnelle. L'analyse de l'échantillon des schémas de classification a permis de révéler quatre ordres de succession des classes différents (i.e.,

ordre alphabétique, ordre chronologique, ordre numérique et ordre logique) utilisés par les employés pour organiser les classes aux 1^{er} et 2^{ème} niveaux de la hiérarchie (Voir Tableau 4.2).

Tableau 4.2 Typologie des différents ordres de succession des classes observés

Ordres de succession des classes	Définition
1. Ordre alphabétique	Les classes sont ordonnées selon un ordre alphabétique. Il s'agit de l'ordre de succession des classes proposé par défaut par le système (par exemple, « Comités et organismes », « Correspondance », « États généraux », etc.).
2. Ordre chronologique	Les classes sont ordonnées selon un ordre chronologique, le plus souvent par année (par exemple, « 2003-2004 », « 2004-2005 », « 2005-2006 », etc.).
3. Ordre numérique	Les classes sont ordonnées selon une numérotation institutionnelle (i.e., numérotation basée sur le système de cotation du Système officiel de classification) (par exemple, « 16.1 Correspondance ») ou personnelle (i.e., numérotation séquentielle) (par exemple, « 01 Comptabilité », « 02 Gestion et conservation des documents », etc.).
4. Ordre logique	Les classes sont ordonnées selon une logique personnelle basée le plus souvent sur leur fréquence d'utilisation par l'employé (par exemple, « Études supérieures », « Accueil », etc.).

Dans les sections suivantes, nous présentons les caractéristiques logiques qui ont été observées dans les schémas de classification personnels des deux catégories d'employés sélectionnées, soit les caractéristiques logiques des schémas de classification des techniciennes à la coordination de travail de bureau et les caractéristiques logiques des schémas de classification des coordonnateurs de stages. En terminant, nous présentons une synthèse des résultats des caractéristiques logiques quelle que soit la nature du poste occupé.

4.2.2.1 Caractéristiques logiques des schémas de classification des techniciennes

Critères de division logique

La thématique est le plus souvent utilisée comme premier critère de division dans cinq schémas de classification des techniciennes (45.46 % des schémas de classification des techniciennes). Le *type de contenu* est utilisé comme premier critère de division dans quatre autres schémas (36.36 %) et le *mélange* (18.19 %) est le premier critère de

division utilisé dans deux autres schémas. Le *temps* est quant à lui un critère de division qui est privilégié au second niveau dans quatre schémas de classification (36.36 %).

Ordre de succession des classes

L'ordre de succession des classes principales (1^{er} niveau hiérarchique) est exclusivement *alphabétique* dans six schémas de classification (54.54 %). Une technicienne utilise à la fois un ordre *numérique* (95 % de ses classes principales) et un ordre *alphabétique* (5 % de ses classes principales). Deux autres techniciennes privilégient un ordre *alphabétique* (93 % à 95 % des classes) combiné à un ordre *logique* (5 % à 7 % des classes). L'ordre de succession des classes secondaires (2^{ème} niveau hiérarchique) pour l'ensemble des schémas de classification des techniciennes est principalement *alphabétique*.

4.2.2.2 Caractéristiques logiques des schémas de classification des coordonnateurs de stages

Critères de division logique

Le *thème* est utilisé comme premier critère de division dans cinq schémas de classification des coordonnateurs de stages (50 % des schémas de classification des coordonnateurs). Le *mélange* est utilisé comme premier critère de division dans quatre autres schémas de classification (40 %). Il n'y a pas de second critère de division qui prédomine dans ce groupe d'employés bien que le *thème* soit privilégié au second niveau hiérarchique dans trois schémas de classification (30 %).

Ordre de succession des classes

Les classes principales sont ordonnées alphabétiquement dans huit schémas de classification (80 % des schémas de classification des coordonnateurs de stages). Un coordonnateur de stages utilise un ordre *numérique* personnel lui permettant, selon ses propos, d'ordonner ses classes en fonction de leur fréquence d'utilisation (par exemple, la classe « 01 Stages » est utilisée quotidiennement et apparaît en « haut de la liste »). L'ordre de succession des classes secondaires (2^{ème} niveau) est strictement *alphabétique*

dans quatre schémas de classification (40 %). On observe que l'ordre *alphabétique* est combiné à un ordre *chronologique* pour les quatre autres schémas observés (40%).

4.2.2.3 Synthèse

Nous présentons dans cette section une synthèse des résultats obtenus pour les deux indicateurs de mesure de la dimension logique des schémas de classification observés, soit les critères de division logique et l'ordre de succession des classes utilisés dans les schémas de classification observés, quel que soit le rôle professionnel de l'employé.

Regroupement par « thème » et « mélange » : Les résultats de l'analyse des caractéristiques logiques des schémas de classification révèlent l'absence d'un critère de division prédominant. Cela dit, l'analyse a mis en évidence une classification qui privilégie soit un regroupement thématique des classes principales (47.61 % des schémas de classification personnels observés) (n = 10) soit un regroupement qui combine plusieurs concepts (i.e., *mélange*) (28.57 %) (n = 6).

Ordre alphabétique de succession des classes : L'ordre de succession des classes principales (classes de 1^{er} niveau) qui est privilégié par les employés interrogés est l'ordre *alphabétique* (80.95 % des schémas de classification) (n = 17) proposé par défaut par le système d'exploitation Windows. L'ordre *numérique* a été observé dans seulement quatre schémas de classification des documents électroniques, tous basés sur le schéma de classification utilisé pour organiser les documents sur support papier. L'ordre de succession des classes secondaires (2^{ème} niveau) est principalement *alphabétique* pour 85.71 % des schémas de classification (n = 18).

Alors que les employés sont généralement satisfaits de l'ordonnement alphabétique de leurs classes, certains d'entre eux mettent en place des stratégies pour imposer un ordre personnel. Par exemple, trois employés intègrent le caractère « _ » devant les noms de certaines classes pour que le système d'exploitation présente automatiquement ces dernières « en haut » de la liste. Une technicienne en coordination de travail de bureau utilise l'ordre alphabétique par défaut sauf pour « ses archives », qui

correspondent à tous les documents électroniques déjà présents sur le poste informatique à son entrée en fonction. Elle fait précéder cette classe « Archives » par la lettre « z » (i.e., « zArchives ») afin que la classe apparaisse à la toute fin de la liste.

Les résultats confirment l'ordonnement alphabétique des classes et la variété des critères de division observés dans les recherches menées par Boardman & Sasse (2004), Henderson (2005) et Khoo *et al.* (2007). Par ailleurs, nos résultats infirment l'importance d'un regroupement par type de document observé par Henderson (2005) et Khoo *et al.* (2007) ou d'un regroupement par projet (Boardman & Sasse, 2004). En effet, le regroupement par type de document a été observé seulement comme premier critère de division dans quatre schémas de classification; les employés interrogés privilégient plutôt un regroupement thématique et utilisent fréquemment des classes qui représentent plusieurs concepts.

La discussion des résultats relatifs aux caractéristiques logiques des schémas de classification personnels est présentée dans le Chapitre 5, section 5.2.2 Caractéristiques logiques des schémas de classification observés.

4.2.3 Caractéristiques sémantiques des schémas de classification observés

L'analyse de la dimension sémantique des schémas de classification s'intéresse à la terminologie utilisée par les employés pour nommer leurs classes. Nous avons réalisé une analyse structurelle et conceptuelle des noms des classes pris hors contexte (hors hiérarchie) (i.e., analyse de la longueur moyenne (en nombre de caractères) des noms de classes et des principales stratégies d'abréviation utilisées) et en contexte (à l'intérieur de la hiérarchie) (i.e., analyse du degré de redondance structurelle et conceptuelle). Nous présentons dans cette section les caractéristiques sémantiques observées dans les schémas de classification personnels pour chaque catégorie d'employés, soit les caractéristiques sémantiques des schémas de classification des techniciennes à la coordination de travail de bureau et les caractéristiques sémantiques des schémas de classification des coordonnateurs de stages. En terminant, nous présentons une synthèse

des observations relatives aux caractéristiques sémantiques quelle que soit la nature du poste occupé.

4.2.3.1 Caractéristiques sémantiques des schémas de classification des techniciennes

Longueur moyenne des noms de classes

La longueur moyenne (en nombre de caractères) des noms de classes varie d'un schéma de classification à l'autre de 10.50 à 22.41 caractères. Bien que la moyenne du nombre de caractères soit de 15.99, l'écart-type de 4.06 confirme une légère dispersion. Dans notre échantillon, nous observons que les schémas de classification qui ont les noms de classes les plus longs sont basés sur le *Système officiel de classification* de l'Université de Montréal.

Les résultats montrent que la longueur moyenne (en nombre de caractères) des noms de classes pourrait varier selon le caractère partagé ou non d'un schéma de classification. Les schémas de classification des techniciennes qui ne sont pas partagés avec des collègues de travail ont généralement des noms plus courts (i.e., moyenne de 14.61 caractères; écart-type = 3.07) que les schémas de classification partagés (i.e., moyenne de 17.65 caractères; écart-type = 4.81).

Principales stratégies d'abréviation utilisées

En moyenne, 40.55 % (écart-type = 11.71) des noms de classes des schémas de classification des techniciennes est abrégée (Voir Tableau 9.9 en Annexe 9 Caractéristiques des schémas de classification personnels observés). L'utilisation des *acronymes* (par exemple, « CV » pour « Curriculum Vitae ») est le mode d'abréviation des classes principales qui est privilégié dans cinq schémas de classification (45.46 % schémas de classification des techniciennes). Les acronymes sont aussi très présents dans la dénomination des classes subordonnées de six schémas de classification (54.54 %). Ces acronymes désignent le plus souvent des organismes institutionnels (par exemple, le « BRI » pour « Bureau de la recherche institutionnelle »), des organismes subventionnaires (par exemple, le « FQRSC » pour « Fonds québécois de la recherche

sur la société et la culture »), des organismes externes (par exemple, la « CREPUQ » pour « Conférence des recteurs et principaux des universités du Québec »), des associations (par exemple, la « FAÉCUM » pour « Fédération des étudiants de l'Université de Montréal ») et des types de contenu (par exemple, « PV » pour « Procès-verbal »).

Redondance structurelle et conceptuelle

On observe une redondance structurelle (répétition d'une sous chaîne de caractères) et conceptuelle (répétition d'un concept) dans la dénomination des classes principales de cinq schémas de classification (45.46 %). Voici deux exemples de redondance : « Courrier direction », « Courrier étudiants » et « Charges profs 2001-2002 », « Charges profs 2002-2003 », « Charges profs 2003-2004 », etc. Ces redondances pourraient être évitées par des regroupements de classes sous des termes plus génériques tels que « Courrier » et « Charges profs ». De tels regroupements ont été jugés souhaitables par les employés eux-mêmes durant les rencontres individuelles mais ne sont pas réalisés par manque de temps.

4.2.3.2 Caractéristiques sémantiques des schémas de classification des coordonnateurs de stages.

Longueur moyenne des noms de classes

La longueur moyenne (en nombre de caractères) des noms de classes varie de 13.10 à 23.55 caractères d'un schéma de classification à l'autre avec une moyenne de 16.26 caractères (écart-type = 3.27). Ces résultats sont similaires à la longueur moyenne des noms de classes des schémas de classification des techniciennes (15.99 caractères en moyenne, écart-type = 4.06).

Dans notre échantillon, nous observons que le schéma de classification doté de noms de classes les plus longs (i.e., moyenne de 23.55 caractères) est un schéma de classification partagé avec un collègue de travail.

Principales stratégies d'abréviation utilisées

En moyenne, un peu plus du tiers des noms de classes des schémas de classification des coordonnateurs est abrégé (33.20 %) (écart-type = 21.77). On observe un usage privilégié des *acronymes* à tous les niveaux hiérarchiques de trois schémas de classification des coordonnateurs de stages (30 %) et un usage fréquent de *codes* (par exemple, le sigle d'un cours) dans trois autres schémas (30 %). On remarque aussi une forte proportion d'acronymes au premier niveau pour quatre schémas de classification (40 %). Ces acronymes représentent le plus souvent des organismes institutionnels (par exemple, « FES » pour « Faculté des études supérieures »), des organismes subventionnaires (par exemple, « FQRSC » pour « Fonds québécois de la recherche sur la société et la culture »), des organismes externes (par exemple, « OFQJ » pour « Office franco-québécois pour la jeunesse »), des associations (par exemple, « ACPUM » pour « Association des cadres et professionnels de l'Université de Montréal »), des types d'établissements (par exemple, « ES » pour « Établissements de santé »), des lieux (par exemple, « NB » pour « Nouveau Brunswick ») et des types de documents (par exemple, « CV » pour « Curriculum vitae »). La forte proportion de *codes* s'explique par la forte présence des sigles de cours et des groupes d'étudiants dans les noms des classes.

Redondance structurelle et conceptuelle

On observe une redondance structurelle et conceptuelle au plan des classes principales (1^{er} niveau) de la hiérarchie de six schémas de classification (60 %). Voici quelques exemples de redondance structurelle et conceptuelle au premier niveau : « Formation à la supervision », « Formation continue », « Formation clinique », « Rencontres annuelles avec les milieux », « Rencontres d'équipe », « Évaluation des stages », « Évaluation étud. des stages ». Ces redondances indiquent une spécificité des classes dès le premier niveau de la hiérarchie et pourraient être évitées par des regroupements de ces classes sous des termes plus génériques tels que « Formations », « Rencontres » et « Évaluations ».

4.2.3.3 Synthèse

Cette section présente une synthèse des résultats obtenus pour les trois indicateurs de mesure de la dimension sémantique des schémas de classification observés, soit la longueur moyenne (en nombre de caractères) des noms des classes, les principales stratégies d'abréviation utilisées et le degré de redondance structurelle et conceptuelle.

Longueur moyenne des noms de classes : La longueur moyenne des noms de classes créées par l'ensemble des employés interrogés est de 16.12 caractères (écart-type = 0.19). Ces résultats sont légèrement supérieurs aux données recueillies par Gonçalves & Jorge (2003) qui avaient observé un nombre moyen de 12.56 caractères. Les résultats suggèrent que cette longueur varie selon le caractère partagé ou non d'un schéma de classification. En effet, les schémas de classification non partagés observés ont généralement des noms plus courts (i.e., moyenne de 15.13 caractères, écart-type = 2.56) que les schémas de classification partagés (i.e., moyenne de 18.16 caractères, écart-type = 4.88). Le souci d'attribuer des noms de classes à la fois plus longs et plus précis a été exprimé en entrevue par les employés utilisant des schémas de classification partagés mais aussi par des employés soucieux de la dénomination de leurs propres classes (i.e., généralement non partagées) pour « qu'elles soient repérables par n'importe qui ». De plus, les résultats indiquent que les schémas de classification personnels qui utilisent certaines classes du schéma de classification institutionnel ont des noms de classes généralement plus longs (i.e., une moyenne de 20.27 caractères).

Utilisation répandue des acronymes : On note l'usage privilégié des *acronymes* dans les stratégies d'abréviation des techniciennes. En ce qui concerne les coordonnateurs de stages, les stratégies d'abréviation privilégiées sont différentes selon le caractère partagé ou non d'un schéma de classification. Ainsi l'usage des *acronymes* est également répandu pour les schémas de classification non partagés des coordonnateurs de stages alors que l'usage de *codes* est plus fréquent pour leurs schémas de classification partagés. Par ailleurs, on observe une proportion moyenne de classes non abrégées plus faible pour les schémas de classification partagés des techniciennes

(48 %) (n = 5) alors qu'inversement, le partage d'un schéma de classification d'un coordonnateur de stages ne semble pas diminuer cette proportion (68.5 %) (n = 2).

Forte redondance structurelle et conceptuelle : Comme l'indique le Tableau 4.3, la proportion moyenne des classes redondantes au niveau structurel et conceptuel est différente selon le caractère partagé ou non d'un schéma de classification.

Tableau 4.3 Proportion moyenne de classes redondantes dans les schémas de classification partagés et non partagés des techniciennes et des coordonnateurs de stages

	<i>Schémas non partagés</i>		<i>Schémas partagés</i>	
	Techniciennes (N = 6)	Coordonnateurs (N = 8)	Techniciennes (N = 5)	Coordonnateurs (N = 2)
Proportion moyenne des classes redondantes (tous niveaux confondus)	15.95 %	12.47 %	32.65 %	31.04 %

On constate une redondance structurelle et conceptuelle généralement plus élevée dans les schémas de classification partagés des techniciennes et des coordonnateurs de stages. Concomitante à cette redondance, on observe l'utilisation d'une dénomination très spécifique des classes dès le premier niveau de la hiérarchie.

En résumé, les schémas de classification personnels observés se caractérisent au plan sémantique par des classes ayant des noms composés d'un nombre moyen de caractères relativement homogène d'un schéma à un autre. Les résultats révèlent des stratégies d'abréviation qui privilégient l'usage d'acronymes et une redondance structurelle et conceptuelle à l'intérieur des noms de classes généralement élevée dès le premier niveau de la hiérarchie.

La discussion des résultats relatifs aux caractéristiques sémantiques des schémas de classification personnels est présentée dans le Chapitre 5, section 5.2.3 Caractéristiques sémantiques des schémas de classification observés.

4.3 Résultats de la simulation du repérage

Cette section présente les résultats de l'analyse statistique des données collectées lors de la simulation du repérage d'un ensemble de documents ciblés. La collecte et l'analyse de ces données visaient à répondre aux questions spécifiques de recherche suivantes :

Existe-t-il des différences sur le plan de la capacité à repérer tous les documents ciblés à l'aide de schémas de classification ayant des caractéristiques variées?

Existe-t-il des différences sur le plan de la capacité à repérer rapidement les documents ciblés à l'aide de schémas de classification ayant des caractéristiques variées?

Existe-t-il des différences sur le plan de la capacité à repérer les documents ciblés dès le premier essai à l'aide de schémas de classification ayant des caractéristiques variées?

L'hypothèse générale que nous souhaitons vérifier est formulée ainsi : « Il existe des différences sur le plan de l'efficacité du repérage des documents électroniques réalisé à l'aide de schémas de classification distincts aux plans structurel, logique et sémantique ».

Les hypothèses spécifiques de recherche sont les suivantes :

HS1. Un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus de documents.

HS2. Un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus rapidement les documents.

HS3. Un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus de documents dès le premier essai.

Pour répondre aux questions et tester les hypothèses, nous avons demandé à 70 répondants de repérer un ensemble de documents ciblés en utilisant cinq schémas de classification présentant des caractéristiques structurelles, logiques et sémantiques variées. Pour chaque document à repérer, la chercheuse a pris la mesure du nombre d'essais (pour un maximum de trois essais), du temps requis à chaque essai (temps mesuré en secondes) et de la capacité à repérer le document (i.e., succès, échec après trois essais ou abandon).

L'analyse de variance à un facteur répété à cinq niveaux (schémas) a été retenue comme approche d'analyse statistique des données collectées lors de la simulation. L'analyse de variance a permis de comparer, pour chaque mesure, les moyennes obtenues entre les schémas et de déterminer si les différences observées sont significatives (Fortin, Côté & Filion, 2006). Dans notre recherche, toutes les hypothèses spécifiques ont été testées à un seuil de signification établi à 5 % ($p \leq 0.05$) (Ancelle, 2002).

Dans un premier temps, ce chapitre présente des observations générales sur le protocole expérimental qui ressortent de l'analyse statistique des données collectées durant la simulation. Par la suite, la présentation des résultats est divisée selon la vérification des trois hypothèses spécifiques de notre recherche, soit l'évaluation de l'efficacité du repérage, en fonction des schémas de classification utilisés, sur le plan (1) de la proportion de documents repérés (quel que soit le nombre d'essais nécessaires), (2) du temps moyen (mesuré en secondes) requis pour repérer un document et (3) de la proportion de documents repérés dès le premier essai.

4.3.1 Observations générales

L'analyse statistique des résultats de la simulation permet de formuler des observations générales sur le protocole expérimental et sur les hypothèses spécifiques de recherche. Ces observations sont relatives aux thématiques suivantes : (1) le niveau de difficulté de l'expérimentation, (2) la validation du choix du nombre d'essais maximum et (3) les différences observées entre les cinq schémas de classification.

4.3.1.1 Niveau de difficulté de l'expérimentation

Le Tableau 4.4 permet de déterminer *a posteriori* le niveau de difficulté de notre expérimentation en indiquant le nombre total de documents repérés (sur un maximum de 20 documents) et le nombre de répondants ayant été capables de repérer ce même nombre total de documents (i.e., fréquences pour le nombre de documents repérés).

Tableau 4.4 Fréquences pour le nombre de documents repérés

<i>Nombre de documents repérés</i>	<i>Nombre de répondants (N)</i>	<i>Proportion de répondants (%)</i>	<i>Pourcentage valide (%)</i>	<i>Pourcentage cumulé (%)</i>
11.00	2	2.9	2.9	2.9
13.00	3	4.3	4.3	7.1
14.00	4	5.7	5.7	12.9
15.00	8	11.4	11.4	24.3
16.00	10	14.3	14.3	38.6
17.00	15	21.4	21.4	60.0
18.00	14	20.0	20.0	80.0
19.00	11	15.7	15.7	95.7
20.00	3	4.3	4.3	100.0
Total	70	100.0	100.0	

Seulement trois répondants (4.3 % des répondants) ont été capables de repérer l'ensemble des 20 documents. Tous les répondants ont été capables de repérer au moins 11 documents sur les 20 documents ciblés (55 % des documents ciblés).

Ces données empiriques nous permettent d'affirmer que l'expérimentation possédait un niveau de difficulté « moyen » dans l'accomplissement des tâches demandées aux répondants. Autrement dit, le repérage des documents ciblés n'a été ni trop facile ni trop difficile à réaliser afin d'observer suffisamment de succès, d'échecs ou d'abandons nécessaires pour réaliser les inférences statistiques. Un niveau de difficulté plus faible aurait probablement permis à l'ensemble des répondants de repérer les 20 documents. Un niveau de difficulté plus élevé aurait augmenté le risque de découragement ou de désintérêt de la part des répondants. Ce risque aurait fort

probablement pu se manifester par un taux d'échec et d'abandon plus élevé que ce qui a été observé dans la présente recherche (Voir plus loin le Tableau 4.5 Proportion de documents repérés et proportion de documents non repérés en rapport avec le nombre d'essais pour l'ensemble des schémas de classification).

4.3.1.2 Validation du choix du nombre d'essais

Dans le cadre de notre expérimentation, nous avons retenu de travailler avec un nombre de trois essais maximum pour repérer un document. Ce nombre avait été jugé satisfaisant dans la recherche de Chapman (1999) qui étudiait la dénomination des noms de fichiers. Le pré-test d'une séance de simulation réalisée avec deux volontaires a permis également de s'assurer que ce nombre d'essais ne rendait pas « trop longue » et « ennuyeuse » une séance de simulation. D'un point de vue statistique, il aurait été souhaitable de ne pas permettre aux répondants d'abandonner le processus de repérage d'un document et d'avoir un nombre d'essais illimité jusqu'au succès du repérage en vue d'observer, pour chaque répondant, le nombre réel d'essais nécessaires pour repérer un document et ainsi d'éviter les données censurées. En ce qui concerne la validité des résultats, ce choix aurait pu cependant donner lieu à des données invalides causées par des réponses hâtives de répondants pressés « d'en finir » pour passer au repérage d'un autre document.

Le Tableau 4.5 décrit le nombre de documents repérés mis en relation avec le nombre d'essais qui a été requis pour repérer ces documents, quel que soit le schéma de classification utilisé et quel que soit le document ciblé.

Tableau 4.5 Proportion de documents repérés et proportion de documents non repérés en rapport avec le nombre d'essais pour l'ensemble des schémas de classification

Nombre d'essais	<i>Succès</i> <i>(Documents repérés)</i>		<i>Échec</i> <i>(Documents non repérés)</i>		<i>Total des documents</i>	
	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>
1	97.3	930	2.7	26	100	956
2	73.8	177	26.3	63	100	240
3	32.4	66	67.6	138	100	204
Total	83.8	1173	16.2	227	100	1400

Les 70 répondants avaient chacun pour tâche de repérer 20 documents, ce qui porte à 1400 le nombre total d'observations. Alors qu'une grande proportion de documents a été repérée dès le premier essai (97.3 %) (n = 930), on observe une proportion de documents repérés qui va en décroissant par rapport à l'augmentation du nombre d'essais.

Ces données empiriques nous confortent dans la décision de limiter à trois le nombre d'essais ou de tentatives possibles pour repérer un document. La faible proportion de documents repérés au troisième essai (32.4 %) pourrait indiquer l'absence d'effet d'apprentissage ou l'influence de l'effet de fatigue.

4.3.1.3 Différences observées sur le plan de l'efficacité du repérage

Notre recherche visait à vérifier l'hypothèse qu'il existe des différences sur le plan de l'efficacité du repérage de documents ciblés réalisé à l'aide de schémas de classification distincts aux plans structurel, logique et sémantique. Cette section présente les premières observations générales qui ont pu être réalisées à partir des statistiques descriptives décrites ci-après.

Le Tableau 4.6 décrit, pour chacun des cinq schémas de classification testés, le nombre et le pourcentage des documents qui ont été repérés (i.e., succès du repérage), le nombre et le pourcentage de documents qui n'ont pas été repérés (i.e., échec à repérer un document après trois essais), et le nombre et le pourcentage d'abandons (i.e., répondants

qui ne se sont pas rendus jusqu'au 3^{ème} essai ou qui ont choisi d'abandonner la recherche au cours du 3^{ème} essai).

Tableau 4.6 Schéma de classification et succès/échec/abandon

Schémas	Succès		Échec (après 3 essais)		Abandon		Total	
	%	N	%	N	%	N	%	N
A	16.6	233	1.9	27	1.4	20	20.0	280
B	17.1	239	1.7	24	1.2	17	20.0	280
C	15.7	220	2.9	40	1.4	20	20.0	280
D	16.1	226	2.4	33	1.5	21	20.0	280
E	18.2	255	1.0	14	0.8	11	20.0	280
Total	83.8	1173	9.9	138	6.4	89	100.0	1400

Ces statistiques descriptives relatives aux proportions de succès, d'échec et d'abandon suggèrent qu'il existe des différences sur le plan de l'efficacité du repérage réalisé à l'aide des schémas de classification testés. Ces premières observations vont dans le sens des hypothèses puisque c'est le schéma E, dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations, qui a le plus faible taux d'échec (1 %) et d'abandon (0.8 %) (Voir Tableau 4.6 Schéma de classification et succès/échec/abandon) ainsi qu'une proportion plus élevée de documents repérés dès le premier essai (78.6 %) (Voir Tableau 4.7).

Tableau 4.7 Proportion de documents repérés pour chacun des schémas de classification selon le nombre d'essais

Schémas	Proportion de documents repérés			
	Essai 1	Essai 2	Essai 3	Total des essais
A	62.1 %	15.0 %	6.1 %	83.2 %
B	66.1 %	14.3 %	5.0 %	85.4 %
C	61.8 %	11.8 %	5.0 %	78.6 %
D	63.6 %	12.1 %	5.0 %	80.7 %
E	78.6 %	10 %	2.5 %	91.1 %

Ces données empiriques ne permettent toutefois pas de faire des inférences statistiques et d'affirmer que ces différences sont statistiquement significatives car les 1400 observations ne sont pas toutes indépendantes les unes des autres (i.e., existence d'une relation entre les mesures à l'intérieur d'un même répondant en raison, par exemple, de l'effet de fatigue et de l'effet d'apprentissage). Les sections suivantes présentent les résultats de l'analyse de variance à mesures répétées qui nous a permis de déterminer si les différences observées doivent être attribuées à la variation naturelle à l'intérieur de l'échantillon ou s'il y a des raisons de croire que l'efficacité du repérage diffère en fonction du schéma de classification utilisé.

4.3.2 Tests d'hypothèses sur la proportion de documents repérés

La première hypothèse spécifique de recherche était :

HS1. Un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus de documents.

La variable dépendante utilisée pour tester cette hypothèse était la proportion de documents repérés pour chacun des cinq schémas de classification. Nous rappelons que chaque répondant devait repérer 20 documents en utilisant à tous les quatre documents un schéma de classification différent. Pour chaque schéma de classification, on a calculé la proportion de documents repérés parmi les quatre documents recherchés. Le Tableau 4.8 décrit le nombre de documents repérés à l'aide de chacun des schémas de classification ainsi que le nombre et la proportion de répondants ayant été capables de réaliser ce repérage.

Tableau 4.8 Fréquences de la proportion de documents repérés par schéma de classification

	<i>Nombre de documents repérés</i>	<i>Nombre de répondants (N)</i>	<i>Proportion des répondants (%)</i>
Schéma A	1	2	2.9
	2	7	10.0
	3	27	38.6
	4	34	48.6
	Total	70	100.0
Schéma B	2	8	11.4
	3	25	35.7
	4	37	52.9
	Total	70	100.0
Schéma C	1	2	2.9
	2	9	12.9
	3	36	51.4
	4	23	32.9
	Total	70	100.0
Schéma D	1	1	1.4
	2	11	15.7
	3	29	41.4
	4	29	41.4
	Total	70	100.0
Schéma E	2	4	5.7
	3	17	24.3
	4	49	70.0
	Total	70	100.0

On observe qu'une plus grande proportion de répondants (70 %) (n = 49) est capable de repérer les quatre documents en utilisant le schéma E, dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations.

Le Tableau 4.9 décrit les différences entre les schémas de classification en rapport avec la proportion moyenne de documents repérés. L'analyse des données empiriques indique que le schéma E obtient une proportion moyenne de documents repérés la plus élevée (91.07 %) alors que le schéma C obtient la proportion de documents repérés la plus faible (78.57 %).

Tableau 4.9 Proportion moyenne de documents repérés pour chacun des schémas de classification

<i>Schémas</i>	<i>Moyenne</i>	<i>Écart-type</i>	<i>N</i>
A	.8321	.19375	70
B	.8536	.17285	70
C	.7857	.18692	70
D	.8071	.19116	70
E	.9107	.14766	70

L'analyse de variance à mesures répétées sur la proportion de documents repérés par schéma de classification a permis de vérifier si ces différences sont significatives. Les résultats du Tableau 4.10 nous autorisent à rejeter l'hypothèse nulle qu'en moyenne la proportion de documents repérés pour chacun des schémas de classification est la même car la valeur p a été quantifiée très inférieure de 0.05 ($p < 0.001$). Les différences entre la proportion de documents repérés à l'aide des cinq schémas de classification sont significatives au seuil de signification 0.05.

Tableau 4.10 Test des effets intra répondants

<i>Source</i>	<i>Somme des carrés de type III</i>	<i>ddl</i>	<i>Moyenne des carrés</i>	<i>F</i>	<i>Signification</i>
struct	.648	4	.162	5.877	.000
Erreur(struct)	7.603	276	.028		

Comme l'indique le Tableau 4.11, les différences entre les proportions de documents repérés sont significatives uniquement entre le schéma E et les schémas A ($p = 0.032$), C ($p < 0.001$) et D ($p = 0.009$). On n'observe pas de différence significative entre le schéma E et le schéma B ($p = 0.379$).

Tableau 4.11 Comparaisons des moyennes par paires de schémas de classification

<i>Schémas comparés</i>	<i>Différences entre les moyennes</i>	<i>Seuil de signification^a</i>
A vs. B	-.021	1.000
A vs. C	.046	.963
A vs. D	.025	1.000
A vs. E	-.079(*)	.032
B vs. C	.068	.230
B vs. D	.046	.963
B vs. E	-.057	.379
C vs. D	-.021	1.000
C vs. E	-.125(*)	.000
D vs. E	-.104(*)	.009

(*) La différence de moyennes est significative au niveau .05
^a Ajustement des comparaisons multiples : Test de Bonferroni¹³

La première hypothèse spécifique de recherche (HS1) est partiellement confirmée :

Il est possible de confirmer l'hypothèse qu'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations (Schéma E) permet de repérer plus de documents, à la différence d'un schéma de classification personnel (Schéma A), d'un schéma de classification à un seul niveau hiérarchique (Schéma C) et d'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique personnelle est celle de l'employé A (Schéma D).

Les résultats ne permettent pas de confirmer cette hypothèse quand la sémantique d'un schéma de classification personnel a été rendue explicite par une définition des classes (Schéma B).

¹³ Le test de Bonferroni pour les comparaisons sur l'ensemble des schémas de classification permet d'ajouter une pénalité et de garantir ainsi le niveau de confiance global. Ce test permet d'avoir globalement un niveau significatif de 5 %.

4.3.3 Tests d'hypothèses sur le temps moyen du repérage

La deuxième hypothèse spécifique de recherche était :

HS2. Un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus rapidement les documents.

La variable dépendante utilisée pour tester cette hypothèse était le temps moyen requis (mesuré en secondes) pour repérer les documents ciblés à l'aide de chacun des cinq schémas de classification. Lorsqu'un répondant ne réussissait pas à repérer un document (i.e., échec après trois essais ou abandon), le temps mesuré pour le repérage de ce document était considéré comme une donnée censurée. Autrement dit, ce temps n'était pas inclus dans la moyenne (i.e., valeur manquante) et on considérait uniquement le temps moyen des autres documents repérés à l'aide du même schéma de classification.

Deux analyses statistiques différentes ont été réalisées en vue de s'assurer de la validité des résultats : la première analyse fut réalisée pour les situations où le répondant trouvait un minimum de deux documents sur quatre (i.e., analyse d'un minimum de 50 % des observations) à l'aide d'un même schéma de classification. La deuxième analyse était réalisée lorsque le répondant trouvait un minimum de trois documents sur quatre (i.e., analyse d'un minimum de 75 % des observations) toujours à l'aide d'un même schéma de classification.

Dans la présente expérimentation, seulement quatre répondants (5.71 % des répondants) ont trouvé moins de deux documents sur quatre en utilisant un des cinq schémas de classification et ont donc été exclus de la première analyse (n = 66) (Voir Tableau 4.12).

Tableau 4.12 Moyenne du temps moyen des réussites (mesuré en secondes) pour les répondants ayant repéré deux documents ou plus

<i>Schémas</i>	<i>Moyenne</i>	<i>Ecart-type</i>	<i>N</i>
A	70.3840	47.10418	66
B	79.5943	50.32045	66
C	64.7149	38.44922	66
D	55.3266	41.07871	66
E	47.4734	36.77577	66

Trente-deux répondants (45.71 % des répondants) ont trouvé moins de trois documents sur quatre dans un des cinq schémas de classification et ont été exclus de la deuxième analyse (n = 38) (Voir Tableau 4.13).

Tableau 4.13 Moyenne du temps moyen requis pour repérer les documents ciblés (mesuré en secondes) pour les répondants ayant repéré trois documents ou plus

<i>Schémas</i>	<i>Moyenne</i>	<i>Ecart-type</i>	<i>N</i>
A	67.6909	40.90768	38
B	81.1393	52.51817	38
C	67.8499	38.48468	38
D	59.5896	37.98877	38
E	44.6657	23.65979	38

Les deux analyses ont donné des résultats semblables et des conclusions identiques : il est possible d'observer des différences entre les cinq schémas de classification en rapport avec le temps moyen requis pour repérer les documents ciblés. Le schéma E permet de repérer plus rapidement les documents ciblés (entre 44.66 et 47.47 secondes en moyenne).

L'analyse de variance à mesures répétées sur le temps moyen requis pour repérer les documents ciblés a été calculée pour les répondants ayant repéré deux documents ou plus par schéma de classification (Voir Tableau 4.14). Cette analyse statistique nous autorise à rejeter l'hypothèse nulle que toutes les moyennes sont égales dans l'absolu car $p < 0.05$. Les différences de performance du repérage entre les cinq schémas de classification sont significatives.

Tableau 4.14 Tests des effets intra-répondants

<i>Source</i>	<i>Somme des carrés de type III</i>	<i>ddl</i>	<i>Moyenne des carrés</i>	<i>F</i>	<i>Signification</i>
structur	41682.227	4	10420.557	7.605	.000
Erreur(structur)	356249.929	260	1370.192		

Comme l'indique le Tableau 4.15, les différences entre le temps moyen requis pour repérer les documents ciblés sont significatives entre le schéma E et les schémas A ($p = 0.002$), B ($p < 0.001$) et C ($p = 0.037$). On n'observe pas de différence significative entre le schéma E et le schéma D ($p = 1$).

Tableau 4.15 Comparaisons des moyennes par paires de schémas de classification

<i>Schémas comparés</i>	<i>Différences entre les moyennes</i>	<i>Seuil de signification^a</i>
A vs. B	-9.210	1.000
A vs. C	5.669	1.000
A vs. D	15.057	.221
A vs. E	22.911(*)	.002
B vs. C	14.879	.181
B vs. D	24.268(*)	.021
B vs. E	32.121(*)	.000
C vs. D	9.388	1.000
C vs. E	17.241(*)	.037
D vs. E	7.853	1.000

(*) La différence de moyennes est significative au niveau .05

^a Ajustement des comparaisons multiples : Test de Bonferroni

La deuxième hypothèse spécifique de recherche (HS2) est partiellement confirmée :

Il est possible de confirmer l'hypothèse qu'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations (Schéma E) permet de repérer plus rapidement les documents, à la différence d'un schéma de classification personnel (Schéma A), d'un schéma de classification à un seul niveau (Schéma C) et d'un schéma de classification personnel dont les classes ont été définies (Schéma B).

Les résultats ne permettent pas de confirmer cette hypothèse quand le repérage des documents est réalisé à l'aide d'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique personnelle est celle de l'employé A (Schéma D).

4.3.4 Tests d'hypothèses sur la proportion de documents repérés dès le premier essai

La troisième hypothèse spécifique de recherche était :

HS3. Un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus de documents dès le premier essai.

La variable dépendante utilisée pour tester cette hypothèse était la proportion de documents repérés dès le premier essai en rapport avec chacun des cinq schémas de classification. Tout comme pour la variable dépendante « temps moyen » (Voir section 4.3.3 Tests d'hypothèses sur le temps moyen du repérage), deux analyses statistiques différentes ont été réalisées en vue de s'assurer de la validité des résultats. Une première analyse fut réalisée pour les situations où le répondant repérait un minimum de deux documents sur quatre à l'aide d'un même schéma de classification. Une deuxième analyse fut réalisée lorsque le répondant trouvait un minimum de trois documents sur quatre toujours à l'aide d'un même schéma de classification. Lorsqu'un répondant

trouvait moins de deux documents (pour la première analyse) ou moins de trois documents (pour la deuxième analyse), cette valeur n'était pas calculée et on considérait la proportion de documents trouvés en un essai comme une valeur manquante. Si le répondant abandonnait au cours du premier essai ou suivant le premier essai, on considérait également la proportion de documents trouvés en un essai comme une valeur manquante (Voir Tableau 4.16).

Tableau 4.16 Données valides et manquantes pour la proportion de documents trouvés en un essai (au moins deux documents repérés) par schéma de classification

		<i>Schéma A</i>	<i>Schéma B</i>	<i>Schéma C</i>	<i>Schéma D</i>	<i>Schéma E</i>
N	Valide	68	70	68	69	70
	Manquante	2	0	2	1	0

Les Tableaux 17.1 à 17.5 en Annexe 17 Analyse de la variable « Essai » indiquent les fréquences pour la proportion de documents trouvés dès le premier essai dans les situations où le répondant repère au moins deux documents sur quatre en utilisant un même schéma de classification. On observe notamment que pour le schéma E, 40 % des répondants ont trouvé les quatre documents dès le premier essai (Voir Tableau 17.5 en Annexe 17 Analyse de la variable « Essai »).

Le Tableau 4.17 décrit la proportion moyenne de documents trouvés en un essai pour chacun des schémas de classification.

Tableau 4.17 Proportion de documents trouvés en un essai pour chacun des schémas de classification (un minimum de deux documents repérés)

<i>Schémas</i>	<i>Moyenne</i>	<i>Ecart-type</i>	<i>N</i>
A	.6402	.19675	66
B	.6705	.17628	66
C	.6250	.24515	66
D	.6439	.20597	66
E	.7955	.20526	66

On observe des différences entre les schémas de classification et surtout le schéma E au moyen duquel la proportion moyenne des documents repérés dès le premier

essai est la plus élevée (79.55 %) comparativement aux schémas A (64.02 %), B (67.05 %), C (62.50 %) et D (64.39 %).

L'analyse de variance à mesures répétées sur la proportion de documents trouvés en un essai pour les répondants ayant repéré un minimum de deux documents par schéma de classification (Voir Tableau 4.18) nous autorise à rejeter l'hypothèse nulle que toutes les moyennes sont égales (puisque $p < 0.001$).

Tableau 4.18 Tests des effets intra-répondants

<i>Source</i>	<i>Somme des carrés de type III</i>	<i>ddl</i>	<i>Moyenne des carrés</i>	<i>F</i>	<i>Signification</i>
proessai	1.268	4	.317	8.255	.000
Erreur(proessai)	9.982	260	.038		

Comme l'indique le Tableau 4.19, les différences entre la proportion de documents trouvés en un essai pour les répondants ayant repéré deux documents ou plus sont significatives entre le schéma E et tous les autres schémas.

Tableau 4.19 Comparaisons des moyennes par paires de schémas de classification

<i>Schémas comparés</i>	<i>Différences entre les moyennes</i>	<i>Seuil de signification^a</i>
A vs. B	-.030	1.000
A vs. C	.015	1.000
A vs. D	-.004	1.000
A vs. E	-.155(*)	.000
B vs. C	.045	1.000
B vs. D	.027	1.000
B vs. E	-.125(*)	.005
C vs. D	-.019	1.000
C vs. E	-.170(*)	.001
D vs. E	-.152(*)	.001

* La différence de moyennes est significative au niveau .05

^a Ajustement des comparaisons multiples : Bonferroni

Les résultats de l'analyse statistique effectuée lorsque les répondants trouvaient un minimum de trois documents (Voir Tableau 4.20 et Tableaux 17.10 à 17.18 en Annexe 17 Analyse de la variable « Essai ») ne changent pas ces conclusions.

Tableau 4.20 Proportion de documents trouvés en un essai pour chacun des schémas de classification (un minimum de trois documents repérés)

<i>Schémas</i>	<i>Moyenne</i>	<i>Ecart-type</i>	<i>N</i>
A	.6513	.19745	38
B	.6776	.16345	38
C	.6447	.25089	38
D	.6842	.20701	38
E	.8355	.17703	38

On observe également des différences au plan de la proportion de documents repérés dès le premier essai lorsque les répondants repèrent un minimum de trois documents sur quatre. Le schéma E obtient la proportion moyenne des documents repérés dès le premier essai la plus élevée (83.55 %) comparativement aux schémas A (65.13 %), B (67.76 %), C (64.47 %) et D (68.42 %).

L'analyse de variance à mesures répétées sur la proportion de documents repérés dès le premier essai pour les répondants ayant repéré trois documents ou plus par schéma de classification (Voir les résultats du Tableau 17.17 en Annexe 17 Analyse de la variable « Essai ») nous autorise à rejeter l'hypothèse nulle que toutes les moyennes sont égales (puisque $p < 0.001$). La proportion moyenne de documents repérés dès le premier essai diffère de manière significative en fonction des cinq schémas de classification pour les situations où les répondants repèrent un minimum de deux ou un minimum de trois documents sur quatre.

Les résultats confirment la troisième hypothèse spécifique de recherche (HS3) :

Il est possible de confirmer l'hypothèse qu'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations (Schéma E) augmente de manière significative la proportion de documents repérés dès le premier essai, à la différence d'un schéma de classification personnel (Schéma A), d'un schéma de classification personnel dont les classes ont été définies (Schéma B), d'un schéma de classification à un seul niveau (Schéma C) et d'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique personnelle est celle de l'employé A (Schéma D).

En résumé, les résultats de la simulation confirment l'existence de différences significatives sur le plan de l'efficacité du repérage des documents électroniques réalisé à l'aide de schémas de classification distincts aux plans structurel, logique et sémantique. Le Tableau 4.21 résume les résultats par rapport aux hypothèses spécifiques de recherche.

Tableau 4.21 Synthèse de résultats par rapport aux hypothèses spécifiques de recherche

<i>Paires de schémas de classification</i>	<i>Critères d'efficacité du repérage</i>		
	<i>Proportion de réussite (HS1)</i>	<i>Temps moyen (HS2)</i>	<i>Nombre d'essais (HS3)</i>
Schéma E vs. Schéma A	Hypothèse confirmée	Hypothèse confirmée	Hypothèse confirmée
Schéma E vs. Schéma B	Non confirmée	Hypothèse confirmée	Hypothèse confirmée
Schéma E vs. Schéma C	Hypothèse confirmée	Hypothèse confirmée	Hypothèse confirmée
Schéma E vs. Schéma D	Hypothèse confirmée	Non confirmée	Hypothèse confirmée
	Hypothèse (HS1) partiellement confirmée	Hypothèse (HS2) partiellement confirmée	Hypothèse (HS3) totalement confirmée

Une seule hypothèse de recherche (HS3) a été totalement confirmée :

On observe des différences significatives entre les proportions de documents repérés dès le premier essai à l'aide du schéma de classification dont la macro-structure est peu étendue, la logique est partiellement basée sur des classes d'activités et la sémantique utilise peu d'abréviations (Schéma E) et les schémas A, B, C et D.

Les deux autres hypothèses de recherche (HS1, HS2) ont été partiellement confirmées :

On observe des différences significatives entre les proportions de documents repérés à l'aide du schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations (Schéma E) et les schémas A, C et D. Cette différence n'est toutefois pas significative entre le schéma E et un schéma de classification personnel dont la sémantique personnelle a été rendue explicite par une définition des classes (Schéma B).

On observe des différences significatives entre le temps moyen requis pour repérer des documents ciblés à l'aide du schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations (Schéma E) et les schémas A, B et C. On n'observe pas de différence significative entre le schéma E et un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique personnelle est celle de l'employé A (Schéma D).

4.4 Conclusion

Notre recherche visait à décrire les caractéristiques structurelles, logiques et sémantiques d'un échantillon de schémas de classification utilisés par deux catégories d'employés pour l'organisation des documents administratifs électroniques qui sont sous leur contrôle direct. Notre recherche avait également comme objectif de vérifier les différences sur le plan de l'efficacité du repérage de documents électroniques réalisé à l'aide de schémas de classification ayant des caractéristiques structurelles, logiques et sémantiques variées.

Examinés globalement, les schémas de classification personnels diffèrent sensiblement les uns des autres aux plans structurel, logique et sémantique. Cependant, les résultats de l'analyse systématique de chacun des indicateurs de mesure des dimensions structurelles, logiques et sémantiques ont révélé plusieurs traits communs à une majorité de ces schémas. Ces derniers se caractérisent au plan structurel par une macro-structure généralement étendue et une micro-structure peu étendue, une structure peu profonde, complexe et déséquilibrée. Les employés privilégient un regroupement thématique des classes et un ordonnancement alphabétique. Les schémas de classification personnels observés se caractérisent au plan sémantique par des noms de classes dont la longueur (en nombre de caractères) est relativement homogène d'un schéma à un autre, par des stratégies d'abréviation qui privilégient l'usage d'acronymes et par une redondance structurelle et conceptuelle à l'intérieur des noms de classes généralement élevée dès le premier niveau de la hiérarchie.

Les résultats obtenus révèlent des différences significatives sur le plan de l'efficacité du repérage réalisé à l'aide de schémas de classification ayant des caractéristiques variées. Par ailleurs, les résultats n'ont pas permis d'identifier un modèle de schéma de classification personnel qui permettrait à la fois d'augmenter de manière significative la probabilité de repérer avec succès un document électronique, tout en minimisant le temps requis pour compléter la tâche ainsi que le risque d'erreur. Les résultats ont toutefois permis d'identifier les caractéristiques qui ont une influence soit

sur le succès, soit sur la rapidité, soit sur le risque d'erreur du repérage de documents électroniques ciblés.

Les résultats ont révélé qu'une sémantique personnelle rendue explicite par des définitions de classes permettrait de repérer les documents quelles que soient les caractéristiques structurelles et logiques du schéma de classification utilisé. Par ailleurs, les résultats révèlent qu'une macro-structure peu étendue et une logique basée partiellement sur une division par classes d'activités affectent de manière significative la rapidité du repérage des documents ciblés quelle que soit la sémantique utilisée (personnelle ou institutionnelle). Enfin, une macro-structure peu étendue, une logique basée partiellement sur une division par classes d'activités et une sémantique qui utilise peu d'abréviations minimisent de manière significative le risque d'erreur au repérage.

La discussion des résultats fait l'objet du prochain chapitre.

5 Discussion

5.1 Introduction

Nous avons présenté au chapitre précédent les résultats détaillés de l'analyse des données recueillies relatives aux caractéristiques des schémas de classification personnels et à la simulation du repérage de documents. Dans le présent chapitre, nous discutons des résultats de façon générale. La discussion des résultats est organisée en trois sections faisant référence aux deux objectifs spécifiques qui sous-tendent notre étude et aux limites de notre recherche. Dans un premier temps, nous discutons des caractéristiques structurelles, logiques et sémantiques des schémas de classification personnels que nous avons observés. Par la suite, nous interprétons les résultats relatifs à l'efficacité du repérage réalisé à l'aide de cinq schémas de classification testés. Pour terminer, nous présentons les limites de notre recherche.

5.2 Caractéristiques des schémas de classification personnels observés

Nous discutons dans cette section des caractéristiques structurelles, logiques et sémantiques identifiées lors de l'analyse de 21 schémas de classification utilisés par des techniciennes à la coordination de travail de bureau et des coordonnateurs de stages pour organiser les documents administratifs électroniques qui sont sous leur contrôle direct. La discussion des résultats relatifs aux caractéristiques des schémas de classification est organisée en trois sections faisant référence aux trois types de caractéristiques analysés : (1) les caractéristiques structurelles, (2) les caractéristiques logiques et (3) les caractéristiques sémantiques.

5.2.1 Caractéristiques structurelles des schémas de classification observés

Nous présentons dans cette section une discussion des résultats pour les quatre variables relatives à la dimension structurelle des schémas de classification soit l'étendue, la profondeur, la complexité et l'équilibre des structures classificatoires.

5.2.1.1 Étendue de la structure

Les résultats montrent qu'il existe de grandes différences dans l'étendue de la macro-structure (i.e., structure de base) des schémas de classification personnels, laquelle peut varier de 6 à 105 classes principales. La moyenne de 28.14 classes principales est bien au-dessus du maximum de 10 classes jugé suffisant pour décrire efficacement des ressources documentaires dans un domaine ou dans un contexte administratif donné (Roberge, 1985; Hudon, Mas & Gazo, 2005). Étonnamment, le plus grand nombre de classes principales n'est pas systématiquement associé aux structures les plus complexes (i.e., schémas de classification dont le total de classes est le plus élevé).

Deux facteurs peuvent influencer la variété observée dans l'étendue de la macro-structure des schémas de classification personnels : (1) le premier critère de division privilégié par l'employé et (2) l'utilisation d'un schéma de classification qui reproduit la structure classificatoire des documents sur support papier.

L'étendue de la macro-structure varie en fonction du premier critère de division privilégié par les employés. Les deux schémas de classification ayant le plus faible nombre de classes principales (i.e., six à sept classes principales) privilégient une logique basée soit sur des classes d'activités génériques (par exemple, « Administration », « Communications », « Formations », « Publicité ») soit sur des classes d'applications (par exemple, « Word », « WordPerfect », « PowerPoint », « Excel »). À l'inverse, les schémas de classification ayant un plus grand nombre de classes principales (i.e., au moins une quarantaine de classes principales) privilégient un regroupement thématique plus spécifique (par exemple, « Plagiat ») ou un regroupement mixte qui combine plusieurs concepts (par exemple, « Correspondance 2004 »).

L'utilisation d'un schéma de classification des documents électroniques qui reproduit intégralement la structure classificatoire conçue pour organiser les documents sur support papier peut également être un facteur d'influence de l'étendue de la macro-structure. Lors des entrevues, quatre employés interrogés ont dit utiliser un schéma de classification des documents électroniques identique au schéma de classification conçu pour l'organisation des documents sur support papier. En rapport avec un nombre souvent élevé de documents relatifs à diverses activités, le schéma de classification conçu pour l'organisation des documents sur support papier comporte généralement de nombreuses classes principales, ce qui explique la macro-structure également étendue du schéma de classification des documents électroniques correspondant.

Bien que l'utilisation d'un même schéma de classification soit théoriquement possible dans n'importe quel environnement (analogique ou électronique) (Molholt, 1995), certaines caractéristiques (par exemple, le nombre de classes principales ou subordonnées) du schéma de classification originellement conçu pour organiser les documents sur support papier sont le plus souvent modifiées par les archivistes pour être adaptées à l'environnement électronique (Sabourin, 2001; Beaupré, 2004; Chamard, 2004). Or, les données empiriques indiquent que les employés ne cherchent pas à modifier dans l'environnement électronique le schéma de classification conçu pour organiser les documents administratifs sur support papier qui sont sous leur contrôle direct. Bien au contraire, nos résultats révèlent la présence et le maintien délibéré de nombreux « répertoires vides » (i.e., classes utilisées dans l'environnement papier mais qui ne permettent de classer aucun document électronique) au niveau de la macro-structure de certains schémas de classification des documents électroniques.

L'utilisation d'un seul schéma de classification personnel pour l'organisation de l'ensemble des documents (quel que soit le support) permet de simplifier la tâche des employés non seulement en diminuant l'effort cognitif lié à la maîtrise et à la gestion de plusieurs structures classificatoires (International Organization for Standardization, 2001; Benfell, 2002; Tough & Moss, 2003; Conférence des recteurs et des principaux des universités du Québec, 2004; Poivre, 2004) mais aussi en leur rappelant l'existence

d'un dossier papier sur le même sujet ou la réalisation d'activités : « Quand les répertoires [classes] sont vides, cela veut dire qu'il n'y a rien eu cette année ».

En résumé, les employés interrogés s'accommodent le plus souvent de schémas de classification ayant un nombre de classes principales généralement élevé (macro-structure étendue). Les résultats ont révélé que l'étendue de la macro-structure peut être influencée par le premier critère de division privilégié et l'utilisation d'un schéma de classification identique au schéma utilisé pour organiser les documents sur support papier. Ces présomptions devront toutefois être vérifiées sur un échantillon plus large de schémas de classification personnels ayant des caractéristiques logiques et macro-structurelles variées.

5.2.1.2 Profondeur de la structure

Les résultats indiquent que les schémas de classification personnels observés ont un nombre moyen de 1.67 niveaux hiérarchiques. Ces résultats sont très inférieurs au nombre de niveaux hiérarchiques maximum préconisé (i.e., trois à quatre niveaux) pour la conception des schémas de classification documentaires et institutionnels (Beaupré, 2004; Chamard, 2004). Deux catégories de facteurs peuvent expliquer la préférence des employés pour des structures aussi peu profondes : (1) des facteurs cognitifs (par exemple, faciliter la navigation à travers la structure) et (2) des facteurs pragmatiques (par exemple, besoin d'accéder rapidement à l'information).

Un nombre de niveaux peu élevé éviterait aux employés d'être « désorientés » durant le processus de navigation (Jones, 2007). Les schémas de classification sont représentés dans l'environnement électronique par une structure arborescente de répertoires (ou classes) qui offre un « menu » de termes principaux à côté d'une icône représentant un dossier (Kirsh, 2001). En sélectionnant un des répertoires ou « points d'entrée » de premier niveau, les répertoires de second niveau sont dévoilés et ainsi de suite jusqu'à ce que l'on atteigne le document recherché. Plusieurs employés ont déclaré en entrevue être insatisfaits lorsque le nombre de niveaux est trop élevé car ils n'aiment pas « chercher trop loin », « c'est là qu'on se perd ». L'observation de schémas de

classification personnels dont la macro-structure est étendue mais peu profonde suggère que certains employés privilégient une lecture horizontale (plutôt que verticale) et une localisation spatiale de leurs classes qui s'appuie sur la capacité de la mémoire (Berlin *et al.*, 1993; Barreau & Nardi, 1995; Berchtold, Biliris & Panagos, 1999; Hertzum, 1999). En effet, pour localiser une classe il faut que cette dernière soit visible, ce qui n'est pas toujours le cas dans l'environnement électronique où les classes subordonnées organisées aux niveaux inférieurs sont visibles seulement à l'ouverture des répertoires.

Une structure peu profonde possède aussi l'avantage d'offrir un accès plus rapide aux documents électroniques en limitant le nombre d'étapes ou de manipulations avant d'atteindre le document ciblé. Un coordonnateur de stages a délibérément conçu un schéma de classification à un seul niveau hiérarchique en vue d'éviter un trop grand nombre de manipulations dans l'ouverture des répertoires. Il a également exprimé le désir de pouvoir « d'un clic arriver tout de suite à l'endroit où il aimerait aller » et d'éviter d'« aller à une place, à une place, à une place ». Il est cependant permis de douter qu'une telle structure demeure efficace au fil des ans pour organiser un nombre sans cesse croissant de documents administratifs électroniques.

5.2.1.3 Complexité de la structure

Les résultats indiquent un nombre total de classes très variable selon les schémas de classification personnels observés. Outre la relation connue entre le nombre total de classes et le nombre total de documents à organiser (Couture & collaborateurs, 1999; Gonçalves & Jorge, 2003), plusieurs facteurs professionnels peuvent influencer le niveau de complexité des schémas de classification personnels observés : (1) l'ancienneté de l'employé au sein de l'unité, (2) l'expérience de l'employé dans le poste occupé, (3) la formation reçue en gestion de documents et (4) la nature du poste occupé.

Il se peut qu'il y ait une relation entre l'ancienneté de l'employé et le niveau de complexité de son schéma de classification (par exemple, présence de nombreuses classes, nombreux regroupements). Un nouvel employé ne sait généralement pas d'emblée quels documents et informations lui seront utiles ou pertinents dans le cadre de

son travail (Whittaker & Hirschberg, 2001). Par précaution, cet employé peut être réticent à « supprimer » les documents déjà présents sur le poste informatique qu'il occupe et être porté à conserver une grande proportion de documents reçus. Cette présomption s'appuie sur l'observation de l'existence de documents dont la valeur informationnelle (i.e., l'utilité personnelle ou institutionnelle) est incertaine pour l'employé; ces documents sont organisés au premier niveau de la hiérarchie, à l'intérieur de classes « provisoires » ou directement sous le « répertoire racine » (i.e., « Mes Documents » ou son équivalent), dans l'attente d'une évaluation ultérieure de l'utilité de ces documents dans l'accomplissement des tâches.

L'attitude prudente des nouveaux employés envers la disposition des documents mentionnée par Whittaker & Hirschberg (2001) a été observée dans notre recherche non seulement envers les documents mais aussi envers le schéma de classification créé par l'employé qui occupait précédemment le poste. Certains employés en poste depuis cinq ans ou moins ont dit qu'ils ne modifieraient pas tout de suite le schéma de classification utilisé par leur prédécesseur. Au fil des années (entre un et cinq ans selon les employés interrogés), les employés disent acquérir une meilleure connaissance de leurs tâches et des documents utiles à l'accomplissement de celles-ci. Le schéma de classification existant est alors réorganisé, ou bien un nouveau schéma de classification est créé en fonction d'une meilleure connaissance de la valeur informationnelle des documents, des tâches à réaliser et de l'évolution des responsabilités de l'employé.

Plusieurs chercheurs ont observé que l'expérience dans le poste occupé peut aussi entraîner l'accumulation d'une plus grande quantité de documents incitant l'employé à consacrer des efforts à l'implantation d'un mode d'organisation plus élaboré (Lansdale, 1988; Barreau, 1995; Whittaker & Sidner, 1996; Jones, Dumais & Bruce, 2002; Henderson, 2003). Effectivement, les employés les plus expérimentés semblent utiliser des structures plus complexes et développent des stratégies, des « petits trucs observés à droite et à gauche », permettant d'améliorer l'organisation et le repérage ultérieur des documents administratifs électroniques qui sont sous leur contrôle direct.

De même, il se peut qu'il y ait un lien entre la formation reçue en gestion des documents et la complexité d'un schéma de classification personnel. Tous les employés ayant suivi une telle formation semblent plus à l'aise pour mettre en place et gérer des structures très complexes basées ou non sur le schéma de classification institutionnel. Plusieurs employés ont reconnu ne pas avoir eu connaissance de l'existence d'un tel outil et attendaient avec impatience de recevoir une formation en gestion des documents pour les aider à mieux structurer l'organisation des documents administratifs électroniques qui sont sous leur contrôle direct.

Plusieurs chercheurs ont déjà observé des différences dans la gestion des documents personnels basées sur le rôle professionnel (Kidd, 1994; Whittaker & Hirschberg, 2001). Ainsi, les schémas de classification des techniciennes et des coordonnateurs de stages ont des traits structurels communs sauf au plan de la complexité. Nos résultats confirment que le poste occupé peut avoir une influence sur la complexité des schémas de classification personnels développés. Dans notre échantillon, les employés qui conservent le plus grand nombre de documents électroniques et qui utilisent le plus grand nombre de classes sont des techniciennes, toutes responsables de la gestion des documents et mandatées ou enclines à vouloir « tout conserver ».

5.2.1.4 Équilibre de la structure

L'équilibre correspond à l'écart-type de la moyenne du nombre de classes à chaque niveau. Cette variable est un indicateur du degré d'équilibre des schémas de classification sur le plan de l'étendue de la structure. Autrement dit, l'équilibre nous permet de déterminer si un schéma de classification a un nombre moyen de sous-classes à chaque niveau qui est similaire d'une classe à l'autre (Gonçalves & Jorge, 2003). Les schémas de classification personnels observés sont généralement des structures peu équilibrées, le nombre moyen de classes variant considérablement d'un niveau à l'autre.

Un contexte de création « pragmatique » des classes et les différentes fonctionnalités que les employés attribuent à ces classes peuvent expliquer ce déséquilibre.

Un schéma de classification hiérarchique spécialisé doit être spécifiquement adapté à un fonds documentaire particulier en tenant compte à la fois du nombre de documents à organiser (par une approche inductive) et de la structure logique du domaine à classifier (par une approche déductive) (Maniez, 1987; Langridge, 1992; Canonne, 1993; Iyer, 1995). Le mode de conception des schémas de classification personnels observés correspond à la première de ces deux approches. Au fil des années, les schémas de classification personnels s'enrichissent de plus en plus de classes et de sous-classes pour répondre au besoin d'organiser un nombre croissant de documents électroniques portant sur des sujets variés ou relatifs à diverses activités. Autrement dit, certains employés créent naturellement une nouvelle classe lorsque la quantité de documents électroniques l'exige. Par exemple, un employé a déclaré en entrevue ne pas subdiviser une classe en sous-classes tant qu'il y a moins d'une dizaine de documents à organiser à l'intérieur de cette classe. Cette pratique applique intuitivement le principe classificatoire de « caution bibliographique » (« *Literary warrant* ») qui consiste à tenir compte du nombre de documents existants sur un sujet en vue de s'assurer d'élaborer un schéma de classification spécifique à l'ensemble documentaire (Langridge, 1992; Canonne, 1993; Iyer, 1995).

Les employés peuvent aussi créer des classes quel que soit le nombre de documents à organiser (« un seul document ou une dizaine »). Nous avons mentionné dans la section 5.2.1.1 Étendue que les classes, même « vides », permettent à l'employé de se rappeler de l'existence de documents sur support papier sur le même sujet. La création de certaines classes peut aussi avoir lieu dès l'ouverture d'un dossier, au moment de la création ou de la réception d'un document (par exemple, le curriculum vitae d'un étudiant pour un stage). Enfin, certaines classes « provisoires » sont créées en vue de regrouper les documents « éphémères » et d'en faciliter l'élimination ultérieure (Voir section 5.2.2.1 Critères de division logique).

Le contexte de création pragmatique des classes et leurs différentes fonctionnalités peuvent expliquer l'observation de schémas de classification personnels déséquilibrés et pas nécessairement compatibles entre eux aux plans structurel et logique.

En résumé, les caractéristiques structurelles des schémas de classification personnels observés ont été comprises et interprétées en tenant compte de plusieurs contextes. Il s'agit du contexte documentaire (par exemple, quantité de documents à organiser), du contexte professionnel (par exemple, nature du poste occupé), du contexte cognitif (par exemple, localisation spatiale des classes) et du contexte pragmatique (par exemple, besoin d'accéder rapidement à l'information). Les résultats confirment la nécessité de concevoir des schémas de classification très peu profonds dans un contexte de gestion décentralisée des ressources informationnelles pour répondre à la fois à des besoins d'économie cognitive (i.e., faciliter la navigation) et pragmatiques (i.e., accéder rapidement à l'information). De façon surprenante, les résultats nous apprennent que les employés s'accommodent d'une structure déséquilibrée, d'un nombre élevé de classes principales et de l'utilisation d'un schéma de classification identique pour l'organisation de l'ensemble des documents (analogiques et électroniques). De plus, les résultats indiquent que les employés ayant suivi une formation en gestion des documents sont plus enclins à concevoir et à utiliser des structures classificatoires très complexes pour organiser les documents administratifs électroniques qui sont sous leur contrôle direct. Enfin, les résultats révèlent qu'un schéma de classification personnel est bien plus qu'un outil d'organisation documentaire pour les employés puisqu'il peut être utilisé par ces derniers comme mémoire de travail (par exemple, rappel de l'existence d'une documentation sur support papier sur un même sujet) ou comme outil de planification de tâches (par exemple, évaluation d'un document) ou d'activités à venir (par exemple, évaluation des étudiants).

5.2.2 Caractéristiques logiques des schémas de classification observés

Nous présentons dans cette section une discussion des résultats pour les deux indicateurs de mesure retenus pour la dimension logique des schémas de classification personnels observés, soit les critères de division logique et l'ordre de succession des classes.

5.2.2.1 Critères de division logique

La classification désigne l'organisation d'entités (par exemple, des documents) en classes sur la base de caractéristiques ou critères de division communs (Maniez, 1987; Hunter, 2000). Le premier critère de division est un des éléments le plus important puisque c'est lui qui décrit et organise le fonds documentaire en classes logiques (Kwasnik, 1999). Cependant, même les documents les plus simples ont plusieurs caractéristiques (par exemple, le type de document ou le thème représenté) qui peuvent être utilisées pour faire des regroupements de documents. Ainsi, les possibilités de regroupements peuvent être infinies et les mêmes documents peuvent être classifiés de différentes manières (Langridge, 1992; Mai, 2000). Les résultats de notre recherche indiquent que les employés n'appliquent pas le principe de division logique puisqu'aucun critère de division ne prédomine aux deux premiers niveaux hiérarchiques. Les techniciennes et les coordonnateurs de stages interrogés font plutôt preuve de beaucoup de créativité dans l'organisation des documents administratifs électroniques qui sont sous leur contrôle direct, ce qui suggère une faible cohérence et prévisibilité de leur schéma de classification. En effet, d'un point de vue théorique, l'utilisation de plusieurs critères de division à un même niveau hiérarchique génère des classes qui ne sont pas mutuellement exclusives (i.e., classes distinctes les unes des autres) (Iyer, 1995; Kwasnik, 1999; Bowker & Star, 2000) et favorise l'incertitude au moment de la sélection d'une classe particulière (Van Der Walt, 1998). Or, les employés privilégient comme base de classification tantôt le thème du document, tantôt le type de document, tantôt l'activité dans laquelle il s'insère; le document peut de fait être classifié et repéré à plusieurs endroits.

Les employés interrogés ont cependant dit être généralement satisfaits de leur schéma de classification personnel et n'éprouvent que rarement de la difficulté à repérer un document donné conformément aux observations de Chapman (1999) et Boardman & Sasse (2004). Les résultats confirment que les employés, qui utilisent indifféremment plusieurs critères de division sans éprouver de difficulté à repérer ultérieurement les documents, se basent bien sur la capacité de leur mémoire (i.e., mémorisation d'une ou

de plusieurs caractéristiques du document recherché) pour classer et repérer les documents plutôt que sur un raisonnement logique (Lansdale, 1988). Cela dit, le processus continu de développement ou de mise à jour inhérent aux schémas de classification documentaire, est l'une des difficultés majeures exprimées par les employés qui remettent toujours à plus tard le « ménage » de leur schéma de classification personnel. Cette difficulté dans la « mise à jour », caractéristique d'un schéma de classification peu flexible (Molholt, 1995; Hertzum, 2004), peut non seulement s'expliquer par l'usage de critères de division variés qui ne facilite pas la mise à jour du schéma de classification mais aussi par un contexte de travail et un effort cognitif lié à la classification le plus souvent contraignants.

Qu'est-ce qui peut expliquer cette variété dans les critères de division logique des schémas de classification personnels observés? Les recherches antérieures suggèrent que le choix d'une stratégie de classification mise en place par les individus peut être influencé, entre autres, par (1) la nature des tâches à réaliser (Henderson, 2003), (2) le manque de temps (Whittaker & Hirschberg, 2001) et (3) la difficulté à catégoriser (Cole, 1981; Malone, 1983; Lansdale, 1988). Nos résultats révèlent également que le caractère partagé ou non d'un schéma de classification peut influencer le critère de division logique.

La nature des tâches à réaliser suggère qu'un employé peut avoir besoin d'accéder plus fréquemment et plus rapidement à certaines catégories spécifiques de documents. Par exemple, les techniciennes travaillent de manière récurrente avec des modèles et des types de documents (par exemple, des formulaires ou des étiquettes), de la documentation de référence sur différentes thématiques (par exemple, sur le plagiat) et sont appelées à gérer les documents de réunions (par exemple, des avis de convocation et des procès-verbaux). Cette catégorie particulière d'employé est plus encline à concevoir dès le premier niveau de la hiérarchie des classes spécifiques, représentant par exemple des « types de contenu » et des « thèmes », plutôt que des classes génériques telles que des classes d'activités (implicites pour l'employé) qui ajouteraient un niveau

hiérarchique et augmenteraient inutilement la profondeur de la structure classificatoire et le temps d'accès au document pertinent.

La variété des critères de division peut aussi s'expliquer par des contraintes de temps qui obligent souvent les employés à organiser rapidement les documents créés ou reçus dans le cadre de tâches souvent interrompues (Jones, 2007). Plutôt que de prendre le temps d'assigner le document à une classe déjà existante, les employés vont rapidement créer une classe dans laquelle ils « sauvegardent » le document, sans souci d'une quelconque cohérence avec les autres classes et en oubliant le plus souvent qu'une classe, souvent nommée différemment, existe déjà pour regrouper des documents similaires. Ce dédoublement des classes, déclaré indésirable par plusieurs employés, se manifeste à l'intérieur d'un même schéma de classification par une proportion élevée de classes principales redondantes et parfois équivalentes (par exemple, « Rencontres annuelles avec les milieux » et « Rencontres d'équipe »). De telles pratiques confirment l'éparpillement des documents relatifs à un même sujet et l'incohérence dans la dénomination personnelle des classes observés par Whittaker & Sidner (1996), Skupsky (1999) et Guercio (2002).

La difficulté cognitive reliée à la décision à prendre au moment de classer un document (i.e., assigner le document à une classe pré-existante ou concevoir une nouvelle classe) (Malone, 1983) conduit plusieurs employés à retarder le processus décisionnel en laissant les documents administratifs électroniques qui sont sous leur contrôle direct « à la racine », sous le répertoire « C:\Mes Documents » (ou son équivalent). Suite aux propos d'une technicienne : « Si on perd du temps à chercher où mettre un document, c'est qu'on ne sait pas où le classer et qu'il y a une ambiguïté sur les sujets », on pourrait envisager que la rapidité du classement des documents à l'aide d'un schéma de classification, et non seulement le repérage, puisse être un indicateur de mesure intéressant pour évaluer la qualité d'un schéma de classification.

Une autre stratégie observée visant à limiter l'effort cognitif lié à la catégorisation des documents consiste à organiser ces derniers sous une classe « Divers » ou son

équivalent. La présence de cette classe qualifiée par les employés de « fourre-tout » a été observée dans une majorité de schémas de classification personnels. L'utilisation de cette classe « Divers » pourrait correspondre à la classe « Généralités », très souvent observée et utilisée dans les schémas de classification documentaires et institutionnels. L'utilisation d'une telle classe est théoriquement justifiée lorsqu'elle est réservée à la classification des documents qui ne peuvent être classifiés dans aucune autre classe. Cependant, l'utilisation de cette classe, jugée parfois excessive par les employés, suggère une fonction davantage pragmatique que théorique dans un contexte de gestion décentralisée des ressources informationnelles. Autrement dit, l'utilisation fréquente par les employés de la classe « Divers » répondrait à un besoin d'économie cognitive et serait finalement révélatrice d'un schéma de classification peu fonctionnel.

Nos résultats nous apprennent que le partage d'un schéma de classification peut également influencer le critère de division logique. Les deux coordonnateurs de stages qui utilisent à la fois un schéma de classification partagé et un schéma de classification non partagé ont dit avoir conscience de mieux structurer l'organisation des documents électroniques quand ces derniers sont partagés avec un ou plusieurs collègues de travail. Les résultats indiquent la présence de classes très spécifiques et informatives à l'intérieur des schémas de classification partagés par des coordonnateurs qui privilégient un regroupement par « thème » ou qui « mélangent » ou combinent plusieurs concepts (par exemple, « Calendrier des tâches 2005-2006 »).

5.2.2.2 Ordre de succession des classes

L'ordre de succession des classes est, en théorie, relativement indifférent puisque ce qui importe c'est que ces classes demeurent exclusives (i.e., distinctes les unes des autres) (Maniez, 1987). Les résultats indiquent que les employés se contentent le plus souvent d'utiliser l'ordre de succession alphabétique des classes proposé par défaut par le système d'exploitation Windows. Quant aux quatre employés qui ont basé leur schéma de classification sur le schéma de classification institutionnel ou sur le schéma de classification personnel utilisé dans l'environnement papier, ils ne sont pas non plus enclins à changer l'ordre numérique proposé.

Si certains employés interrogés sont généralement satisfaits de l'ordre alphabétique ou numérique de leurs classes, l'analyse des résultats révèle que plus du quart d'entre eux manipulent occasionnellement les noms de classes pour imposer un ordre personnel. Cette manipulation consiste le plus souvent à inclure des caractères alphabétiques, numériques ou symboliques devant le premier caractère du nom de la classe. Ces observations ont été également relevées dans les travaux de Jones *et al.* (2005).

Plusieurs motivations mentionnées en entrevue par les employés sont à l'origine des modifications personnelles apportées à l'ordre alphabétique. Par exemple, il s'agit de l'importance d'un ordre qui permet la localisation « spatiale » de classes particulières (par exemple, les classes « importantes » ou « provisoires » seront localisées « en haut ou en bas de la liste ») et de la proximité souhaitée (par exemple, « à côté l'une de l'autre ») entre les classes ayant un lien entre elles. Nos résultats confirment le besoin de proximité spatiale entre les classes déjà observé par Jones *et al.* (2005). Ces observations pourraient être utiles dans la conception de schémas de classification qui permettent dans l'environnement électronique de modifier l'ordre (alphabétique ou numérique) de succession des classes pour répondre aux besoins pragmatiques et cognitifs exprimés par les employés.

En résumé, le non respect du principe de division logique semble être la règle dans un contexte de gestion décentralisée des ressources informationnelles. La variété des critères de division logique utilisés par les employés pour organiser les documents administratifs électroniques qui sont sous leur contrôle direct peut être expliquée par différents facteurs professionnels, cognitifs et pragmatiques tels que la nature des tâches à réaliser (et les besoins documentaires qui y sont rattachés), le manque de temps, la difficulté cognitive à catégoriser et le caractère partagé ou non d'un schéma de classification personnel. L'ordre de succession alphabétique des classes proposé par défaut dans l'environnement Windows semble satisfaire les employés. Cependant, cet ordonnancement alphabétique ne permet pas de contrôler l'ordre de présentation de classes souvent nommées différemment d'un individu à l'autre. Or, selon plusieurs

employés ayant été amenés à effectuer de fréquents remplacements, une présentation uniforme et familière des classes principales des schémas de classification personnels en faciliterait la compréhension et donc l'utilisation.

5.2.3 Caractéristiques sémantiques des schémas de classification observés

Nous présentons dans cette section une discussion des résultats pour les deux niveaux d'analyse retenus de la dimension sémantique des schémas de classification observés soit (1) l'analyse structurelle et conceptuelle des intitulés des classes pris hors-contexte (hors hiérarchie) (i.e., la longueur moyenne (en nombre de caractères) des intitulés d'un échantillon de classes et les principales stratégies d'abréviation utilisées) et (2) l'analyse structurelle et conceptuelle des intitulés des classes pris en contexte (à l'intérieur de la hiérarchie) (i.e., le degré de redondance structurelle et conceptuelle).

5.2.3.1 Analyse structurelle et conceptuelle des intitulés des classes pris hors-contexte

Pendant de nombreuses années, les archivistes ont recommandé l'utilisation d'un maximum de huit caractères pour nommer les classes, particulièrement en raison de contraintes informatiques (Saulnier, 1997-1998). Or, nos résultats indiquent que les employés utilisent des noms de classes d'une longueur moyenne de 16 caractères, soit le double de la limite longtemps préconisée. De plus, les résultats révèlent que l'utilisation d'un schéma de classification institutionnel ou d'un schéma de classification partagé avec des collègues de travail peut augmenter la longueur moyenne des noms de classes que nous avons observée. Ceci laisse présumer que les employés nomment leurs classes de façon plus brève lorsqu'ils conçoivent un schéma de classification pour leurs propres besoins.

Deux facteurs peuvent expliquer ces résultats : (1) l'assertion selon laquelle plus un énoncé est long plus il est significatif et (2) l'observation d'une forte proportion de noms de classes abrégés qui a comme corollaire une diminution de la longueur des noms de ces classes.

Il est reconnu que plus un descripteur est long, plus il représente un sujet de manière précise et détaillée (Éthier, 1989; Saffady, 2002). Nos résultats semblent indiquer que plus le nom d'une classe est long, plus il pourrait être significatif non seulement pour l'employé lui-même mais aussi pour une tierce personne. Le souci d'attribuer des noms plus longs, donc plus significatifs, aux classes destinées à être communiquées ou partagées par d'autres collègues de travail a été formellement exprimé par les employés utilisant des schémas de classification partagés mais aussi par des employés soucieux de la qualité de la dénomination de leurs propres classes (i.e., non partagées) de manière à ce « qu'elles soient repérables par n'importe qui ». L'attention particulière parfois accordée au nom des classes en vue d'augmenter leur fonction descriptive a déjà été relevée par Chapman (1999).

La longueur moyenne des noms de classes doit aussi être mise en relation avec les résultats qui indiquent une proportion élevée d'abréviations et particulièrement d'acronymes. Selon Jones (2004), la dénomination des classes, le plus souvent hâtive dans un contexte de travail, peut être une des raisons qui motive l'emploi de titres courts, le plus souvent abrégés et pas toujours significatifs pour une tierce personne (par exemple, « nn » ou « nnn »). Les employés interrogés n'hésitent pas à désigner leurs classes par des acronymes ou des « bouts de noms » qui sont facilement reconnaissables et significatifs au niveau institutionnel (par exemple, « BRI » pour « Bureau de la recherche institutionnelle ») ou local (i.e., utilisées à l'intérieur d'une même unité) (par exemple, « CO » pour « Communautaire »). On peut toutefois s'interroger sur les conséquences de telles pratiques au niveau du repérage par une tierce personne non familière avec le lexique personnel, local ou institutionnel.

Les noms des classes créés pour des besoins classificatoires personnels ne sont ni immuables ni constants, une même classe pouvant être nommée différemment par un même individu. Les « incohérences » (« *inconsistencies* ») dans la dénomination personnelle des classes relevées par Berlin *et al.* (1993, p. 29), Kidd (1994), Whittaker & Sidner (1996) et Whittaker & Hirschberg (2001) ont également été observées dans notre recherche. De plus, l'analyse des données recueillies permet d'observer une tendance à la

« troncature » des noms des classes au fil des années. Par exemple, la classe « Assignations en milieu communautaire » devient l'année suivante « Assignations CO » puis l'année d'après « CO ». Dans le même ordre d'idée, la classe « Suivi des cas problèmes » est renommée « Cas problèmes » puis « Cas prob ». Une telle simplification serait significative seulement pour les employés qui connaîtraient la relation entre « CO » et « Assignations en milieu communautaire » ou entre « Cas prob » et « Suivi des cas problèmes ». Il serait intéressant de vérifier l'existence d'une corrélation entre l'évolution de la longueur des noms de classes à travers le temps et l'ancienneté de l'employé, par exemple dans le cadre d'une étude longitudinale. Il ne serait pas étonnant de constater que la longueur moyenne des noms de classes diminue au fur et à mesure que l'employé s'approprie le discours et le lexique rattachés à sa fonction. On peut ainsi s'attendre à ce que certains employés ayant le plus d'ancienneté utilisent davantage d'abréviations personnelles, locales ou institutionnelles, devenues familières avec le temps.

5.2.3.2 Analyse structurelle et conceptuelle des intitulés des classes pris en contexte

Les résultats montrent que la proportion de classes redondantes structurellement (i.e., même sous-chaîne de caractères) et conceptuellement (i.e., même concept) varie selon le caractère partagé ou non d'un schéma de classification personnel. Les schémas de classification non partagés observés ont généralement un nombre de classes redondantes moins élevé que les schémas de classification partagés.

Le degré de redondance plus élevé observé dans les schémas de classification partagés doit être mis en relation avec les résultats qui indiquent la présence de classes très spécifiques et informatives à l'intérieur des schémas qui privilégient un regroupement thématique ou mixte (i.e., « mélange » de plusieurs concepts). La redondance structurelle et conceptuelle des noms de classes nous paraît inévitable quand les classes principales combinent plusieurs concepts (Voir Figure 5.1).

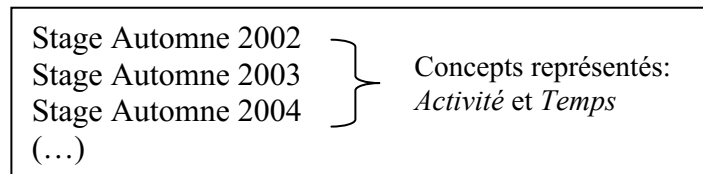


Figure 5.1 Exemple de classes redondantes au premier niveau.

Le degré de redondance structurelle et conceptuelle au niveau des classes principales nous apparaît être un signe d'éparpillement et pourrait être facilement corrigé en effectuant des regroupements. La redondance dans les classes principales n'est d'ailleurs pas préconisée car elle entraîne une certaine confusion et un dédoublement inutile des classes. Par exemple, un employé s'est rendu compte durant l'entrevue de la présence de deux classes intitulées « Dossier Vaccination permis travail » et « Permis de travail » et ne savait plus lui-même ce qui distinguait ces deux classes.

Alors que la redondance conceptuelle et structurelle observée au premier niveau hiérarchique d'un schéma de classification apparaît être un indicateur d'ambiguïté des classes, la redondance observée aux niveaux hiérarchiques inférieurs apparaît au contraire en augmenter l'univocité (Voir Figure 5.2).

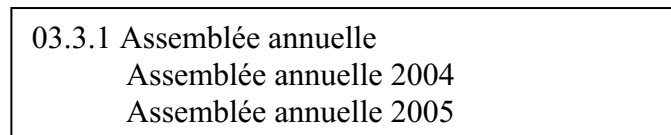


Figure 5.2 Exemple de classes redondantes au deuxième niveau.

La redondance dans les classes subordonnées permet de rendre ces dernières plus explicites en incluant dans leur intitulé le titre de la classe supérieure. Cette règle d'écriture a l'avantage de préciser le contexte dans lequel la classe doit être interprétée et contribue à éviter à l'employé d'être « désorienté » dans sa navigation à travers la structure.

En résumé, les employés qui utilisent des schémas de classification personnels partagés avec des collègues de travail ou qui ont été amenés à effectuer fréquemment des remplacements ont dit être soucieux et attentifs à créer des classes dont l'intitulé soit

compréhensible par autrui. Les classes « significatives » se caractérisent par des intitulés généralement plus longs et spécifiques (i.e., degré de redondance élevé) que ceux utilisés pour désigner les classes qui ne sont pas destinées à être communiquées ou partagées. Par ailleurs, dans un contexte de travail exécuté rapidement et souvent interrompu, les employés peuvent utiliser un nombre important d'abréviations pour nommer plus rapidement leurs classes. La proportion de classes abrégées peut augmenter au fil du temps, au point de rendre ces classes peu explicites pour un individu non familier avec les abréviations utilisées. Il serait intéressant d'effectuer une analyse lexicale et d'identifier systématiquement les abréviations les plus fréquemment utilisées par une même catégorie d'employés pour nommer leur domaine d'activités. Existe-t-il une nomenclature personnelle qui repose sur l'emploi d'un certain nombre de désignations officielles, à l'exclusion d'autres? Suite à une telle analyse, il serait possible d'envisager la conception et l'emploi d'un référentiel commun mis en place par la direction (par exemple, une liste d'abréviations ou une liste de sigles) qui permettrait d'officialiser et de définir les acronymes, troncatures et autres expressions abrégées qui ont un sens au niveau local ou au niveau institutionnel.

5.2.4 Conclusion

Les schémas de classification observés ont été qualifiés, avec raison, de « personnels » car ils correspondent bien, pour reprendre l'expression de Peyrelong & Accart (2002, p. 3), « à une logique d'acteur singulière dans un contexte de travail particulier ».

Si nous sommes d'accord avec les propos de Jones (2007) pour qui la gestion personnelle de l'information réalisée par un individu est unique, les résultats de notre recherche révèlent cependant plusieurs caractéristiques structurelles, logiques et sémantiques communes à une majorité de schémas de classification personnels observés. Le schéma de classification personnel « type » est composé d'une macro-structure étendue, d'une structure très peu profonde tout en étant déséquilibrée et dont la complexité varie selon la nature du poste occupé. Les employés utilisent des schémas de classification dont les critères de division sont variables, les intitulés des classes sont

généralement longs et souvent composés d'abréviations. On observe également un degré de redondance structurelle et conceptuelle élevé, tant au niveau de la macro-structure que de la micro-structure, ce qui est caractéristique d'un usage de classes très spécifiques.

L'analyse des résultats de notre recherche suggère que plusieurs facteurs, mentionnés dans les écrits, peuvent influencer les caractéristiques des schémas de classification, soit la quantité de documents à organiser (que cela soit à l'entrée en poste ou accumulés au fil des ans), la nature du poste occupé et les besoins documentaires qui y sont rattachés, les contraintes de temps, la formation reçue en gestion de documents, l'effort cognitif lié à la classification, l'ancienneté et l'expérience dans le poste occupé. Nos résultats révèlent l'existence possible d'autres facteurs pouvant influencer les caractéristiques des schémas, non mentionnés dans les écrits, soit le caractère partagé ou non partagé d'un schéma de classification ainsi que les multiples fonctions des classes. Enfin, nos résultats ont révélé qu'un schéma de classification personnel est non seulement conçu par les employés pour organiser et repérer les documents administratifs électroniques qui sont sous leur contrôle direct mais aussi pour assumer un rôle original de mémoire de travail (par exemple, rappel de l'existence d'une documentation papier sur un même sujet).

Certains schémas de classification personnels observés, qui peuvent être qualifiés de « pragmatiques », sont créés selon une approche inductive (« *bottom-up* ») (i.e., basée sur la quantité de documents à organiser). Un tel contexte de création peut expliquer pourquoi les schémas de classification personnels sont uniques et personnels, particulièrement au plan logique (i.e., critères de division très variés). D'autres schémas de classification personnels observés, que l'on peut qualifier d'« objectifs » (i.e., créés selon les besoins à plus long terme), sont conçus selon une approche déductive (« *top-down* »), (i.e., basée sur un cadre institutionnel ou sur un schéma de classification des documents sur support papier déjà existant). Ces derniers, plus rares, sont créés par des employés ayant acquis des connaissances théoriques ou une formation pratique en gestion des documents administratifs et qui ont investi du temps pour concevoir un

schéma de classification qui soit fonctionnel à long terme et non seulement pour répondre à des besoins immédiats.

Nous avons discuté dans cette section des caractéristiques structurelles, logiques et sémantiques des schémas de classification personnels. Les archivistes pourraient être tentés de se baser sur ces caractéristiques pour proposer un schéma de classification davantage adapté aux besoins des employés. Or, notre recherche est partie de la présomption que les caractéristiques des schémas de classification personnels ne peuvent être toutes satisfaisantes au plan de l'efficacité du repérage réalisé par une tierce personne. Si les individus paraissent pouvoir se retrouver plus ou moins efficacement dans leur propre schéma de classification, qu'en est-il de l'efficacité du repérage de documents réalisé par autrui à l'aide de l'un de ces schémas de classification personnels? Nous discutons ci-après des résultats de la deuxième partie de notre recherche, laquelle visait à vérifier les différences sur le plan de l'efficacité du repérage de documents électroniques dans un environnement contrôlé qui sont fonction du schéma de classification utilisé.

5.3 Schémas de classification et repérage

Notre recherche a permis de vérifier objectivement, dans le cadre d'une simulation, avec quelle efficacité cinq schémas de classification distincts aux plans structurel, logique et sémantique ont permis à un ensemble de 70 répondants de repérer des documents électroniques ciblés. L'efficacité du repérage de documents a été mesurée à l'aide de trois variables : la proportion de documents repérés, le temps moyen requis pour repérer les documents et la proportion de documents repérés dès le premier essai. Nous présentons dans cette section une discussion des résultats de l'analyse des données issues de cette expérimentation.

5.3.1 Schémas de classification et proportion de documents repérés

Nous discutons dans cette section de l'efficacité du repérage mesurée en termes de proportion de documents repérés à l'aide de chacun des cinq schémas de classification testés. Ce faisant, nous voulions identifier les caractéristiques d'un schéma de classification qui permettent d'augmenter la probabilité de repérer avec succès un document ciblé quels que soient le temps et le nombre d'essais requis.

La Figure 5.3 rappelle les résultats obtenus relatifs à la proportion moyenne de documents repérés à l'aide des schémas de classification.

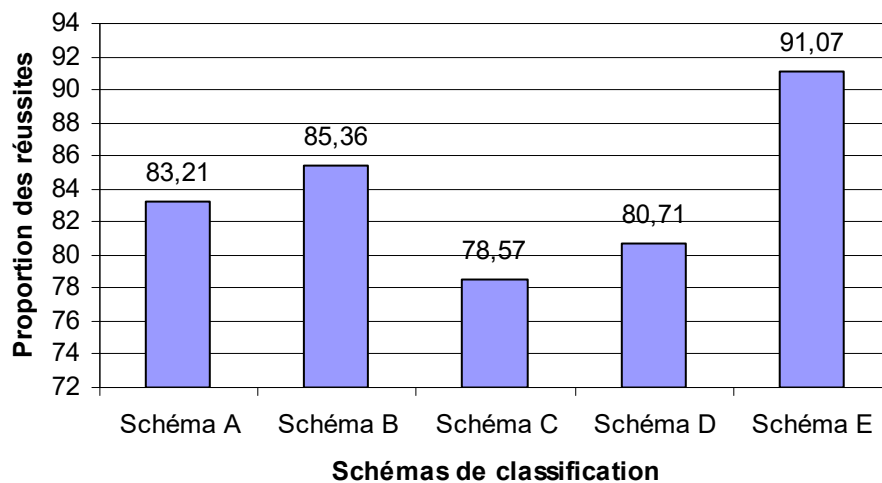


Figure 5.3 Proportion des documents repérés à l'aide des schémas de classification.

Nos résultats nous ont permis de confirmer partiellement l'hypothèse (HS1) qu'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations (Schéma E) permet de repérer plus de documents, à la différence d'un schéma de classification personnel (Schéma A), d'un schéma de classification à un seul niveau hiérarchique (Schéma C) et d'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique est celle de l'employé A (Schéma D). Il n'a pas été possible de

confirmer cette hypothèse quand les classes d'un schéma de classification personnel sont définies (Schéma B). Autrement dit, les différences observées dans la proportion moyenne de documents repérés à l'aide des schémas de classification B et E ne sont pas statistiquement significatives. Nous discutons ci-après de ces résultats.

Un schéma de classification hiérarchique est définitionnel par nature (Svenonius, 1997). Par exemple, la classe « Évaluation » peut avoir plusieurs sens dans un contexte académique et désigner différentes activités : « Évaluation des professeurs », « Évaluation des programmes », « Évaluation des superviseurs de stage », etc. Incorporée sous la classe « Étudiants », la signification de la classe « Évaluation » se précise tout en demeurant ambiguë. Par exemple, la classe « Évaluation » désigne-t-elle l'« Évaluation des étudiants » ou l'« Évaluation par les étudiants »? Sans pouvoir supprimer toute ambiguïté, la position d'une classe dans une structure hiérarchique est implicitement définitoire car elle fournit le contexte nécessaire à la désambiguïsation du terme utilisé pour nommer cette classe.

La conception du schéma B (i.e., schéma de classification personnel dont les classes sont définies) s'est inspirée du besoin de recréer artificiellement la situation où les répondants utilisent un schéma de classification personnel dont l'ambiguïté des classes serait minimisée non seulement par la position des classes dans la hiérarchie mais aussi par la possibilité d'interroger au besoin le propriétaire du schéma (i.e., l'employé A) sur la signification des classes. Pour simuler cette situation, les répondants ont consulté la définition de l'ensemble des classes avant de commencer à repérer les documents. Les répondants ont pu par la suite se référer à tout moment à ces définitions durant la simulation du repérage des documents à l'aide du schéma B. Ce faisant, nous voulions tester l'influence sur le repérage de documents d'une sémantique rendue explicite par la définition des classes d'un schéma de classification personnel.

Nos résultats confirment que l'utilisation d'un schéma de classification personnel qui comprendrait une définition des classes ou qu'une « visite guidée du schéma » avant le départ d'un employé permet d'augmenter sensiblement la proportion de documents

repérés par autrui. En l'absence de définition des classes (Schéma A), le même schéma ne permet qu'à 48.6 % des répondants de retrouver tous les documents (Voir Tableau 4.8 Fréquences de la proportion de documents repérés par schéma de classification, section 4.3.2 Tests d'hypothèses sur la proportion de documents repérés). Quand le même schéma de classification personnel est « expliqué », on peut s'attendre à ce qu'un individu autre que le propriétaire du schéma soit capable de repérer un document dans 52.9 % des cas même si la logique et la structure du schéma utilisé demeurent très personnelles.

Dans le cas de la présente expérimentation, le repérage des documents avait lieu immédiatement après la lecture des définitions de classes. On peut toutefois s'interroger sur la validité de ces résultats si on ajoute un délai de plusieurs heures ou de plusieurs jours entre la lecture des définitions et le repérage des documents. Est-on toujours capable de repérer les documents après un intervalle de temps prolongé? Il serait donc recommandé de consigner ces définitions pour pouvoir y référer à tout moment, comme c'est le cas dans les guides de classification en archivistique.

Si les définitions ont été généralement appréciées par les répondants car elles « aident », « sécurisent » et « rendent moins difficile le repérage », leur consultation sur un support papier a été jugée « déconcertante » par plusieurs répondants qui préféreraient une approche plus « intuitive » de repérage direct à l'écran, sans avoir à utiliser un « outil à côté [de l'écran d'ordinateur] ». La consultation des définitions sur un support papier en alternance avec la consultation à l'écran de la structure arborescente de répertoires correspondante « désoriente » en raison du changement de contexte (consultation papier vs. consultation à l'écran). Cette désorientation pourrait être évitée au moyen d'une interface technologique qui autorise l'affichage de la définition des classes directement à l'écran, par exemple au moyen d'« info bulles ».

Un schéma de classification hiérarchique a l'avantage de montrer explicitement les relations entre les différentes classes et d'offrir la possibilité de naviguer systématiquement de sujets généraux à sujets spécifiques, ce qui permet d'augmenter la

précision du repérage (Lin & Chan, 1999). Les résultats de notre recherche confirment l'avantage d'une structure hiérarchique, même personnelle, sur le plan du succès du repérage puisque le schéma de classification le moins performant en terme de proportion de documents repérés est le schéma C, constitué d'une longue liste alphabétique de classes. Nous rappelons que le schéma C a été artificiellement créé pour présenter des caractéristiques structurelles et sémantiques « extrêmes » (i.e., macro-structure très étendue, absence de profondeur, forte redondance structurelle et conceptuelle, intitulés exagérément longs) en vue d'évaluer l'efficacité du repérage réalisé à l'aide d'un schéma de classification dont la lecture est uniquement horizontale.

Les répondants ont manifesté durant la simulation des signes de surprise (par exemple, « Oh, mon Dieu! »), de découragement (par exemple, « C'est long à lire et à chercher ») et parfois de frustration (i.e., nombreux soupirs, beaucoup d'abandons) en utilisant le schéma C. Ce dernier contient beaucoup d'informations qualifiées par certains répondants de « bruit », la structure et les noms des classes ayant été jugés « trop longs » par plusieurs répondants. La piètre performance du schéma C sur le plan de la proportion de documents repérés confirme l'importance de la simplicité de l'apparence visuelle (i.e., structure physique) dans l'utilisabilité d'un instrument de recherche dans l'environnement électronique déjà relevée par Prom (2004), en vue d'éviter une surcharge cognitive et un découragement des utilisateurs.

L'accès à un document précis est réalisé à partir d'une lecture rapide des classes et mobilise des facultés sensorielles propres à la lecture ou au « balayage » d'un texte. La faible proportion de documents repérés à l'aide du schéma C mis en relation avec l'étendue de la structure devrait encourager les concepteurs de schémas à regrouper les classes en vue d'offrir une structure dont l'étendue (i.e., nombre de classes à chaque niveau hiérarchique) doit éviter de dépasser les dimensions de l'écran, ce qui correspond à une trentaine de classes. Le besoin de pouvoir visualiser simultanément l'ensemble de classes sans avoir à faire défiler l'écran (« pas pratique du tout ») a été exprimé verbalement par les répondants durant la simulation. Ce besoin a été également relevé dans la recherche menée par Gonçalves & Jorge (2003).

Les résultats confirment que le schéma E, dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations, augmente la probabilité de repérer avec succès les documents ciblés. Ce schéma a été rendu non seulement prévisible et cohérent par une organisation logique et hiérarchique de ses classes, mais aussi compréhensible en raison d'une sémantique rendue explicite par un usage de classes abrégées relativement modéré.

De manière surprenante, le schéma D, qui se distingue uniquement du schéma E sur le plan de la sémantique (le schéma D utilise la sémantique personnelle de l'employé A), mène à un repérage relativement faible au plan de la proportion de documents repérés comparativement aux autres schémas. La faible proportion des documents repérés à l'aide du schéma D mise en relation avec la forte proportion des documents repérés à l'aide des schémas E et B révèle l'influence d'une sémantique explicite (quelles que soient les caractéristiques structurelles et logiques) sur la probabilité de repérer avec succès des documents ciblés.

En résumé, nos résultats montrent l'influence de la dimension sémantique d'un schéma de classification sur la proportion de documents repérés. Les résultats démontrent et rappellent qu'un schéma de classification est non seulement un langage classificatoire (Maniez, 1987) mais aussi un langage définitoire dont les termes doivent être soigneusement choisis (selon la théorie sémantique, certains termes seraient plus significatifs que d'autres) ou définis (Dumais & Landauer, 1983; Hjørland, 2007). Une dénomination explicite des classes (par exemple, par l'utilisation de définitions ou en évitant l'usage des acronymes et des abréviations personnelles) améliore l'intelligibilité d'un schéma de classification et augmente la probabilité de repérer avec succès des documents ciblés quels que soient le temps et le nombre d'essais requis.

5.3.2 Schémas de classification et temps moyen du repérage

Nous discutons dans cette section de l'efficacité du repérage mesurée en termes de temps moyen requis pour repérer les documents à l'aide de chacun des cinq schémas de classification testés. Ce faisant, nous voulions identifier les caractéristiques d'un schéma de classification qui permettent d'augmenter la probabilité de repérer le plus rapidement les documents ciblés quels que soient la proportion de documents repérés et le nombre d'essais requis pour le repérage.

La Figure 5.4 rappelle les résultats obtenus relatifs au temps moyen requis (mesuré en secondes) pour repérer les documents ciblés à l'aide des schémas de classification.

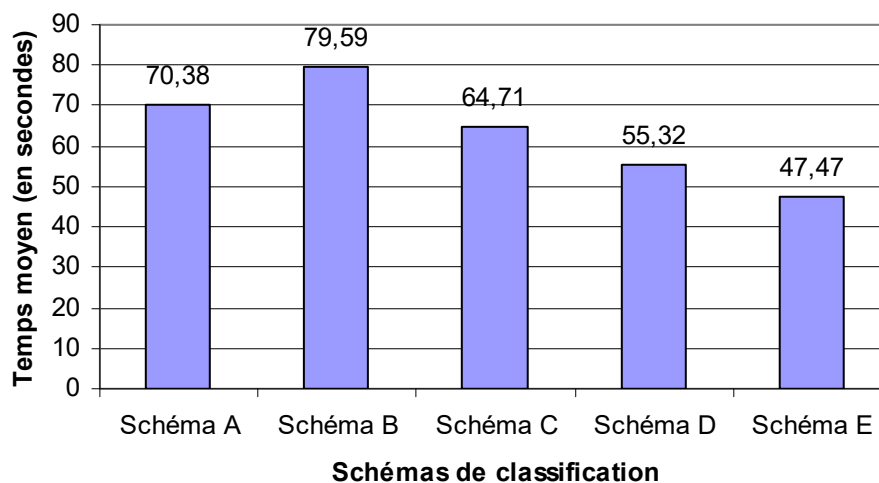


Figure 5.4 Temps moyen requis pour repérer les documents à l'aide des schémas de classification.

Les résultats nous ont permis de confirmer partiellement l'hypothèse (HS2) qu'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations (Schéma E) permet de repérer plus rapidement les documents, à la différence d'un schéma de classification personnel (Schéma A), d'un schéma de classification à un seul niveau (Schéma C) et d'un schéma de classification personnel

dont les classes ont été définies (Schéma B). Il n'a pas été possible de confirmer cette hypothèse quand le repérage des documents est réalisé à l'aide d'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique personnelle est celle de l'employé A (Schéma D). Autrement dit, les différences observées dans le temps moyen requis pour repérer les documents à l'aide des schémas de classification D et E ne sont pas significatives. Nous discutons ci-après de ces résultats.

La classification a toujours été étroitement liée au besoin de « sauver du temps ». L'histoire de la classification montre que lorsqu'un schéma de classification n'est plus utilisé, c'est plutôt en raison d'un mauvais rendement qu'en raison de non conformité théorique (i.e., application des meilleurs principes) (Maltby & Marcella, 2000). Nos résultats révèlent l'importance de privilégier une macro-structure peu étendue (i.e., nombre de classes principales peu élevé) basée, par exemple, sur un regroupement par classes d'activités pour accélérer le repérage des documents (Schémas E et D). Les résultats ne sont guère surprenants car l'importance de reconnaître la logique des langages classificatoires qui permet une exploration méthodique et un repérage plus rapide a déjà été soulignée par Hunter (2000). Ce qui est plus surprenant, c'est que ces caractéristiques structurelles (i.e., macro-structure peu étendue) et logiques (i.e., regroupement par classes d'activités) pourraient accélérer non seulement le processus de lecture (« à chaque niveau, pas plus qu'une dizaine de choix, on parcourt assez vite ») et de discrimination des classes, mais offrirait également l'avantage de permettre aux répondants de procéder par élimination lorsque la sémantique personnelle est peu explicite. Autrement dit, les hésitations entre le choix des classes et sous-classes pourraient être à la fois plus rares et moins longues (« il est rare d'hésiter entre plus que deux classes », « tout de suite plus facile d'éliminer certaines classes ») même si la signification de certaines classes peut demeurer ambiguë (par exemple, « S.S. »).

Plusieurs améliorations des caractéristiques structurelles et logiques peuvent cependant être encore apportées aux schémas D et E à partir des commentaires des répondants. Par exemple, plusieurs répondants ont qualifié l'ordre numérique de

succession des classes qui caractérise les schémas D et E de « mélangeant », « surtout quand le titre de la classe comporte aussi des données numériques ». D'autres répondants considèrent insuffisant le nombre de huit classes principales (i.e., macro-structure trop « étroite ») car il oblige à « descendre dans la hiérarchie » : « c'est agaçant, on n'en finit plus de cliquer » (i.e., structure jugée trop profonde). Enfin, l'incohérence de la logique caractérisée par la présence de « répertoires vides » a incommodé plusieurs répondants (« Dans stages, à chaque année, il n'y a pas la même arborescence. J'aimerais mieux trouver une arborescence partout plutôt que les répertoires soient vides. Comme cela on n'a pas à faire un apprentissage d'une année à l'autre »). Ces commentaires intéressants mériteraient d'être validés dans des recherches ultérieures en rapport avec les caractéristiques structurelles, logiques et sémantiques observées dans les schémas de classification personnels et les pratiques classificatoires institutionnelles. Les commentaires des répondants suscitent d'emblée des questionnements et des recommandations. On peut ainsi s'interroger sur la pertinence de préconiser un ordre numérique de succession des classes dans l'environnement électronique quand la majorité des employés interrogés et les répondants amenés à utiliser ces schémas de classification personnels manifestent ou expriment eux-mêmes une préférence pour un ordre alphabétique de succession des classes. La cote de classification pourrait être jointe à la fin du nom de la classe (par exemple, « Stages (1100) »), et non au début, comme c'est déjà pratiqué dans certains organismes. De même, les employés interrogés, tout comme les répondants amenés à utiliser les schémas de classification personnels, préfèrent utiliser des schémas de classification dont la macro-structure est étendue et la structure peu profonde, à l'inverse des caractéristiques structurelles des schémas D et E testés, plus proches des caractéristiques des schémas de classification institutionnels, ce qui pourrait expliquer la réticence des employés à utiliser ces derniers. Enfin, la présence de « répertoires vides » dans certains schémas de classification personnels, en réponse à divers besoins déjà mentionnés par les employés (par exemple, se rappeler de l'existence d'un « dossier papier » sur le même sujet), ne favorise pas l'apprentissage du schéma par une tierce personne en raison d'un manque de cohérence et génère de la confusion chez

une tierce personne appelée à utiliser le schéma de classification pour le repérage de documents ciblés.

Les pires résultats au niveau du temps moyen requis pour repérer les documents sont obtenus avec le schéma B (schéma personnel de l'employé dont les classes sont définies) et le schéma A (schéma personnel de l'employé). En ce qui concerne le schéma B, ces résultats ne sont guère surprenants car ils incluent le temps de lecture des définitions des classes effectuée par les répondants à tout moment durant le repérage des documents électroniques. Le temps serait probablement moins long si les définitions étaient visibles directement à l'écran, comme nous l'avons déjà suggéré. Cependant, si le repérage des documents réalisé à l'aide du schéma B est généralement plus long, cet « investissement » en termes de temps moyen requis pour le repérage est compensé par la probabilité plus élevée de repérer avec succès les documents.

Le faible score observé sur le plan de la rapidité du repérage réalisé à l'aide du schéma personnel A peut s'expliquer par le nombre élevé de classes principales (i.e., 42 classes principales) combiné à des subdivisions sur un ou deux niveaux qui obligent une lecture à la fois horizontale et verticale des classes rallongeant le temps de lecture. De plus, mais ces présomptions n'ont pas été statistiquement vérifiées, il se peut que la proportion élevée de classes redondantes et la variété des critères de division privilégiés aux deux premiers niveaux hiérarchiques contribuent à placer le répondant dans un univers d'incertitude qui rallonge la durée de la prise de décision dans l'identification des classes pertinentes.

En résumé, les résultats démontrent l'importance des caractéristiques structurelles et logiques (quelles que soient les caractéristiques sémantiques) d'un schéma de classification pour la rapidité avec laquelle on repère un document. Un schéma de classification agit comme outil de communication dont la simplicité structurelle et logique favorise la prise de décision ainsi qu'une exploration plus rapide et méthodique des classes. Nos résultats révèlent que la conception d'un schéma de classification peut s'inspirer du principe de parcimonie (ou principe de simplicité) pour être compris

rapidement. Autrement dit, à performance comparable au plan de la rapidité du repérage, on devrait privilégier le modèle le plus simple. Cette simplicité peut se manifester par un schéma dont le nombre de classes principales est peu élevé et la logique est basée partiellement sur un seul critère de division.

5.3.3 Schémas de classification et proportion de documents repérés dès le premier essai

Nous discutons dans cette section de l'efficacité du repérage mesurée en termes du nombre d'essais nécessaires pour repérer des documents ciblés à l'aide de chacun des cinq schémas de classification testés. Ce faisant, nous voulions identifier les caractéristiques d'un schéma de classification qui minimisent le risque d'erreur au repérage quels que soient la proportion de documents repérés et le temps moyen requis pour repérer les documents.

La Figure 5.5 rappelle les résultats obtenus relativement à la proportion de documents repérés dès le premier essai à l'aide des schémas de classification.

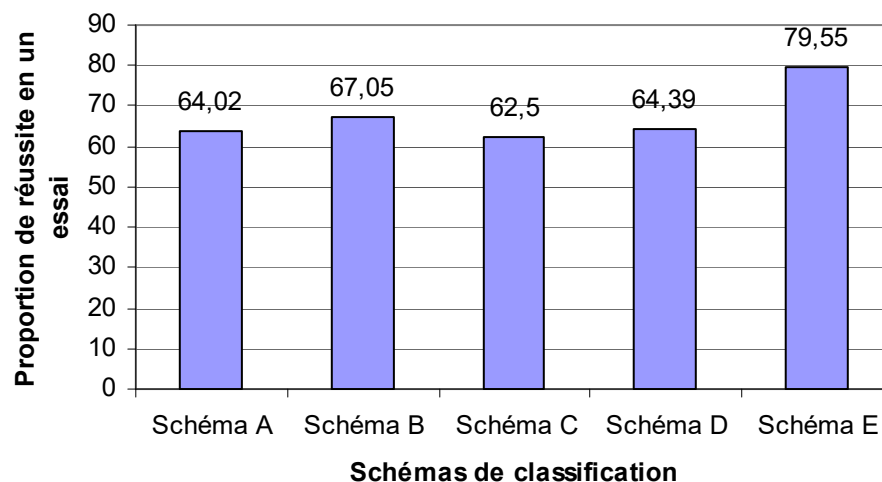


Figure 5.5 Proportion de documents repérés dès le premier essai à l'aide des schémas de classification.

Les résultats nous ont permis de confirmer l'hypothèse (HS3) qu'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations (Schéma E) augmente la proportion de documents repérés dès le premier essai, à la différence d'un schéma de classification personnel (Schéma A), d'un schéma de classification personnel dont les classes ont été définies (Schéma B), d'un schéma de classification à un seul niveau (Schéma C) et d'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique personnelle est celle de l'employé A (Schéma D).

D'un point de vue théorique, l'utilisation d'un seul critère de division (ou d'un critère de division dominant) à chaque niveau de la hiérarchie augmente la cohérence et permet une recherche méthodique à l'intérieur de la structure (Maniez, 1987; Fallis & Mathiesen, 2000; Hunter, 2000). Nous avons déjà rappelé l'importance du premier critère de division car c'est lui qui détermine la macro-structure et décrit le domaine à classer (Kwasnik, 1999). Les schémas D et E privilégient tous deux un regroupement par classes d'activités, tel que préconisé par les archivistes pour l'organisation des documents administratifs (Couture & Rousseau, 1982; Roberge, 1985; Héon, 1995) ainsi qu'une macro-structure peu étendue. Or, seul le schéma E, qui se distingue du schéma D au plan sémantique, permet de minimiser le risque d'erreur au repérage.

L'approche utilisée dans notre recherche ne permet pas de connaître avec précision les raisons liées à l'incapacité à repérer un document dès le premier essai. Cela dit, les causes de l'incapacité à repérer les documents ont été mentionnées dans des recherches sur l'évaluation de différents systèmes de repérage (Lancaster, 1971; Lytle, 1980). Par exemple, les résultats de la recherche menée par Lytle (1980) indiquent quatre raisons pouvant être à l'origine des échecs sur le plan du repérage. Il s'agit : (1) d'une mauvaise utilisation du système de repérage (par exemple, sélectionner la mauvaise classe principale dans la hiérarchie), (2) de mauvaises inférences (par exemple, échec dans la localisation des documents), (3) des erreurs de lecture (par exemple, les classes pertinentes sont vues mais non sélectionnées), et (4) d'un système de repérage de

mauvaise qualité (par exemple, des noms de classes non significatifs). Les raisons évoquées par Lytle (1980) pour expliquer l'échec du repérage permettent *a contrario* d'expliquer le succès du repérage dès le premier essai.

Conformément à nos attentes, le schéma E dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations obtient le score le plus élevé sur le plan de la proportion de documents repérés dès le premier essai. Un nombre limité de classes principales, par exemple, minimiserait le risque d'échec en favorisant la sélection et la localisation des classes pertinentes. Les caractéristiques du schéma E encourageraient aussi les inférences adéquates (par exemple, en proposant une division logique des classes) et contribueraient à éviter les erreurs de lecture (par exemple, en évitant des noms de classes abrégées). La confirmation de notre troisième hypothèse (i.e., seul le schéma E dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus de documents dès le premier essai) suggère que les caractéristiques structurelles, logiques et sémantiques du schéma E agissent en synergie pour minimiser le risque d'erreur lors du repérage.

En rapport avec les résultats relatifs à la proportion de documents repérés, il ne nous apparaît pas étonnant que le score le plus faible soit obtenu avec le schéma C, caractérisé par l'absence de structure arborescente, l'utilisation d'énoncés extrêmement longs et une plus forte proportion de classes abrégées et redondantes. L'ensemble de ces caractéristiques peut générer de la confusion (« on s'y perd ») et parfois de l'impatience (« on s'arrête aux premiers [répertoires] qui semblent pertinents »), sentiments qui peuvent encourager une réponse souvent erronée ou hâtive d'où un nombre d'essais plus élevé. L'importance d'une apparence visuelle parcimonieuse d'un schéma de classification, déjà soulignée, est également valable dans ce contexte-ci.

Les résultats relatifs à la proportion de documents repérés dès le premier essai obtenue à l'aide du schéma D (i.e., le schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique personnelle est celle de l'employé A) sont surprenants puisqu'ils sont sensiblement identiques aux résultats obtenus à l'aide du schéma A (i.e., le schéma de classification personnel de l'employé), soit une proportion de documents repérés dès le premier essai d'environ 64 %. À l'inverse, les résultats relatifs à la proportion de documents repérés dès le premier essai obtenue à l'aide du schéma E (i.e., le schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations) sont différents des résultats obtenus à l'aide du schéma D. Autrement dit, la dimension sémantique qui distingue le schéma E du schéma D peut influencer la probabilité de retrouver une plus grande proportion de documents dès le premier essai. La forte proportion de documents repérés dès le premier essai avec le schéma B, dont la sémantique personnelle est rendue explicite par une définition des classes, va dans le sens de cette présomption. Cela dit, les différences statistiquement significatives observées entre le schéma E et le schéma B suggèrent bien que c'est plutôt une action coordonnée de l'ensemble des caractéristiques structurelles, logiques et sémantiques qui minimise le risque d'erreur au repérage plutôt que l'influence particulière de l'une ou l'autre de ces caractéristiques.

En résumé, nos résultats montrent l'influence des caractéristiques structurelles, logiques et sémantiques d'un schéma de classification sur la proportion des documents repérés dès le premier essai. Un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations minimise le risque d'erreur dans le repérage. Ces caractéristiques agissent en synergie, c'est-à-dire qu'elles se renforcent les unes les autres, contribuant ainsi à clarifier le sens à donner à l'organisation logique et hiérarchique des classes et à minimiser le risque d'erreur dans le repérage.

5.3.4 Conclusion

Notre recherche a démontré que les schémas de classification ne sont pas tous égaux sur le plan de l'efficacité du repérage. En corrélant les principes classificatoires avec les résultats obtenus, notre recherche a permis d'observer l'influence des caractéristiques structurelles, logiques et sémantiques d'un schéma de classification sur l'efficacité du repérage en termes de proportion de documents repérés, de temps moyen requis pour le repérage et de la proportion de documents repérés dès le premier essai.

Les résultats révèlent qu'il est possible de repérer des documents à l'aide d'un schéma de classification personnel même si l'efficacité du repérage n'est pas optimale. En vue d'améliorer le succès du repérage (i.e., augmenter la proportion de documents repérés), l'optimisation de la logique et de la structure seulement ne suffit pas; les résultats révèlent l'importance de la nature définitoire d'un schéma de classification et d'une sémantique personnelle ou institutionnelle rendue explicite par l'usage de termes conventionnels, contrôlés ou définis. Une analyse lexicale des schémas de classification personnels pourrait être envisagée en vue de concevoir un référentiel commun qui contribuerait à rendre un schéma de classification intelligible pour une majorité d'utilisateurs.

En ce qui concerne la rapidité du repérage, les résultats confirment que si une sémantique personnelle rendue explicite par une définition des classes améliore le succès du repérage, la lecture de ces définitions a pour corollaire d'augmenter la durée du repérage ce qui peut être particulièrement contraignant dans un contexte de productivité au travail. Nos résultats révèlent que si un regroupement par activité et une macro-structure simplifiée ne permettent pas nécessairement de repérer les documents lorsque la sémantique n'est pas explicite, les documents repérés le sont toutefois avec plus de rapidité.

Enfin, si la qualité d'un schéma de classification devait être mesurée par la proportion de documents repérés dès le premier essai, qualité qui peut être particulièrement valorisée en milieu de travail, une attention toute particulière devrait

être portée à l'ensemble des caractéristiques structurelles, logiques et sémantiques d'un schéma de classification de manière à en contrôler l'action coordonnée sur le repérage. La simplicité d'utilisation et la clarté du schéma de classification doivent être privilégiées. De telles propriétés ont été retrouvées dans notre recherche dans un schéma de classification présentant un faible nombre de classes principales, un regroupement par classes d'activités et un nombre peu élevé de classes abrégées.

L'observation de différences sur le plan de l'efficacité du repérage de documents ciblés réalisé à l'aide de schémas de classification présentant des caractéristiques variées permet d'identifier des méthodes et de confirmer des principes classificatoires visant à favoriser le succès du repérage, à en augmenter la rapidité et à diminuer le risque d'erreur du repérage de documents administratifs électroniques réalisé par une tierce personne.

Il est souhaitable d'inclure dans un schéma de classification personnel une définition de ses classes qui soit directement visible à l'écran, plutôt que sur papier, afin de réduire le temps moyen du repérage et d'augmenter la probabilité de repérer avec succès un document administratif électronique ciblé. Il est conseillé d'éviter d'utiliser des classes abrégées ou d'employer, le cas échéant, des abréviations basées sur une liste de désignations officielles ou consensuelles au sein de l'unité, afin d'augmenter la probabilité de repérer avec succès un document administratif électronique ciblé.

Il est recommandé de regrouper les classes principales sur la base d'un seul critère ou d'un critère dominant (par exemple, par activité) et d'en réduire le nombre pour des raisons de cohérence et d'ergonomie d'écran, afin d'augmenter la probabilité de repérer plus rapidement un document administratif électronique ciblé.

Il est pertinent de concevoir un schéma de classification qui appliquerait les recommandations précédentes (i.e., sémantique explicite, nombre de classes peu élevé, logique basée partiellement sur une division par classes d'activités), afin d'augmenter la probabilité de repérer avec succès un document administratif électronique ciblé dès le premier essai.

Toute recherche comporte des limites inhérentes, à la fois conceptuelles et méthodologiques, qui déterminent le degré de généralisation des résultats. La section suivante présente les limites de notre recherche.

5.4 Limites de la recherche

Cette section présente les précautions qui doivent être prises avant de généraliser les observations relatives aux caractéristiques des schémas de classification personnels et aux résultats de la simulation du repérage.

Une première limite de notre recherche est liée au mode d'échantillonnage des schémas de classification personnels. L'échantillonnage visait à privilégier la sélection d'une vingtaine de schémas de classification conçus par des individus en fonction de traits caractéristiques particuliers (i.e., même rôle professionnel et utilisation d'une structure arborescente de répertoires pour l'organisation des documents électroniques). La taille de l'échantillon des schémas de classification personnels étudiés ($n = 21$) et sa nature non probabiliste (i.e., échantillon à choix raisonné et participation volontaire des employés) ne permettent pas de généraliser les observations relatives aux caractéristiques des schémas de classification personnels à l'ensemble des schémas de classification utilisés par les techniciennes à la coordination de bureau et les coordonnateurs de stages de l'Université de Montréal. Comme tous les employés n'ont pas eu une chance égale de faire partie de l'échantillon, la représentativité des schémas de classification personnels n'est pas assurée.

Une deuxième limite de notre recherche est liée au mode de sélection des répondants ayant participé à l'expérimentation. Les critères d'inclusion de ces répondants ont été conçus en vue de contrôler les variables étrangères et de constituer des groupes équivalents. Nous avons délibérément choisi des étudiants et des diplômés provenant majoritairement de l'École de bibliothéconomie et des sciences de l'information de l'Université de Montréal. En raison de leur connaissance plus ou moins développée des principes et méthodes classificatoires reliées au domaine des sciences de

l'information, les commentaires formulés par les répondants lors de la simulation et qui ont été cités dans cette thèse doivent être examinés avec précaution. Ce biais possible n'affecte cependant pas les résultats de l'expérimentation puisque l'unité principale d'analyse n'était pas la perception et la satisfaction des répondants mais plutôt l'efficacité du repérage. Le nombre élevé de répondants, leur répartition aléatoire à travers les groupes et les tests statistiques réalisés assurent la validité interne et la fiabilité des résultats. Autrement dit, on obtiendrait des résultats similaires dans le cadre d'une autre expérimentation réalisée dans les mêmes conditions.

Enfin, la simulation du repérage de documents a été réalisée avec cinq schémas de classification possédant des caractéristiques structurelles, logiques et sémantiques particulières. Les conclusions de cette recherche de nature exploratoire sont valides uniquement pour les schémas de classification qui ont été inclus dans cette recherche. Il serait toutefois pertinent de comparer, dans des conditions expérimentales identiques, l'efficacité du repérage obtenu à l'aide d'un échantillon de schémas de classification possédant des caractéristiques structurelles, logiques et sémantiques différentes de celles déjà étudiées dans notre recherche.

5.5 Conclusion

Ce chapitre a fourni des éléments de discussion des résultats en reliant ces derniers à la littérature scientifique et aux connaissances actuelles.

Livrés à eux-mêmes, les employés ont des comportements classificatoires que l'on peut qualifier à la fois de créatifs et de réguliers. La créativité se manifeste par l'organisation des documents sur la base de critères de division variés et par une dénomination des classes très personnelle. La régularité se manifeste à travers plusieurs traits structurels, logiques et sémantiques observables dans plusieurs schémas de classification personnels. Par exemple, les employés conçoivent et utilisent des structures classificatoires beaucoup moins profondes que celles des schémas de classification institutionnels, ils sont davantage enclins à utiliser un ordonnancement alphabétique

plutôt que numérique des classes et s'accommodent de classes redondantes aux noms abrégés. L'identification de ces caractéristiques communes suggère un certain décalage entre les schémas de classification personnels et les schémas de classifications institutionnels et permet de mieux comprendre la réticence de certains employés à vouloir utiliser ces derniers pour l'organisation des documents administratifs électroniques qui sont sous leur contrôle direct.

La conception et le processus continu de mise à jour des schémas de classification personnels analysés apparaissent davantage contextuels que basés sur des principes classificatoires théoriques. En effet, nos résultats suggèrent l'influence non négligeable de contraintes ou de facteurs documentaires (par exemple, quantité de documents à organiser), professionnels (par exemple, formation reçue en gestion des documents), pragmatiques (par exemple, besoin d'accéder plus rapidement à certains documents) et cognitifs (par exemple, repérage basé sur la mémoire) sur les caractéristiques structurelles, logiques et sémantiques des schémas de classification personnels observés. Par conséquent, les observations relatives aux caractéristiques structurelles, logiques et sémantiques des schémas de classification personnels observés sont valides pour une période donnée et dans un contexte professionnel, documentaire, pragmatique et cognitif particulier.

L'analyse des résultats de la simulation du repérage de documents confirme une seule de nos hypothèses de recherche : l'utilisation d'un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus de documents dès le premier essai. Les résultats de la simulation du repérage rappellent l'importance de la nature définitoire d'un schéma de classification et confirment que les règles personnelles de conception des schémas de classification ne sont pas nécessairement comprises par d'autres. Nous avons cherché à identifier plus précisément les caractéristiques structurelles, logiques et sémantiques qui facilitent la compréhension d'un schéma de classification conçu par autrui et, par conséquent, améliorent l'efficacité du repérage de documents administratifs électroniques réalisé à l'aide d'un tel schéma.

L'usage d'une sémantique explicite, d'une structure parcimonieuse sur le plan de l'étendue et d'un regroupement cohérent des classes augmente la clarté et la simplicité d'utilisation d'un schéma de classification personnel par une tierce personne.

La nature exploratoire de notre recherche oblige cependant à prendre des précautions quant à la généralisation statistique de ces résultats à d'autres contextes en raison de limites liées au mode d'échantillonnage des schémas de classification personnels, au mode de sélection des répondants et à la taille de l'échantillon des schémas de classification testés.

Les retombées de cette thèse et les pistes de recherche suggérées par nos résultats sont présentées dans le chapitre suivant.

6 Conclusion

La conclusion résume notre recherche, présente la contribution de notre thèse aux sciences de l'information, plus particulièrement à l'archivistique, et recommande des pistes de recherches futures.

6.1 Résumé de la recherche

Notre recherche est venue combler un manque de connaissances sur les schémas de classification et le repérage des documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles. Nous avons pu atteindre le but que nous nous étions fixé, soit d'étudier l'efficacité du repérage de documents électroniques à l'aide de schémas de classification ayant des caractéristiques variées. Nous avons réalisé les deux objectifs spécifiques suivants : (1) décrire les caractéristiques de différents schémas de classification utilisés par des employés pour organiser et classer les documents administratifs électroniques qui sont sous leur contrôle direct et (2) vérifier les différences sur le plan de l'efficacité du repérage de documents électroniques dans un environnement contrôlé qui sont fonction du schéma de classification utilisé. La problématique a soulevé une question simple mais qui demeurait, à ce jour, sans réponse : « Peut-on repérer avec la même efficacité un document électronique à l'aide de schémas de classification ayant des caractéristiques variées? ». Il a été aussi possible de répondre aux questions spécifiques suivantes : (1) Quelles sont les caractéristiques des schémas de classification utilisés par les employés pour organiser et classer les documents administratifs électroniques qui sont sous leur contrôle direct? (2) Existe-t-il des différences sur le plan de la capacité à repérer tous les documents ciblés à l'aide de schémas de classification ayant des caractéristiques variées? (3) Existe-t-il des différences sur le plan de la capacité à repérer rapidement les documents ciblés à l'aide de schémas de classification ayant des caractéristiques variées? (4) Existe-t-il des différences sur le plan de la capacité à repérer

les documents ciblés dès le premier essai à l'aide de schémas de classification ayant des caractéristiques variées?

La littérature ayant servi de base à cette recherche provient de différents domaines de connaissances, soit les sciences de l'information, l'archivistique et la gestion personnelle de l'information. L'analyse des écrits pertinents à notre sujet a permis de présenter les assises théoriques de la classification documentaire, les principes de classification des documents administratifs en archivistique, le contexte général relatif à l'organisation personnelle des documents électroniques et le processus de repérage des documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles. Ce faisant, la recension des écrits a servi de cadre de référence pour (1) développer un modèle d'analyse des caractéristiques multidimensionnelles des schémas de classification, (2) justifier les procédures d'échantillonnage en tenant compte des différents facteurs d'influence possible, (3) justifier les manipulations réalisées sur les schémas de classification, (4) identifier les indicateurs de mesure de l'efficacité du repérage de documents ciblés, (5) concevoir un protocole de simulation qui contrôle les variables étrangères et (6) interpréter les résultats de notre recherche.

L'approche méthodologique retenue pour notre recherche est de type quantitatif exploratoire. Un protocole de recherche en trois étapes fut réalisé pour atteindre les deux objectifs que nous nous étions fixés. La première étape du projet impliquait l'identification et la description des caractéristiques de 21 schémas de classification personnels utilisés par des employés pour organiser les documents administratifs électroniques qui sont sous leur contrôle direct. Cette caractérisation de schémas de classification personnels, effectuée à l'aide d'un modèle d'analyse, nous a permis d'approfondir notre connaissance des schémas de classification personnels aux plans structurel, logique et sémantique. Dans la deuxième étape de notre projet, nous avons procédé à plusieurs manipulations des caractéristiques structurelles, logiques et sémantiques d'un schéma de classification personnel. L'objectif de cette opération était de constituer un corpus de cinq schémas de classification présentant des caractéristiques

variées au plan structurel, logique et sémantique en vue de tester les hypothèses spécifiques suivantes : (1) un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus de documents, (2) un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus rapidement les documents, et (3) un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations permet de repérer plus de documents dès le premier essai. Dans la troisième étape de notre projet, une simulation a été réalisée avec 70 répondants pour collecter des données sur l'efficacité du repérage d'un ensemble de 20 documents électroniques réalisé à l'aide de schémas de classification ayant des caractéristiques variées. Trois ensembles de données ont été utilisés pour une analyse de variance qui nous a permis d'examiner de quelle façon les caractéristiques des différents schémas de classification des documents électroniques influençaient la performance quant à l'efficacité du repérage des documents électroniques : la proportion de documents repérés, le temps moyen (mesuré en secondes) requis pour repérer les documents cibles et la proportion de documents cibles repérés dès le premier essai.

À toutes les étapes de notre recherche, des mesures ont été prises pour assurer la qualité de la recherche en termes de validité et de fiabilité. Pour assurer la validité de la recherche, nous avons tenu compte des critères de validité du construit, de validité interne et de validité externe. Nous avons pris des moyens pour assurer la fidélité des instruments de mesure et la fiabilité des résultats.

Les résultats obtenus sont de deux types : (1) les résultats descriptifs relatifs à la caractérisation des schémas de classification personnels et (2) les résultats des tests d'hypothèses relatifs à l'efficacité du repérage de documents électroniques dans un environnement contrôlé qui est fonction du schéma de classification utilisé.

L'analyse des résultats descriptifs nous a permis de constater que les schémas de classification conçus et utilisés par les employés sont la manifestation de comportements classificatoires à la fois créatifs et réguliers. Les schémas de classification personnels des employés se caractérisent généralement par une macro-structure étendue, une structure peu profonde, complexe et déséquilibrée. Les employés privilégient un ordonnancement alphabétique des classes plutôt que numérique et divisent leurs classes sur la base de différents critères de division, plutôt que de privilégier une division par classes d'activités. Les schémas de classification personnels observés se distinguent des schémas de classification institutionnels par une redondance structurelle et conceptuelle à l'intérieur des noms de classes généralement élevée dès le premier niveau de la hiérarchie.

L'analyse des résultats de la simulation du repérage de documents a permis de confirmer que les schémas de classification ne sont pas tous égaux sur le plan de l'efficacité du repérage d'un ensemble de documents ciblés. Les résultats de la simulation confirment cependant une seule de nos hypothèses de recherche. Les résultats ne nous ont pas permis d'identifier un modèle de schéma de classification personnel qui permettrait à la fois d'augmenter de manière significative la probabilité de repérer avec succès un document électronique, tout en minimisant le temps requis pour compléter la tâche ainsi que le risque d'erreur au repérage. Cependant, nous avons pu observer qu'une sémantique explicite (par une définition des classes ou une faible proportion de noms de classes abrégés) s'est révélée particulièrement influente sur le succès du repérage, quelles que soient les caractéristiques structurelles et logiques d'un schéma de classification. Une macro-structure peu étendue (i.e., moins d'une dizaine de classes principales) et une logique claire et cohérente basée partiellement sur une division par classes d'activités apparaissent suffisantes pour accélérer le repérage de documents ciblés. Enfin, un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations minimise de manière significative le risque d'erreur au repérage.

La discussion des résultats a permis de relever plusieurs points saillants. En ce qui concerne les caractéristiques des schémas de classification personnels, nous avons constaté que l'identification de traits communs suggère un certain décalage entre les caractéristiques des schémas de classification personnels et les caractéristiques des schémas de classifications institutionnels. Ce décalage permet de mieux comprendre la réticence de certains employés à vouloir utiliser les schémas de classification institutionnels pour l'organisation des documents administratifs électroniques qui sont sous leur contrôle direct. Les caractéristiques structurelles, logiques et sémantiques singulières observées dans les schémas de classification personnels des employés pourraient s'expliquer par une conception et une mise à jour davantage contextuelles (influence de facteurs documentaires, professionnels, pragmatiques et cognitifs) plutôt que basées sur une connaissance et une application des principes classificatoires théoriques. Malgré le non respect des principes classificatoires, il s'est avéré possible pour une tierce personne de repérer des documents, dans un environnement contrôlé à l'aide d'un schéma de classification personnel, même si l'efficacité du repérage n'est pas optimale. Il a été possible de cibler avec précision quelles caractéristiques (ou combinaison de caractéristiques) influencent l'efficacité du repérage sur le plan de la proportion de documents repérés, du temps moyen (mesuré en secondes) requis pour le repérage et de la proportion de documents repérés dès le premier essai. Enfin, les limites de la recherche liées au mode d'échantillonnage des schémas de classification personnels, au mode de sélection des répondants et à la taille de l'échantillon des schémas de classification testés ont été clairement identifiées.

Les retombées théoriques, méthodologiques et pratiques de cette recherche sont multiples. La section suivante présente les contributions de cette thèse dans le domaine des sciences de l'information.

6.2 Contributions dans le domaine des sciences de l'information

En ce qui concerne les retombées théoriques, notre recherche a permis d'enrichir la connaissance de types de schémas de classification bien particuliers, ceux créés par des employés dans un contexte de travail de bureau. Les schémas de classification conçus par les employés peuvent être qualifiés de « naïfs », pour reprendre l'expression de Beghtol (2003), par opposition aux schémas de classification « professionnels », conçus par des professionnels de la gestion des documents. L'intérêt théorique de décrire des schémas de classification « naïfs » a été de permettre l'étude des caractéristiques de schémas de classification dont la conception ne se base pas *a priori* sur la connaissance de la théorie de la classification. Il a été toutefois possible d'observer l'application intuitive par les employés du principe de garantie d'existence bibliographique ou « *literary warrant* ». Les résultats suggèrent des interrogations quant à la pertinence du principe de division logique qui n'est pas respecté dans l'organisation des documents administratifs électroniques qui sont sous le contrôle direct de l'employé. Par ailleurs, les résultats de la simulation du repérage révèlent que le non respect du principe de division logique ne permet pas d'atteindre une efficacité optimale du repérage de documents ciblés, quand ce repérage est réalisé par une tierce personne.

En ce qui concerne les retombées méthodologiques, notre recherche a permis la conception d'un modèle d'analyse de schémas de classification original qui prend en compte non seulement les caractéristiques structurelles des schémas de classification mais aussi leurs caractéristiques logiques et sémantiques. Notre modèle a permis une analyse quantitative et qualitative des schémas de classification. Cet outil d'analyse, conçu et validé pour l'étude des schémas de classification personnels, peut être développé et réutilisé dans toute discipline reliée aux sciences de l'information dont l'archivistique ou dans toute recherche qui s'intéresse aux schémas de classification. Cet outil d'analyse a d'ailleurs déjà été employé dans le cadre d'une recherche visant à caractériser et comparer différents répertoires de ressources du Web spécialisées dans le domaine de l'éducation (Hudon, Mas & Gazo, 2005).

Notre protocole de recherche est une contribution au développement, jugé encore embryonnaire (Jones, 2007), de méthodologies spécialement adaptées aux études relatives à la gestion personnelle de l'information. L'originalité de notre méthodologie a été d'évaluer des schémas de classification au moyen du repérage de documents administratifs électroniques. Il existe très peu de recherches expérimentales sur les instruments de repérage en archivistique (Lytle, 1980; Ribeiro, 1996). Aucune, à notre connaissance, ne s'est intéressée à l'étude des schémas de classification de documents administratifs électroniques selon une approche quantitative. Dans notre recherche, à la fois descriptive (décrire les caractéristiques de schémas de classification personnels) et exploratoire (comparer l'efficacité du repérage réalisé à l'aide de différents schémas de classification), nous avons conçu une méthodologie qui combine plusieurs méthodes de collecte et d'analyse de données complémentaires : utilisation d'un logiciel d'analyse de l'espace documentaire, saisies d'écran, analyse de contenu, analyse de variance; cela nous a permis d'enrichir les conclusions de notre étude. Notre recherche a permis de confirmer la faisabilité et la validité d'une approche essentiellement quantitative de nature expérimentale pour l'étude de l'influence des schémas de classification sur l'efficacité du repérage de documents électroniques.

En ce qui concerne les retombées pratiques, notre recherche présente un intérêt certain pour les archivistes qui pourraient s'appuyer sur ses résultats pour bâtir des schémas de classification destinés à l'organisation, au classement et au repérage des documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles. Notre recherche a permis d'examiner des schémas de classification personnels utilisés au sein d'un organisme et d'observer quel a été le schéma de classification le plus performant sur le plan de l'efficacité du repérage dans les conditions de la simulation. En ciblant les schémas de classification qui augmentent la probabilité de repérer efficacement les documents électroniques (en termes de succès, de temps et de nombre d'essais requis pour le repérage), notre recherche présente également un intérêt pour les employés soucieux de concevoir et d'utiliser un schéma de

classification permettant un repérage et une réutilisation possible de leurs documents électroniques par leurs collègues de travail.

6.3 Recommandations pour des recherches futures

Nos résultats et nos observations suggèrent plusieurs pistes de recherche. Nous présentons dans cette section six recommandations pour des recherches futures reliées à l'étude des schémas de classification dans un contexte de gestion décentralisée des ressources informationnelles pistes de recherche.

6.3.1 Valider la relation entre la classification personnelle et la constitution du fonds d'archives

Comme première piste de recherche, nous suggérons de valider le lien existant entre la classification personnelle des documents administratifs électroniques et la constitution du fonds d'archives. La notion de fonds d'archives désigne l'ensemble des documents que tout individu ou organisme a « automatiquement et organiquement » produit ou reçu dans l'exercice de ses fonctions ou de ses activités (Conseil international des archives, 1997). Nous savons que la classification institutionnelle joue un rôle déterminant dans la constitution du fonds d'archives puisqu'elle permet le regroupement logique des documents à l'intérieur des dossiers, le regroupement logique des dossiers selon les activités organiques et la constitution du fonds par l'identification de ces activités dans le temps. Ce faisant, la classification n'a pas seulement pour objectif d'organiser et de faciliter le repérage de l'information que contiennent les documents, elle permet également de préserver la valeur de preuve de ces documents par la mise en valeur du contexte dans lequel ils sont créés. Or, notre recherche a montré que les employés ne classifient pas leurs documents seulement par activité mais utilisent indifféremment plusieurs critères de regroupement des documents. Les pratiques classificatoires personnelles représentent donc un risque de perte de connaissance du contexte de création et d'utilisation des documents, et par conséquent une impossibilité de témoigner fidèlement des activités de l'employé et de comprendre le rôle, la valeur et la signification des documents dans la réalisation de ces activités.

Notre recherche a montré qu'il a été possible à une tierce personne de repérer des documents à l'aide d'un schéma de classification personnel. Il nous apparaît pertinent de vérifier si une tierce personne est également capable d'identifier, à partir du schéma de classification personnel d'un employé, les activités auxquelles se rattachent les documents administratifs électroniques qui sont sous le contrôle direct de cet employé. Une telle recherche nous permettrait de vérifier s'il est possible de garder la trace des circonstances originales de la création et de l'utilisation des documents administratifs électroniques qui sont sous le contrôle direct des employés. Elle nous donnerait ainsi l'occasion de vérifier s'il est possible de faire le lien entre la classification personnelle et la classification institutionnelle qui établit les contours externes du fonds et définit ses divisions internes (Héon, 1999).

6.3.2 Tester les schémas de classification dans un contexte réel d'utilisation

Il serait intéressant de tester le schéma de classification artificiellement créé qui obtient le meilleur score au plan de l'efficacité du repérage dans un contexte réel d'utilisation. Garantir le repérage des documents est une des fonctions d'un schéma de classification personnel. C'est en raison de cette fonctionnalité que nous avons évalué l'efficacité du repérage de documents électroniques à partir de différents schémas de classification. Les résultats de notre recherche sont basés sur des tests réalisés dans un environnement contrôlé. Bien que cela n'enlève rien à la portée de notre recherche, les résultats doivent être validés et raffinés par d'autres approches qui étudient des schémas de classification dans un environnement réel. Une autre fonction principale d'un schéma de classification est d'organiser et de classer les documents. La rapidité du classement de documents par les employés, au même titre que l'efficacité du repérage de documents par autrui, est un indicateur de mesure qui nous semble pertinent pour évaluer les qualités fonctionnelles d'un schéma de classification des documents électroniques dans un contexte de gestion décentralisée des ressources informationnelles. Qu'il nous soit permis de rappeler qu'un schéma de classification peut être théoriquement très bien conçu, mais si ce modèle demande trop de temps et trop de peine à l'utilisateur, ce n'est peut-être pas le meilleur. Afin de valider les résultats de notre recherche dans un autre

contexte, il serait pertinent de demander à des employés volontaires d'utiliser dans des contraintes de travail réelles un schéma de classification dont la macro-structure est peu étendue, la logique est basée partiellement sur une division par classes d'activités et la sémantique utilise peu d'abréviations pour l'organisation et le classement de leurs documents électroniques. Les employés pourraient être invités à noter dans un journal de bord les circonstances de l'utilisation et des modifications apportées au schéma de classification dans le cadre d'une étude de cas longitudinale. Une recherche de cette nature permettrait de vérifier si le schéma de classification testé est fonctionnel et satisfaisant du point de vue des employés et de mieux connaître les raisons de leur insatisfaction.

6.3.3 Vérifier l'influence de facteurs socio-démographiques sur les caractéristiques des schémas de classification personnels

Nous suggérons aussi de vérifier l'influence des facteurs socio-démographiques sur les caractéristiques des schémas de classification personnels. Les conclusions de recherches antérieures sur la gestion personnelle de l'information suggèrent que différents facteurs socio-démographiques pourraient influencer le mode d'organisation de l'information par les individus. Notre recherche a tenu compte de ces présomptions, et particulièrement de la variable « rôle professionnel » lors de l'échantillonnage des schémas de classification. L'analyse des résultats suggère que les caractéristiques structurelles, logiques et sémantiques d'un schéma de classification utilisé par les employés peuvent être influencées, entre autres, par la nature des tâches à réaliser, la formation reçue en matière de gestion de documents ou le caractère partagé ou non d'un schéma de classification. La nature des tâches à réaliser suggère qu'un individu peut avoir besoin d'accéder plus fréquemment à certains types de documents que d'autres, obligeant un niveau de formalisation élevé du mode d'organisation des documents électroniques en vue d'un repérage efficace de ces documents. Par ailleurs, nous avons observé que la formation reçue par les employés en matière de gestion des documents est un facteur qui pourrait encourager l'utilisation de plusieurs classes du schéma de classification institutionnel pour l'organisation décentralisée des documents. Ces

assertions n'ont pas été vérifiées statistiquement dans notre étude. La vérification du lien entre la nature des tâches de l'employé, les années d'expérience, le caractère partagé ou non du mode d'organisation des documents électroniques et le niveau de formalisation d'un schéma de classification pourrait être réalisée dans le cadre d'une recherche ultérieure. Il serait alors pertinent d'élargir l'étude en y incluant des schémas de classification conçus par d'autres types d'employés et de corrélérer les caractéristiques des schémas de classification utilisés par ces employés avec les données démographiques et professionnelles collectées. Ce faisant, il serait possible de mieux cibler les pratiques et besoins des employés en matière d'organisation des documents électroniques et de développer une connaissance plus généralisée, au sens statistique du terme, des caractéristiques des schémas de classification personnels des documents électroniques.

6.3.4 Étudier les schémas de classification personnels utilisés pour l'organisation et le repérage des documents papier et des courriels

Il serait pertinent d'étudier les schémas de classification personnels utilisés pour l'organisation et le repérage des documents papier et des courriels. Dans le cadre de leur travail, les employés utilisent différents formats technologiques (document papier, courriels, pages Web, documents bureautiques, etc.) qui constituent leur espace documentaire personnel. Ces différents formats technologiques sont le plus souvent organisés selon différents systèmes ou applications. Les chercheurs qui se sont intéressés à l'organisation personnelle des documents ont le plus souvent limité leur étude à un seul format technologique en raison de la difficulté d'analyse causée par la diversité d'outils d'organisation proposés ou de besoins personnels différents d'un environnement à l'autre (Henderson, 2003). Afin de simplifier notre démarche, nous avons choisi de limiter notre recherche à l'étude des schémas de classification conçus et utilisés par les employés pour organiser les documents électroniques classés dans « Mes Documents » (ou son équivalent). La raison de ce choix est liée au fait que presque tous les documents de bureau sont depuis quelques années créés grâce à l'outil électronique, ce qui nous permettait de supposer l'application régulière d'une classification pour organiser un corpus sans doute volumineux de documents électroniques. Dans le cadre d'une

recherche ultérieure, il serait possible d'appliquer la même méthodologie pour caractériser et évaluer les schémas de classification personnels utilisés pour l'organisation des documents papier et des courriels. Les messages électroniques sont rarement imprimés mais ils peuvent posséder une valeur administrative, légale ou financière pour l'organisme, au même titre que tout document créé ou reçu dans un contexte de travail. L'examen de l'organisation personnelle de ce type particulier de documents nous apparaît particulièrement pertinent dans un tel contexte.

6.3.5 Identifier plus précisément les facteurs d'influence de l'efficacité du repérage

Il nous apparaît intéressant de poursuivre la recherche en identifiant plus précisément les facteurs d'influence de l'efficacité du repérage de documents ciblés. En raison de la taille de l'échantillon et de la méthode d'analyse statistique retenue (i.e., analyse de variance), nous avons pu étudier l'efficacité du repérage de documents électroniques à l'aide de schémas de classification ayant des caractéristiques variées. Cependant, les calculs de variance ne permettent pas d'identifier avec précision quels indicateurs de mesure des dimensions identifiées pourraient affecter positivement l'efficacité du repérage. Par exemple, s'agit-il de la complexité de la structure? De l'ordre de succession des classes? Une telle précision serait souhaitable pour améliorer la conception de schémas de classification compréhensibles par autrui. À cet effet, une approche corrélationnelle pourrait être envisagée pour vérifier l'existence d'une corrélation entre les indicateurs de mesure constitutifs de notre modèle d'analyse et l'efficacité du repérage. Une telle recherche pourrait être envisagée à condition de disposer d'un échantillon de taille suffisante (i.e., au moins 40 schémas de classification) composé de schémas de classification très distincts aux plans structurels, logiques et sémantiques.

6.3.6 Explorer la possibilité d'utiliser une classification à facettes pour l'organisation personnelle des documents administratifs électroniques

Enfin, il nous paraît pertinent d'explorer la possibilité de remplacer la structure arborescente traditionnelle par une classification à facettes. Les schémas de classification personnels utilisés pour l'organisation des documents administratifs électroniques sont représentés dans l'environnement informatique par une structure arborescente de répertoires et de sous-répertoires. Les limites inhérentes à une telle structure, déjà soulignées par plusieurs chercheurs (Barreau, 1995; Sénécal, 1999; Dourish *et al.*, 2000), obligent les employés à devoir arbitrairement privilégier comme base de classification tantôt le thème du document, tantôt le type de document ou encore l'activité à laquelle il se rattache, alors que le document pourrait de fait être classifié à plusieurs endroits. Beaucoup de recherches ont été réalisées et continuent à être menées pour tenter de changer le paradigme hiérarchique de l'interface technologique (Freeman & Gelernter, 1996; Robertson *et al.*, 1998; Dourish *et al.*, 1999; Rekimoto, 1999; Dourish *et al.*, 2000; Bergman, Beyth-Marom & Nachmias, 2003). La classification à facettes se présente comme une combinaison de groupes conceptuels plutôt que comme une liste structurée de sujets. L'application d'une classification à facettes permettrait d'éliminer certaines contraintes liées au fait que le schéma de classification personnel soit un langage de représentation pré-coordonné. L'analyse des schémas de classification personnels utilisés par deux catégories d'employés a révélé plusieurs facettes communes, dont « type de document » et « activités », qui pourraient servir à décrire et classifier uniformément les documents administratifs électroniques dans un contexte de gestion décentralisée. L'observation d'un usage répandu de classes qui représentent plusieurs concepts ou facettes, comme « Notes de cours [Sigle du cours] Automne 2004 », va dans le sens d'un développement souhaitable de la recherche dans ce domaine.

Pour terminer, nous souhaitons que notre recherche sur les schémas de classification et le repérage de documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles suggère d'autres recherches. Le besoin demeure de poursuivre la recherche sur la classification des documents

administratifs électroniques réalisée jour après jour par de nombreux employés. Toute recherche dans ce domaine contribuerait à repousser les limites de nos connaissances, à développer des principes et des méthodes visant à améliorer l'organisation et le repérage des documents et à assurer la conservation de la mémoire institutionnelle.

7 Bibliographie

Notes : Toutes les adresses URL étaient valides le 16 mai 2007.

- Abrams, D., Baecker, R., & Chignell, M. (1998). Information archiving with bookmarks : Personal web space construction and organization. In *Proceedings of the SIGCHI conference on Human factors in computing systems, April 18-23, 1998, Los Angeles, California, United States*. New York, NY : ACM Press, 41-48.
- Adar, E., Karger, D. R., & Stein, L. (1999). Haystack : Per-user information environment. In *Proceedings of the 8th International Conference on Information and Knowledge Management, November 2-6, 1999, Kansas City, Missouri, United States*.
<http://www.hpl.hp.com/research/idl/people/eytan/p413-adar.pdf>
- Alvarado, C. et al. (2003). *Surviving the information explosion : How people find their electronic information*. Cambridge, MA : Massachusetts Institute of Technology, Artificial Intelligence Laboratory.
<http://haystack.lcs.mit.edu/papers/alvarado.aim03.pdf>
- Ancelle, T. (2002). *Statistique épidémiologie*. Paris : Maloine.
- Archives nationales du Canada. (1996). *Gestion des documents électroniques dans un environnement de travail électronique*. [Ottawa, ON] : Archives nationales du Canada, Division des normes et pratiques liées à la gestion de l'information.
- Archives nationales du Québec. (1996). *Normes et procédures archivistiques des Archives nationales du Québec*. Québec, QC : Les Publications du Québec.
- Bälter, O. (1997). Strategies for organising email. In *Proceedings of Human Computer Interaction on People and Computers XII, London, United Kingdom*. London, UK : Springer-Verlag, 21-38.
- Barreau, D. K. (1995). Context as a factor in personal information management systems. *Journal of the American Society for Information Science*, 46 (5), 327-339.
- Barreau, D. K. (1997). *Information systems for organizations and the problem of ephemeral information*. Dissertation submitted to the Faculty of the Graduate School of the University of Maryland at College Park in partial fulfillment of the requirements for the degree of Doctor of Philosophy.
- Barreau, D. K., & Nardi, B. A. (1995). Finding and reminding : File organization from the desktop. *SIGCHI Bulletin*, 27 (3), 39-43.

- Bates, M. (1977). Factors affecting subject catalog search success. *Journal of the American Society for Information Science*, 28 (3), 161-169.
- Bearman, D. A. (1989). *Archival methods*. Pittsburgh, PA : Archives & Museum Informatic.
- Bearman, D. A. (1997). The physical archives and the virtual archives. *Archivum*, XLIII, 150-167.
- Beaupré, L. (2004). Classification des documents électroniques : expérience à la Sûreté du Québec. In *Actes du 33ème Congrès de l'Association des archivistes du Québec, 27-29 mai 2004, Sainte-Adèle, Québec, Canada*. Québec, QC : Association des archivistes du Québec, 54-57.
- Beghtol, C. (1998). General classification systems : Structural principles for multidisciplinary specification. In *Proceedings of the 5th International ISKO Conference, August 25-29, 1998, Lille, France*. Würzburg, Germany : Ergon, 89-96.
- Beghtol, C. (2003). Classification for information retrieval and classification for knowledge discovery : Relationships between professional and naive classifications. *Knowledge Organization*, 30 (2), 64-73.
- Bellot, P. (2000). *Méthodes de classification et de segmentation locales non supervisées pour la recherche documentaire*. Thèse de doctorat en informatique, Université d'Avignon et des Pays du Vaucluse.
- Bellotti, V. et al. (2003). Taking email to task : The design & evaluation of a task management centered email tool. In *Proceedings of the SIGCHI conference on Human factors in computing systems, April 5-10, 2003, Fort Lauderdale, Florida, United States*, 345-352.
<http://portal.acm.org/citation.cfm?id=642611.642672>
- Benfell, P. (2002). An integrated approach to managing electronic records. *Records Management Journal*, 12 (3), 94-97.
- Berchtold, S., Biliris, A., & Panagos, E. (1999). SaveMe : A system for archiving electronic documents using messaging groupware. In *Proceedings of the international joint conference on work activities coordination and collaboration, February 22-25, 1999, San Francisco, California, United States*, 167-176.
<http://portal.acm.org/citation.cfm?id=295684>
- Bergeron, P. (1992). La gestion des archives électroniques : quelques questions-clés à considérer. *Archives (Revue de l'Association des archivistes du Québec)*, 23 (3), 51-70.

- Bergman, O., Beyth-Marom, R., & Nachmias, R. (2003). The user-subjective approach to personal information management systems. *Journal of the American Society for Information Science and Technology*, 54 (9), 872-878.
- Bergquist, M. & Ljungberg, J. (1999). Genre in action : Negotiating Genres in Practice. In *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences*. Washington, DC : IEEE Computer Society.
<http://ieeexplore.ieee.org/iel5/6293/16782/00772655.pdf>
- Berlin, L. M. *et al.* (1993). Where did you put it? Issues in the design and use of a group memory. In *Proceedings of the SIGCHI conference on Human factors in computing systems, April 24-29, 1993, Amsterdam, The Netherlands*. Amsterdam, The Netherlands : IOS Press, 23-30.
- Bertrand-Gastaldy, S. (1990). L'indexation assistée par ordinateur : un moyen de satisfaire les besoins collectifs et individuels des utilisateurs de bases de données textuelles dans les organisations. *ICO Québec*, 71-91.
- Bertrand-Gastaldy, S., & Deschatelets, G. (2002). *Premiers résultats d'un test d'indexation et de repérage avec différents types de représentations du contenu, dans Archimède, le catalogue en ligne de la Bibliothèque de l'École Polytechnique de l'Université de Montréal*.
<http://www.ling.uqam.ca/sato/publications/bibliographie/Asted.htm>
- Blomberg, J., Suchman, L., & Trigg, R. (1994). Reflections on a work-oriented design project. In *Proceedings of the Participatory Design Conference, October 27-28, 1994, Chapel Hill, North Carolina, United States*. New York, NY : ACM Press, 99-109.
- Boardman, R. (2001a). Category overlaps between hierarchies in user workspace. In *Proceedings of Interact'01, the 8th IFIP TC 13 International Conference on Human-Computer Interaction, Tokyo, Japan*.
<http://www.iis.ee.ic.ac.uk/~rick/research/pubs/categories-interact2001.pdf>
- Boardman, R. (2001b). Multiple hierarchies in user workspace. In *Proceedings of Conference on Human Factors in Computing Systems, March 31-April 5, 2001, Seattle, Washington*, 403-404.
<http://www.iis.ee.ic.ac.uk/~rick/research/pubs/workspace-chi2001.pdf>
- Boardman, R. (2002). Workspaces that work : Towards unified personal information management. In *Proceedings of the 16th British HCI Conference, People and Computers XVI, September 2-6, 2002, South Bank University, London, United Kingdom*, 216-217.
<http://www.iis.ee.ic.ac.uk/~rick/research/pubs/wthatw-hci2002.pdf>

- Boardman, R., Spence, R., & Sasse, M. A. (2003). Too many hierarchies? The daily struggle for control of the workspace. In *Proceedings of HCI International*, 616-620.
<http://www.iis.ee.ic.ac.uk/~rick/research/pubs/struggle-hcii2003.pdf>
- Boardman, R., & Sasse, M. A. (2004). Stuff goes into the computer and doesn't come out : A cross-tool study of personal information management. In *Proceedings of the SIGCHI conference on Human factors in computing systems, April 24-29, 2004, Vienna, Austria*, 583-590.
<http://www.iis.ee.ic.ac.uk/~rick/research/pubs/boardman-chi04.pdf>
- Borlund, P., & Ingwersen, P. (1998). Measures of relative relevance and ranked half-life : Performance indicators for interactive IR. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval, August 24-28, 1998, Melbourne, Australia*, 324-331.
<http://portal.acm.org/citation.cfm?id=291019>
- Boutet, L. (2000). *Le plan de classification et les documents électroniques : rapport du Groupe de travail sur la classification des documents électroniques*. [Québec, QC] : Ministère des relations internationales.
- Bowker, G. C., & Star, S. L. (2000). *Sorting things out : Classification and its consequences*. Cambridge, MA : MIT Press.
- Broughton, V. (2000). Classification schemes revisited : Applications to web indexing and searching. *Journal of Internet Cataloging*, 2 (3-4), 143-155.
- Buchanan, B. (1979). *Theory of library classification*. London, UK : C. Bingley; K.F. Saur.
- Busha, C. H., & Harter, S. P. (1980). *Research methods in librarianship : Techniques and interpretation*. New York, NY : Academic Press.
- Canonne, A. (1993). *Vocabulaire élémentaire des classifications*. Liège, Belgique : Éditions du CÉFAL. (Collection Bibliothèque du bibliothécaire; 2).
- Carroll, J. M. (1982). Creative names for personal files in an interactive computing environment. *International Journal of Man-Machine Studies*, 16, 405-438.
- Casault, D. (1995). *Plan de classification type des documents de gestion gouvernementaux*. Québec, QC : Direction des communications, Ministère de la culture et des communications.
- Casault, D., & Lévesque, M. (1997). *Guide de classification et de classement des documents des associations et autres organismes de même nature*. Québec, QC : Association des archivistes du Québec.

- Chabin, M.-A. (1997). La cinquième dimension de l'archive numérique. *Document numérique*, 1 (2), 205-216.
- Chamard, J.-P. (2004). La classification des documents électroniques : expérience du ministère de la Justice du Québec. In *Actes du 33ème Congrès de l'Association des archivistes du Québec, 27-29 mai 2004, Sainte-Adèle, Québec, Canada*. Québec, QC : Association des archivistes du Québec, 58-62.
- Champagne, M. & Chouinard, D. (1987). *Le traitement d'un fonds d'archives : ses documents historiques*. La Pocatière, QC : Documentor.
- Chan, L. M. (1995). Classification, present and future. *Cataloging & Classification Quarterly*, 21 (2), 5-17.
- Chapman, H. (1999). *The file naming habits of personal computer users*. A Master's paper for the M.S. in I.S. degree, University of North Carolina at Chapel Hill.
- Chouinard, D., Rousseau, J.-Y., & Roy, M. (1990). La gestion des archives informatiques : l'expérience de l'Université de Montréal. *Archives (Revue de l'Association des archivistes du Québec)*, 21 (4), 41-58.
- Chu, H. (2003). *Information representation and retrieval in the digital age*. Medford, NJ : American Society for Information Science and Technology by Information Today.
- Cleverdon, C. (1967). The Cranfield tests on index language devices. *ASLIB Proceedings*, 19 (6), 173-194.
- Cochrane, P. A. (1985). Preparing for the use of classification in online computer cataloging systems and in online catalogs. *Information Technology and Libraries*, 4, 91-111.
- Cole, I. (1982). Human aspects of office filing : Implications for the electronic office. In *Proceedings of the Human Factors Society 26th Annual Meeting, Santa Monica, California, United States*. Santa Monica, CA : Human Factors Society, 59-63.
- Conférence des recteurs et des principaux des universités du Québec (CRÉPUQ). (1994). *La gestion des archives informatiques*. Sainte-Foy, QC : Presses de l'Université du Québec.
- Conférence des recteurs et des principaux des universités du Québec (CRÉPUQ). (2004). *La gestion des documents numériques des établissements universitaires du Québec : état de situation et planification stratégique*. <http://www.crepucq.ca/documents/arch/Rapport-GGDN.htm>

- Conseil international des archives (1997). *Guide pour la gestion archivistique des documents électroniques*. Paris : Conseil international des archives.
- Contandriopoulos, A.-P. et al. (1990). *Savoir préparer une recherche : la définir, la structurer, la financer*. Montréal, QC : Presses de l'Université de Montréal.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation : Design and analysis issues for field settings*. Chicago, IL : David McNally.
- Couture, C. (1999). La politique de gestion des archives. In Couture, C., & collaborateurs. *Les fonctions de l'archivistique contemporaine*. Sainte-Foy, QC : Presses de l'Université du Québec, 3-30.
- Couture, C., & Rousseau, J-Y. (1982). *Les archives au XXème siècle : une réponse aux besoins de l'administration et de la recherche*. Montréal, QC : Université de Montréal, Secrétariat général.
- Dahlberg, I. (1988). Concept and definition theory. In *Classification theory in the computer age : Conversations across the disciplines. Proceedings from the Conference, November 18-19, 1988, Albany, New York, United States*. New York, NY : Nelson A. Rockefeller College of Public Affairs and Policy, University at Albany, State University of New York, 12-24.
- Dahlberg, I. (1998). Classification structure principles : Investigations, experiences, conclusions. In *Proceedings of the 5th International ISKO Conference, August 25-29, 1998, Lille, France*. Würzburg, Germany : Ergon, 80-88.
- Daum, P. B. (1997). Technology and the four-level information hierarchy. *Records Management Quarterly*, 31 (4), 8-14.
- De Grolier, É. (1988). Taxilogie et classification : un essai de mise au point et quelques notes de prospective. *Bulletin des bibliothèques de France*, 33 (6), 468-489.
- Delsalle, P. (1996). *Lexique des archives et documents historiques du papyrus au vidéodisque*. Paris : Nathan. (Collection 128; 118).
- Dhérent, C., & collaborateurs. (2002). *Les archives électroniques : manuel pratique*. [Paris] : Direction des archives de France.
<http://www.archivesdefrance.culture.gouv.fr/fr/archivistique/>
- Direction des archives de France. (1993). *La Pratique archivistique française*. Paris : Archives nationales.

- Douglas, M. (1995). Forgotten knowledge. In Strathern, M. (Ed.). *Shifting contexts*. London, UK : Routledge, 13-30.
- Dourish, P. *et al.* (1999). Presto : An experimental architecture for fluid interactive document spaces. *ACM Transactions on Computer Human Interaction*, 6 (2), 133-161.
<http://portal.acm.org/citation.cfm?id=319099>
- Dourish, P. *et al.* (2000). Extending document management systems with user-specific active properties. *ACM Transaction on Information Systems*, 18 (2), 140-170.
<http://www.cc.gatech.edu/~keith/pubs/tois-placeless.pdf>
- Duchain, M. (1977). Le respect des fonds en archivistique : principes théoriques et problèmes pratiques. *La Gazette des archives*, (97), 71-96.
- Duchain, M. (1998). Le principe de provenance et la pratique du tri, du classement et de la description en archivistique contemporaine. *Janus*, (1), 87-100.
- Dumais, S. T., & Landauer, T. K. (1983). Using examples to describe categories. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems, December 12-15, 1983, Boston, Massachusetts, United States*, 112-115.
<http://portal.acm.org/citation.cfm?id=800045.801592>
- Dumais, S. T. *et al.* (2003). Stuff I've seen : A system for personal information retrieval and re-use. In *Proceedings of the 26th annual international ACM SIGIR Conference on Research an Development in Information Retrieval, July 28-August 1, 2003, Toronto, Canada*, 72-79.
<http://portal.acm.org/citation.cfm?id=860435.860451>
- Éthier, G. (1989). *Introduction à la gestion des documents*. Boucherville, QC : G. Vermette.
- Evans, F. B., Harrison, D. F., & Thompson, E. A. (1974). A basic glossary for archivists, manuscript curators, and records managers. *The American Archivist*, 37 (3), 415-433.
- Faillès, B. (1997). Les nouvelles dispositions de la législation américaine sur l'accès aux documents classifiés. In *Transparence et secret. L'accès aux archives contemporaines*. Paris : L'Association des archivistes français, 145-148.
- Fallis, D., & Mathiesen, K. (2000). Consistency rules for classification schemes (or how to organize your beanie babies). In *Proceedings of the 6th International ISKO Conference, July 10-13, 2000, Toronto, Canada*. Würzburg, Germany : Ergon, 339-344.

- Farrell, S. *et al.* (2002). Information programming for personal user interfaces. In *Proceedings of the 7th international conference on intelligent user interfaces, January 13-16, 2002, San Francisco, California, United States*, 190-191.
<http://portal.acm.org/citation.cfm?id=502751>
- Ferraioli, L. (2005). An exploratory study of metadata creation in a health care agency. *Cataloging & Classification Quarterly*, 40 (3/4), 75-102.
- Fertig, S., Freeman, E., & Gelernter, D. (1996). Lifestream : An alternative to the desktop metaphor. In *ACM Conference in Computer Human Interaction (CHI), April 13-18, 1996*, 410-411.
<http://people.cs.vt.edu/~north/infoviz/lifestreams.pdf>
- Fortin, M-F., Côté, J., & Fillion, F. (2006). *Fondements et étapes du processus de recherche*. Montréal, QC : Chenelière éducation.
- Freeman, E., & Gelernter, D. (1996). Lifestreams : A storage model for personal data. *SIGMOD Record*, 25 (1), 80-86.
<http://portal.acm.org/citation.cfm?id=381893>
- Gagnon-Arguin, L. (1999). La création. In Couture, C., & collaborateurs. *Les fonctions de l'archivistique contemporaine*. Sainte-Foy, QC : Presses de l'Université du Québec, 69-101.
- Gagnon-Arguin, L., & Bannouri, R. (1998). *La classification des documents institutionnels dans les ministères du gouvernement du Québec et le défi informatique : rapport de recherche*. Montréal, QC : Université de Montréal, École de bibliothéconomie et des sciences de l'information.
<http://www.msg.gouv.qc.ca/fr/publications/enligne/administration/ingenierie/classification.pdf>
- Garon, L. (1992). Le fonds d'archives : objet premier de la normalisation en archivistique. In Couture, C. (dir.). *La normalisation en archivistique : un pas de plus dans l'évolution d'une discipline. Mélanges Jacques Ducharme*. Québec, QC : Documentor et Association des archivistes du Québec, 105-125.
- Gemmell, J. *et al.* (2002). MyLifeBits : Fulfilling the Memex vision. In *Proceedings of ACM Multimedia' 02*, 235-238.
<http://research.microsoft.com/~jgemzell/pubs/MyLifeBitsMM02.pdf>
- Gonçalves, D. J., & Jorge, J. A. (2003). An empirical study of personal document spaces. In *Proceedings DSV-IS'03, June 4-6, 2003, Funchal, Madeira Island, Portugal*, 46-60.
http://immi.inesc-id.pt/publication.php?publication_id=34

- Guercio, M. (2001). Principles, methods, and instruments for the creation, preservation, and use of archival records in the digital environment. *The American Archivist*, 64, 238-269.
- Guercio, M. (2002). Records classification and content management : Old functions and new requirements in the legislations and standards for electronic record-keeping systems. In *Proceedings of the DLM Forum 2002. Access and preservation of electronic information : Best practices and solutions, May 6-8, 2002, Barcelona, Spain*. Luxembourg : Office for Official Publications of the European Communities, 432-442.
- Hackett, Y. (1997). Le programme relatif aux documents informatiques aux Archives nationales du Canada. In *Actes du 26ème Congrès de l'Association des archivistes du Québec, 28-31 mai 1997, Aylmer*. Québec, QC : Association des archivistes du Québec, 163-173.
- Hedstrom, M. (1996). Electronic records research : What have archivists learned from the mistakes of the past? *Archives and Museum Informatics*, 10 (4), 313-325.
- Henderson, S. (2003). *Information workspaces : Investigating the information behaviour of knowledge workers and its implications for the design of usable information workspaces*. PhD Dissertation Proposal, University of Auckland.
- Henderson, S. (2004). Personal digital document management. In *6th Asia-Pacific Conference on Computer-Human Interaction, June 29-July 2, 2004, Rotorua, New Zealand*. Berlin : Springer, 651-655.
- Henderson, S. (2005). Genre, task, topic and time : Facets of personal digital document management. In *Proceedings of the 6th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction, Auckland, New Zealand*, 75-82.
<http://doi.acm.org/10.1145/1073943.1073957>
- Héon, G. (1995). Les plans de classification en archivistique. *Archives (Revue de l'Association des archivistes du Québec)*, 27 (1), 73-89.
- Héon, G. (1999). La classification. In Couture, C., & collaborateurs. *Les fonctions de l'archivistique contemporaine*. Sainte-Foy, QC : Presses de l'Université du Québec, 219-254.

- Hertzum, M. (1993). Information retrieval in a work setting : A case study of the documentation part of chemists' work. In *Proceedings of the 16th IRIS. Information Systems Research Seminar in Scandinavia, August 7-10, 1993, Copenhagen, Denmark*, 786-798.
<http://akira.ruc.dk/~mhz/Research/Publ/IRIS1993.pdf>
- Hertzum, M. (1999). Six roles of documents in professionals' work. In *Proceedings of the 6th European Conference on Computer Supported Cooperative Work, September 12-16, 1999, Copenhagen, Denmark*, 41-60.
<http://www.dcs.qmw.ac.uk/~mounia/PROJECTS/DocRoles.doc>
- Hertzum, M. (2004). Small-scale classification schemes : A field study of requirements engineering. *Computer Supported Cooperative Work*, 13 (1), 35-61.
http://akira.ruc.dk/~mhz/Research/Publ/JCSCW2004_preprint.pdf
- Hertzum, M., & Frøkjær, E. (1996). Browsing and querying in online documentation : A study of user interfaces and the interaction process. *ACM Transactions on Computer-Human Interaction*, 3 (2), 136-161.
http://akira.ruc.dk/~mhz/Research/Publ/TOCHI1996_preprint.pdf
- Hildreth, C. R. (2000). Are Web-based OPACs more effective retrieval systems than their conventional predecessors? An experimental study. In *Proceedings of the 6th International ISKO Conference, July 10-13, 2000, Toronto, Canada. Würzburg, Germany : Ergon*, 237-246.
- Hjørland, B. (2007). Semantics and knowledge organization. *Annual Review of Information Science and Technology*, 41, 367-405.
- Hjørland, B., & Albrechtsen, H. (1999). An analysis of some trends in classification research. *Knowledge organization*, 26 (3), 131-139.
- Hudon, M. (2000). Innovation and tradition in knowledge organization schemes on the Internet, or, finding one's way in the virtual library. In *Proceedings of the 6th International ISKO Conference, July 10-13, 2000, Toronto, Canada. Würzburg, Germany : Ergon*, 35-40.
- Hudon, M. (2001). Structuration du savoir et organisation des collections dans les répertoires du Web. *Bulletin des bibliothèques de France*, 46 (1), 57-62.

- Hudon, M., & Mas, S. (2001). *Analyse des facettes pour la classification des documents institutionnels au gouvernement du Québec*. Rapport présenté pour le Groupe de travail en classification et indexation. Montréal, QC : Université de Montréal, École de bibliothéconomie et des sciences de l'information. (Collection en ingénierie documentaire; 13).
http://www.msg.gouv.qc.ca/fr/publications/enligne/administration/ingenierie/classification_analyse.pdf
- Hudon, M., Mas, S., & Gazo, D. (2005). Structure, logic, and semantics in *ad hoc* classification schemes applied to web-based libraries in the field of education. *Canadian Journal of Information and Library Science/Revue canadienne des sciences de l'information et de bibliothéconomie*, 29 (3), 265-288.
- Hunter, E. J. (2000). Do we still need classification? In Marcella, R., & Maltby, A. (eds.). *The future of classification*. London, UK : Gower, 1-17.
- Hunter, E. J. (2002). *Classification made simple*. Aldershot, Hants, England ; Burlington, Vermont : Ashgate.
- International Organization for Standardization. (2001). *Norme internationale ISO 15489-1 : Information et documentation - "Records management". Partie 1 : Principes directeurs*. [Genève, Suisse] : International Organization for Standardization.
- Iyer, H. (1995). *Classificatory structures : Concepts, relations and representation*. Frankfurt/Main : Indeks. (Textbooks for knowledge organization; v. 2).
- Jacob, E. K. (2004). Classification and categorization : A difference that makes a difference. *Library Trends*, 52 (3), 515-540.
- Jones, R. A. (2000). *Méthodes de recherche en sciences humaines*. Paris : De Boeck Université.
- Jones, W. P. (2004). Finders, keepers? The present and future perfect in support of personal information management. *First Monday*, 9 (3).
http://www.firstmonday.dk/issues/issue9_3/jones/
- Jones, W. P. (2007). Personal information management. *Annual Review of Information Science and Technology*, 41, 453-504.
- Jones, W. P., Bruce, H., & Dumais, S. T. (2001). Keeping found things found on the web. In *Proceedings of the 2001 ACM CIKM 10th International Conference on Information and Knowledge Management, November 5-10, 2001, Atlanta, United States*, 119-134.
http://kftf.ischool.washington.edu/KFTF_Web.pdf

- Jones, W. P., Dumais, S. T., & Bruce, H. (2002). Once found, what then? : A study of "keeping" behaviors in the personal use of web information. In *Proceedings of the 65th ASIST Annual Meeting, Philadelphia, United States*, 391-402.
[http://kftf.ischool.washington.edu/Jones %202002 %20ASIST.pdf](http://kftf.ischool.washington.edu/Jones%202002%20ASIST.pdf)
- Jones, W. P., & Maier, D. (2003). Personal information management group report. In *National Science Foundation (NSF) Information and Data Management (IDM) Workshop, September 14-16, 2003, Seattle, Washington, United States*.
[http://kftf.ischool.washington.edu/Summary %20of %20IDM03 %20session %20on %20Personal %20Information %20Management.pdf](http://kftf.ischool.washington.edu/Summary%20of%20IDM03%20session%20on%20Personal%20Information%20Management.pdf)
- Jones, W.P. *et al.* (2005). Don't take my folders away! : Organizing personal information to get things done. In *CHI'05 extended abstracts on Human factors in computing systems, April 2-7, 2005, Portland, Oregon, United States*, 1505-1508.
<http://portal.acm.org/citation.cfm?id=1056952&dl=GUIDE&coll=GUIDE&CFID=48509855&CFTOKEN=8097759>
- Kaiser, F. M. (2003). Access to classified information : Seeking security clearances for state and local officials and personnel. *Government Information Quarterly*, 20 (3), 213-232.
- Kaptelinin, V. (1996). Creating computer-based work environments : An empirical study of MacIntosh users. In *Proceedings of the 1996 ACM SIGCPR/SIGMIS Conference on Computer personnel research, April 1996, Denver, Colorado, United States*, 360-366.
<http://portal.acm.org/citation.cfm?id=238921>
- Kaptelinin, V. (2003). UMEA : Translating interaction histories into project contexts. In *Proceedings of CHI, April 5-10, 2003, Fort Lauderdale, Florida, United States*, 353-360.
<http://portal.acm.org/citation.cfm?id=642673>
- Karjalainen, A. *et al.* (2000). Genre-based metadata for enterprise document management. In *Proceedings of the 33rd Hawaii International Conference on System Sciences, January 4-7, 2000, Island of Maui, Hawaii*.
<http://csdl.computer.org/comp/proceedings/hicss/2000/0493/03/04933013.pdf>
- Khoo, C. *et al.* (2007). How users organize electronic files on their workstations in the office environment : A preliminary study of personal information organization behaviour. *Information Research*, 12 (2).
<http://InformationR.net/ir/12-2/paper293.html>

- Kidd, A. (1994). The marks are on the knowledge worker. In *Proceedings of the SIGCHI conference on Human factors in computing systems, April 24-28, 1994, Boston, Massachusetts, United States*, 186-191.
<http://portal.acm.org/citation.cfm?id=191740>
- Kirsh, D. (2001). The context of work. *Human Computer Interaction*.
<http://interactivity.ucsd.edu/articles/HCI/final.html>
- Krawczyk, B. (1999). Cross reference heaven : The abandonment of the fonds as the primary level of arrangement for Ontario government records. *Archivaria*, 48, 131-153.
- Krikelas, J. (1969). Subject searches using two catalogs : A comparative evaluation. *College & Research Libraries*, 30 (6), 506-517.
- Kwasnik, B. H. (1989). How a personal document's intended use or purpose affects its classification in an office. In *Proceedings of the 12th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Cambridge, Massachusetts, United States*, 207-210.
<http://portal.acm.org/citation.cfm?id=75334.75356>
- Kwasnik, B. H. (1991). The importance of factors that are not document attributes in the organisation of personal documents. *Journal of Documentation*, 47 (4), 389-398.
- Kwasnik, B. H. (1999). The role of classification in knowledge representation and discovery. *Library Trends*, 48 (1), 22-47.
- Lancaster, F. W. (1968). *Information retrieval systems characteristics, testing, and evaluation*. New York, NY : Wiley.
- Lancaster, F. W. (1971). Aftermath of an evaluation. *Journal of Documentation*, 27, 1-10.
- Langridge, D. W. (1992). *Classification : Its kinds, elements, systems, and applications*. London, UK : Bowker-Saur.
- Lansdale, M. (1988). The psychology of personal information management. *Applied Ergonomics*, 19 (1), 55-66.
- Lansdale, M., & Edmonds, E. (1992). Using memory for events in the design of personal filing systems. *International Journal of Man-Machine Studies*, 36, 97-126.

- Lemaire, P. (1999). *Psychologie cognitive*. Paris : De Boeck Université.
- Lévesque, M. (1998-1999). Les documents électroniques et le calendrier de conservation, ou les considérations d'un pauvre archiviste qui essaie de suivre l'évolution technologique. *Archives (Revue de l'Association des archivistes du Québec)*, 30 (3-4), 39-54.
- Lin, X., & Chan, L. M. (1999). Personalized knowledge organization and access for the Web. *Library & Information Science Research*, 21 (2), 153-172.
- Lytle, R. H. (1980). Intellectual access to archives : II. Report of an experiment comparing provenance and content indexing methods of subject retrieval. *The American Archivist*, 43, 191-205.
- MacNeil, H. (1992). The context is all : Describing a fonds and its parts in accordance with the Rules for Archival Description. In Eastwood, Terry. *Le fonds d'archives : de la théorie à la pratique*. Ottawa, ON : Bureau canadien des archivistes, 195-225.
- Mai, J.-E. (2000). Likeness : A pragmatic approach. In *Proceedings of the 6th International ISKO Conference, July 10-13, 2000, Toronto, Canada*. Würzburg, Germany : Ergon, 23-27.
- Mai, J.-E. (2004). Classification in context : Relativity, reality, and representation. *Knowledge Organization*, 31 (1), 39-48.
- Malone, T. W. (1983). How do people organize their desks? Implications for the design of office information systems. *ACM Transactions on Office Information Systems*, 1 (1), 99-112.
- Maltby, A., & Marcella, R. (2000). Organizing knowledge : The need for system and unity. In Marcella, R., & Maltby, A. (eds). *The future of classification*. London, UK : Gower, 19-31.
- Maniez, J. (1987). *Les langages documentaires et classificatoires : conception, construction et utilisation dans les systèmes documentaires*. Paris : Éditions d'Organisation.
- Maniez, J. (1991). A decade of research in classification. *International Classification*, 18 (2), 73-77.
- Marcella, R., & Newton, C. (1994). The functions of classification in the library. In Marcella, R., & Newton, C. *A new manual of classification*. Aldershot, England : Gower, 159-181.

- Marsden, G., & Cairns, D. E. (2003). Improving the usability of the hierarchical file system. In *Proceedings of the 2003 Annual research conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT) on Enablement through technology, September 17-19, 2003*, 122-129.
<http://portal.acm.org/citation.cfm?id=954027>
- Mas, S., & Gagnon-Arguin, L. (2003-2004). Pour un approfondissement de la "notion" de dossier dans la gestion de l'information organique et consignée d'une organisation. *Archives (Revue de l'Association des archivistes du Québec)*, 35 (1-2), 29-48.
- Metzger, J.-P., & Lallich-Boidin, G. (2004). Temps et documents numériques. *Document numérique*, 8 (4), 11-21.
- Molholt, P. (1995). Qualities of classification schemes for the information superhighway. *Cataloging & Classification Quarterly*, 21 (2), 19-22.
- Nardi, B., Anderson, K., & Erickson, T. (1994). *Filing and finding computer files*. Cupertino, CA : Apple Computer Inc. (Technical Report; 118).
- Nardi, B., & Barreau, D. (1997). "Finding and reminding" revisited : Appropriate metaphors for file organization at the desktop. *SIGCHI*, 29 (1), 76-78.
<http://sigchi.org/bulletin/1997.1/nardi.html>
- Nasse-Kolmayer, E. (1997). *Contribution à l'analyse des processus cognitifs mis en jeu dans l'interrogation d'une base de données documentaires*. Thèse présentée en vue de l'obtention du Doctorat en Psychologie, Université René Descartes, Paris 5.
<http://www.enssib.fr/bibliotheque/documents/theses/kolmayer/kolmayer.pdf>
- Nelson, T. H. (2003). Structure, tradition and possibility. In *Proceedings of the 14th ACM conference on Hypertext and hypermedia, August 26-30, 2003, Nottingham, United Kingdom*.
<http://portal.acm.org/citation.cfm?doid=900051.900053>
- Norwood, S. L. (2000). *Research strategies*. Upper Saddle River, NJ : Prentice Hall Health.
- O'Day, V. L., & Jeffries, R. (1993). Orienteering in an information landscape : How information seekers get from here to there. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 24-29, 1993, Amsterdam, The Netherlands*, 438-445.
<http://portal.acm.org/citation.cfm?id=169365>

- Orlikowski, W., & Yates, J. (1994). Genre repertoire : The structuring of communicative practices in organizations. *Administrative Sciences Quarterly*, 33 (4), 541-574.
- Patterson, G., & Sprehe, T. J. (2002). Principal challenges facing electronic records management in federal agencies today. *Government Information Quarterly*, 19, 307-315.
- Pérotin, Y. (1961). L'administration et les trois âges des archives. *Seine-et-Paris*, 20, 1-14.
- Pétilat, C. (1999). La gestion des archives électroniques : un enjeu pour les archivistes. *L'Écluse*, 11 (2), 12-20.
- Peyrelong, M-F., & Accart, J-P. (2002). Du système d'information personnel au système d'information collectif : réalités et mirages du partage de l'information en entreprise. In *Proceeding of the 30th Annual Conference of the Canadian Association for Information Science : Advancing knowledge : Expanding Horizons for Information Science, May 30-June 1, 2002, University of Toronto : Faculty of Information Studies*, 136-149.
<http://eprints.relis.org/archive/00007305/>
- Poivre, J. (2004). *L'archivage des documents bureautiques : manuel pratique*. Paris : Direction des archives de France.
- Polit, D. F. & Beck, C. T. (2004). *Nursing research : Principles and methods*. Philadelphia, PA : Lippincott, Williams & Wilkins.
- Prom, C. J. (2004). User interactions with electronic findings aids in a controlled setting. *The American Archivist*, 67 (2), 234-268.
- Québec (Province). Assemblée nationale. (2001). *Loi concernant le cadre juridique des technologies de l'information (Loi 161)*. Québec, QC : Éditeur officiel du Québec.
- Ranganathan, S. R. (1965). *The Colon classification*. New Brunswick, NJ : Rutgers University Press. (Rutgers Series on Systems for the Intellectual Organization of Information; v. 4).
- Ravasio, P., Schar, S. G., & Krueger, H. (2004). In pursuit of desktop evolution : User problems and practices with modern desktop systems. *ACM Transactions on Computer-Human Interaction*, 11 (2), 156-180.
<http://portal.acm.org/citation.cfm?id=1005363>

- Rekimoto, J. (1999). Time machine computing : A time-centric approach for the information environment. In *Proceedings of the 12th annual ACM symposium on User interface software and technology, November 7-10, 1999, Asheville, North Carolina, United States*, 45-54.
<http://portal.acm.org/citation.cfm?id=322582>
- Renaud, J. (1975). Le traitement et la conservation des documents. *Archives*, 75 (1), 13-23.
- Ribeiro, F. (1996). Subject indexing and authority control in archives : The need for subject indexing in archives and for an indexing policy using controlled language. *Journal of the Society of Archivists*, 17 (1), 27-54.
- Richmond, P. A. (1990). General theory of classification. In Bengston, B., & Hill, J.S. *Classification of library materials : Current and future potential for providing access*. New York, NY : Neal Schuman, 16-26.
- Roberge, M. (1985). *La classification universelle des documents administratifs*. La Pocatière, QC : Documentor.
- Roberge, M. (2002). *L'essentiel de la gestion documentaire*. Québec, QC : GESTAR.
- Roberge, M. (2004). *L'essentiel de la gestion documentaire*. 2^{ème} édition revue et augmentée. Québec, QC : GESTAR.
- Robertson, G. *et al.* (1998). Data Mountain : Using spatial memory for document management. In *ACM Symposium on User Interface Software and Technology, November 1-4, 1998, San Francisco, California, United States*, 153-162.
<http://citeseer.ist.psu.edu/robertson98data.html>
- Rolland-Thomas, P. (1994). Les classifications documentaires : bref rappel historique. In Dewey, M. *Classification décimale de Dewey et index*. Montréal, QC : ASTED.
- Rousseau, J.-Y. (1980). Le choix d'un système de classement et son application. *Archives (Revue de l'Association des archivistes du Québec)*, 11 (1), 3-19.
- Rousseau, J.-Y., Couture, C., & collaborateurs. (1994). *Les fondements de la discipline archivistique*. Sainte-Foy, QC : Presses de l'Université du Québec.
- Sabourin, P. (2001). Constructing a function-based records classification system : Business activity structure classification system. *Archivaria*, 51, 137-154.
- Saffady, W. (2002). *Managing electronic records*. Prairie Village, Kansas : ARMA International.

- Salton, G. (1989). *Automatic text processing : The transformation, analysis, and retrieval of information by computer*. Reading, MA. : Addison-Wesley.
- Satija, M. P. (2000). Library classification : An essay in terminology. *Knowledge Organization*, 27 (4), 221-229.
- Saulnier, C. (1997-1998). Prolégomènes à la gestion des documents administratifs électroniques. *Archives (Revue de l'Association des archivistes du Québec)*, 29 (1), 57-75.
- Savic, D. (1995). Automatic classification of office documents : Review of available methods and techniques. *Records Management Quarterly*, 29 (4), 3-18.
- Sayers, W. C. B. (1915). *Canons of classification*. London, UK : Grafton & Co.
- Sayers, W. C. B., & Maltby, A. (1975). *Sayers' manual of classification for librarians*. London, UK : Deutsch.
- Schellenberg, T. R. (1956). *Modern archives principles and techniques*. Chicago, IL : University of Chicago Press.
- Schellenberg, T. R. (1965). *The management of archives*. New York, NY : Columbia University Press. (Columbia University studies in library service; 14).
- Sellen, A., & Harper, R. (2001). *The myth of the paperless office*. Cambridge, MA : MIT Press.
- Sénécal, S. (1997-1998). La lecture et la description archivistique du document. *Archives (Revue de l'Association des archivistes du Québec)*, 29 (3/4), 49-56.
- Sénécal, S. (1999). *La notion de classification et l'utilisation d'un plan de classification pour la gestion des documents dans un environnement électronique*. Montréal, QC : Hydro-Québec.
- Skupsky, D. S. (1999). Applying records retention to electronic records. *The Information Management Journal*, 33 (3), 28-35.
- Smiraglia, R. (2001). Further progress toward theory in knowledge organization. *Canadian Journal of Information and Library Science / Revue canadienne des sciences de l'information et de bibliothéconomie*, 26 (2/3), 31-50.
- Smiraglia, R. (2002). The progress of theory in knowledge organization. *Library Trends*, 50 (3), 330-349.

- Sprehe, T. J., McClure, C. R., & Zellner, P. (2002). The role of situational factors in managing U.S. federal recordkeeping. *Government Information Quarterly*, 19, 289-305.
- Suchman, L., & Wynn, E. (1984). Procedures and problems in the office. *Office Technology and People*, 2, 134-154.
- Sukiasyan, E. (1998). Classification systems in their historical development : Problems of typology and terminology. In *Proceedings of the 5th International ISKO Conference, August 25-29, 1998, Lille, France*. Würzburg, Germany : Ergon, 72-79.
- Svenonius, E. (1983). Use of classification in online retrieval. *Library Resources and Technical Services*, 27 (1), 76-80.
- Svenonius, E. (1997). Definitional approaches in the design of classification and thesauri and their implications for retrieval and for automatic classification. In *Proceedings of the 6th international study conference on classification research, June 16-18, 1997, The Hague, Netherlands*. The Hague, Netherlands : International Federation for Information and Documentation, 12-16.
- Thiétart, R.-A., & collaborateurs (2003). *Méthodes de recherche en management*. Paris : Dunod.
- Toms, E. G., & Campbell, D. G. (1999). Utilizing information "shape" as an interface metaphor based on genre. In *Proceedings of the 27th Annual Conference of the Canadian Association for Information Science, June 9-11, 1999, Université de Sherbrooke, Canada*, 370-386.
http://www.cais-acsi.ca/proceedings/1999/Toms_1999.pdf
- Tough, A., & Moss, M. (2003). Metadata, controlled vocabulary and directories : Electronic document management and standard for records management. *Records Management Journal*, 13 (1), 24-31.
- Trudel, R., & Antonius, R. (1991). *Méthodes quantitatives appliquées aux sciences humaines*. Montréal, QC : Centre éducatif et culturel.
- Université de Montréal. Service des archives. (1995). *Guide de gestion des archives informatiques de l'Université de Montréal : volet bureautique (environnements IBM et Macintosh)*. Montréal, QC : Presses de l'Université de Montréal.

- U.S. National Archives and Records Administration (NARA). (2001). *An overview of three projects relating to the changing federal recordkeeping environment : Report on current recordkeeping practices within the Federal Government*. SRA International.
http://www.archives.gov/records_management/initiatives/report_on_recordkeeping_practices.html
- Van der Walt, M. S. (1998). The structures of classification schemes used in Internet search engines. In *Proceedings of the 5th International ISKO Conference, August 25-29, 1998, Lille, France*. Würzburg, Germany : Ergon, 379-387.
- Van der Walt, M. S. (2004). A classification scheme for the organization of electronic documents in small, medium and micro enterprises (SMMEs). *Knowledge Organization*, 31 (1), 26-38.
- Vigneau, A. (1996). Les documents informatiques : pour une classification efficace. *Archives (Revue de l'Association des archivistes du Québec)*, 27 (3), 29-47.
- Waldron, M. (2002). The new business necessity. *E-doc*, 16 (2), 50-53.
- Walliser, B. (1977). *Systèmes et modèles : introduction critique à l'analyse de systèmes*. Paris : Seuil.
- Whittaker, S., & Hirschberg, J. (2001). The character, value and management of personal paper archives. *ACM Transactions on Computer Human Interaction*, 8 (2), 150-170.
<http://portal.acm.org/citation.cfm?id=376932>
- Whittaker, S., & Sidner, C. (1996). Email overload : Exploring personal information management of email. In *Proceedings of the SIGCHI conference on Human factors in computing systems, April 13-18, 1996, Vancouver, British Columbia, Canada*, 276-283.
<http://delivery.acm.org/10.1145/240000/238530/p276-whittaker.pdf>
- Whittaker, S., Terveen, L., & Nardi, B. (2000). Let's stop pushing the envelope & start addressing it : A reference task agenda for HCI. *Human Computer Interaction*, 15, 75-106.
- Zeller, J.-D. (2004). Documents numériques : à la recherche d'une typologie perdue. *Document numérique*, 8 (2), 101-116.
- Zins, C. (2002). Models for classifying Internet resources. *Knowledge Organisation*, 29 (1), 20-28.

Annexe 1. Certificat d'éthique et formulaires de consentement

FORMULAIRE DE CONSENTEMENT
[Entrevues avec les employés]

Titre du projet :

Schémas de classification et repérage des documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles

Responsable du projet (personne à contacter pour informations ou désistement)

Mas, Sabine, candidate au doctorat

Faculté des arts et des sciences

École de bibliothéconomie et sciences de l'information

C.P. 6128, succursale Centre-ville

Montréal QC H3C 3J7

Téléphone : (514) 343-6111 poste 3045

██

Nom de la directrice de recherche

Gagnon-Arguin, Louise, professeure associée

Nom de la co-directrice de recherche

Hudon, Michèle, professeure agrégée

Objectifs du projet : Cette recherche a pour but de vérifier l'existence d'un lien entre les différents schémas de classification (structures de répertoires) des documents électroniques produits ou reçus par le personnel d'un organisme dans le cadre de ses activités administratives et le repérage de documents électroniques. Plus précisément, cette recherche vise à identifier certains facteurs ou caractéristiques déterminantes à l'efficacité relative du repérage d'un sous-ensemble donné de documents électroniques. Par votre participation, vous contribuerez à une meilleure compréhension de la relation entre les schémas de classification et le repérage des documents administratifs électroniques.

Tâches à exécuter : Dans le cadre d'une entrevue avec le chercheur, vous aurez à décrire oralement la structure arborescente des répertoires ou dossiers électroniques que vous utilisez pour organiser vos documents électroniques et à répondre à diverses questions relatives à vos pratiques d'organisation (i.e., mise à jour de la structure) et d'utilisation (i.e., fréquence d'utilisation) de ces répertoires électroniques.

Durée : De 60 à 90 minutes.

Bénéfices prévus pour les répondants : Sur simple demande, nous vous transmettrons les résultats de cette recherche, une fois l'étude terminée.

Nature des données : Le chercheur effectuera des impressions écrans de la structure de répertoires électroniques créée dans le répertoire « Mes documents ». Une application sera également utilisée pour colliger des données statistiques relatives à cette structure (i.e., nombre total de répertoires). L’entrevue sera enregistrée à l’aide d’un magnétophone.

Caractère confidentiel des informations : Les données recueillies demeureront confidentielles. Un numéro de code vous sera attribué. Les informations recueillies seront donc codées et seule la chercheuse aura accès à la liste des noms et des numéros de codes, de même qu’aux informations recueillies. Les données recueillies seront conservées sous clés par la chercheuse. Seuls des résultats de groupes pourront être rendus publics.

Participation volontaire et retrait : Votre participation à cette étude est tout à fait volontaire. Vous êtes donc libre d’accepter ou de refuser d’y participer et vous pouvez vous retirer de l’étude en tout temps. De même, vous pouvez décider de ne pas répondre à une question.

Consentement du/ de la répondant-e : J’ai pris connaissance du but de l’étude. Je comprends que je peux cesser ma participation en tout temps. J’accepte de mon plein gré de participer à cette étude. Je reconnais avoir 18 ans ou plus.

_____	_____	_____
Nom du participant	Signature du participant	Date
_____	_____	_____
Nom de la chercheuse	Signature de la chercheuse	Date

Toute plainte relative à votre participation à cette recherche peut être adressée à l’ombudsman de l’Université de Montréal, au numéro de téléphone (514) 343-2100 ou à l’adresse de courriel suivante : ombudsman@umontreal.ca

FORMULAIRE DE CONSENTEMENT [Séances de simulation]

Titre du projet :

Schémas de classification et repérage des documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles

Responsable du projet (personne à contacter pour informations ou désistement) :

Mas, Sabine, candidate au doctorat

Faculté des arts et des sciences

École de bibliothéconomie et sciences de l'information

C.P. 6128, succursale Centre-ville

Montréal QC H3C 3J7

Téléphone : (514) 343-2245



Directrice de recherche :

Gagnon-Arguin, Louise, Ph.D., professeure associée

Co-directrice de recherche :

Hudon, Michèle, Ph.D., professeure agrégée

Objectifs du projet : Cette recherche a pour but de vérifier l'existence d'un lien entre les différents schémas de classification (structures de répertoires) des documents électroniques produits ou reçus par le personnel d'un organisme dans le cadre de ses activités administratives et le repérage de documents électroniques. Plus précisément, cette recherche vise à identifier certains facteurs ou caractéristiques déterminantes à l'efficacité relative du repérage d'un sous-ensemble donné de documents électroniques. Par votre participation, vous contribuerez à une meilleure compréhension de la relation entre les schémas de classification et le repérage des documents administratifs électroniques.

Tâches à exécuter : Durant cette étude, vous utiliserez un ordinateur portable pour repérer un sous-ensemble de vingt documents cibles préalablement décrits et organisés selon cinq structures de répertoires électroniques différentes. Nous colligerons également votre appréciation générale de chaque structure de répertoire utilisée pour le repérage.

Durée : 60 à 90 minutes.

Bénéfices prévus pour les participant(e)s : Sur simple demande, nous vous transmettrons les résultats de cette recherche, une fois l'étude terminée. Chaque participant(e) sera également admissible à un tirage au sort d'un certificat-cadeau de 50\$ chez Renaud-Bray (fin juin).

Nature des données : Les mesures suivantes seront colligées par la chercheure : temps pris pour repérer chaque document, nombre d'essais, taux de succès du repérage. Nous colligerons votre appréciation générale de chaque structure de répertoires utilisée pour le repérage. Nous vous demanderons également de répondre à un court questionnaire en vue de colliger des données de nature démographique.

Caractère confidentiel des informations : Les données recueillies demeureront confidentielles. Un numéro de code vous sera attribué. Les informations recueillies seront donc codées et seule la chercheure aura accès à la liste des noms et des numéros de codes, de même qu'aux informations recueillies. Les données recueillies seront conservées sous clés par la chercheure. Seuls des résultats de groupes pourront être rendus publics.

Participation volontaire et retrait : Votre participation à cette étude est tout à fait volontaire. Vous êtes donc libre d'accepter ou de refuser d'y participer et vous pouvez vous retirer de l'étude en tout temps. De même, vous pouvez décider de ne pas répondre à une question.

Consentement du/ de la participant-e : J'ai pris connaissance du but de l'étude. Je comprends que je peux cesser ma participation en tout temps. J'accepte de mon plein gré de participer à cette étude. Je reconnais avoir 18 ans ou plus.

_____	_____	_____
Nom du participant(e)	Signature du participant(e)	Date
_____	_____	_____
Nom de la chercheure	Signature de la chercheure	Date

Toute plainte relative à votre participation à cette recherche peut être adressée à l'ombudsman de l'Université de Montréal, au numéro de téléphone (514) 343-2100 ou à l'adresse de courriel suivante : ombudsman@umontreal.ca

Annexe 2. Lettre d'appel à participation (employés)

-----Message d'origine-----

Envoyé : [Date] [Heure]

À : [Nom de l'employé(e)]

Objet : Recherche doctorale sur l'organisation des documents électroniques en milieu de travail

Bonjour [Madame/Monsieur] [Nom de l'employé(e)],

J'effectue actuellement une recherche doctorale qui porte sur l'organisation des documents administratifs électroniques en milieu de travail. Un des objectifs de cette recherche est d'étudier la façon dont les individus organisent leurs documents électroniques sur leur poste de travail personnel. À cet effet, nous avons préparé un questionnaire afin d'explorer les pratiques d'organisation et de repérage des documents électroniques par certaines catégories d'employé(e)s de l'Université de Montréal et d'identifier parmi les répondants les personnes qui accepteraient de participer à une entrevue d'une heure et demie environ en vue de colliger des données descriptives sur la structure de répertoires mise en place.

Nous vous serons très reconnaissants si vous prenez le temps (20 à 30 minutes) de répondre à ce questionnaire et de nous le retourner si possible AVANT LE [Date]. Pour ce faire, vous pouvez imprimer le fichier ci-joint et nous le retourner dûment complété par courrier interne à l'adresse ci-dessous. Vous pouvez aussi remplir le questionnaire directement à l'écran, enregistrer le fichier (très important !) et nous le renvoyer par courriel à l'adresse suivante [redacted]. Soyez assurée que vos réponses demeureront strictement anonymes.

Votre participation est essentielle si nous voulons mieux comprendre la relation entre les structures de répertoires existantes et le repérage des documents électroniques. Ce que nous apprendrons nous aidera à formuler des recommandations pour l'organisation des documents électroniques tout en tenant compte à la fois des besoins institutionnels et des besoins individuels de différents groupes d'employés.

Nous vous remercions d'avance pour votre aimable participation à cette recherche et pour votre contribution à la formation des étudiants au doctorat en sciences de l'information de l'Université de Montréal.

Dans l'attente de vos nouvelles, je vous prie de recevoir mes salutations les meilleures.

Sabine Mas

--

Sabine Mas

Candidate au doctorat

Membre de la Commission des archives de l'Université de Montréal

École de bibliothéconomie et des sciences de l'information (EBSI) Faculté des arts et des sciences Université de Montréal Pavillon Lionel-Groulx Local C-2004

Adresse postale :

C.P. 6128, succursale Centre-ville

Montréal QC

H3C 3J7

Téléphone : (514) 343-2245

Télécopieur : (514) 343-5753

[redacted]

Annexe 3. Questionnaire adressé aux techniciennes à la coordination de travail de bureau

Note: Ce document a été envoyé par courriel aux techniciennes à la coordination de bureau avant les entrevues

Organisation et repérage des documents électroniques en milieu de travail

Enquête sur les pratiques de certaines catégories d'employés de l'Université de Montréal

Ce questionnaire fait partie d'une recherche doctorale sur l'organisation des documents électroniques en milieu de travail. Un des buts visés par cette recherche est d'étudier l'influence des structures de répertoires électroniques mises en place par différents groupes d'employés sur le repérage des documents électroniques.

À cet effet, nous avons préparé ce questionnaire afin d'explorer les pratiques d'organisation et de repérage des documents électroniques par certaines catégories d'employés et d'identifier parmi les répondants les personnes qui accepteraient de participer à une entrevue d'une heure et demie environ en vue de colliger des données descriptives sur la structure de répertoires mise en place. Si vous acceptez de nous accorder cette entrevue nous vous demandons de bien vouloir indiquer votre adresse de courriel à la fin du questionnaire afin de pouvoir communiquer avec vous. Votre participation est essentielle si nous voulons mieux comprendre la relation entre les structures de répertoires existantes et le repérage des documents électroniques. Ce que nous apprendrons nous aidera à formuler des recommandations pour l'organisation des documents électroniques tout en tenant compte à la fois des besoins institutionnels et individuels de différents groupes d'employés.

Nous vous serons très reconnaissants si vous prenez le temps (20 à 30 minutes) de répondre à ce questionnaire. Soyez assuré(e) que vos réponses demeureront strictement confidentielles.

Sabine Mas,
Candidate au doctorat
Membre de la Commission des archives de l'Université de Montréal

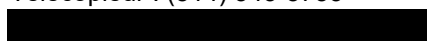
Louise Gagnon-Arguin, Ph.D. (directrice de recherche)
Professeure associée

Hudon Michèle, Ph.D. (co-directrice de recherche)
Professeure agrégée

École de bibliothéconomie et des sciences de l'information
Faculté des arts et des sciences
Université de Montréal
Pavillon Lionel-Groulx
Local C-2004

Adresse postale :

C.P. 6128, succursale Centre-ville
Montréal, QC
H3C 3J7
Téléphone : (514) 343-2245
Télécopieur : (514) 343-5753



SECTION 1. Informations générales

Q 1.1 Vous êtes :

Une femme

Un homme

Q 1.2 Quel est votre âge ?

18 - 25

26 - 34

35 - 49

50 et +

Q 1.3 Depuis combien d'années travaillez-vous à l'Université de Montréal ?

< 1 an

1 à 5 ans

6 à 10 ans

11 à 15 ans

16 ans et +

Q 1.4 Depuis combien d'années travaillez-vous dans ce département/faculté/école ?

< 1 an

1 à 5 ans

6 à 10 ans

11 à 15 ans

16 ans et +

Q 1.5 Quel est le titre du poste actuellement occupé ?

Q 1.6 Depuis combien d'années occupez-vous ce poste ?

< 1 an

1 à 5 ans

6 à 10 ans

11 à 15 ans

16 ans et +

Q 1.7 **Quels sont les énoncés qui correspondent à vos tâches ? (merci de cocher tous les énoncés qui s'appliquent)**

- Je participe à différentes opérations reliées à la gestion des études
- Je participe à différentes opérations reliées à la gestion des ressources humaines
- Je participe à différentes opérations reliées à la gestion des ressources financières
- Je participe à différentes opérations reliées à la gestion des ressources matérielles
- J'effectue le suivi des décisions prises par mon ou ma supérieure et prend les mesures appropriées afin d'en assurer la réalisation.
- Je prends les dispositions relatives à l'agenda de mon ou ma supérieure. À cette fin, je prends et fixe les rendez-vous, les lui rappelle et organise des rencontres selon les priorités.
- Je tiens à jour la comptabilité de l'unité. J'enregistre les dépenses effectuées et fait, sur demande, un rapport des états de compte. Je compare les écritures comptables avec les pièces justificatives sur réception périodique des relevés budgétaires.
- Je tiens à jour les dossiers du personnel et prépare les documents requis lors des affichages, des engagements, des promotions, des nominations, des affectations, des congés sabbatiques, des absences, etc.
- J'accomplis les tâches de la technicienne ou du technicien à la gestion des dossiers étudiants lorsque dans le secteur de travail il n'y a pas de poste de cette nature ou que la charge de travail de la technicienne ou du technicien à la gestion des dossiers étudiants le requiert.
- Je reçois et filtre les appels téléphoniques et donne des renseignements. J'accueille les visiteuses et les visiteurs, m'informe de l'objet de l'entretien et prépare les documents nécessaires. J'oriente ceux dont le problème peut être réglé par une autre autorité et fixe des rendez-vous si, à mon avis, il est nécessaire de tenir un entretien avec mon ou ma supérieure.
- Je prends connaissance du courrier et le trie en fonction des priorités. Je prépare les dossiers en faisant les recherches et les consultations préalables. Je réponds aux demandes relevant de ma compétence. Je rédige des projets de réponse complexes. Je note les suites à donner et effectue le suivi des dossiers.
- J'organise des réunions en faisant les convocations, les réservations des salles et la préparation du matériel et des documents nécessaires. J'assiste aux réunions, prends note des délibérations, rédige les procès-verbaux et donne suite aux décisions qui relèvent de ma compétence.
- J'assume la responsabilité du système de classement et applique la procédure d'archivage. J'effectue la mise à jour régulière des dossiers et des fichiers informatiques de gestion.
- Je prends des notes, dactylographie divers documents tels que lettres, rapports, mémos, etc. J'effectue la mise en page des documents et en vérifie l'orthographe et la grammaire. Je signe certains documents.
- J'utilise les logiciels en usage. Je tiens à jour les applications informatisées courantes et participe à leur développement afin d'améliorer l'efficacité des opérations.
- Je participe à la planification et à l'organisation matérielle d'activités spéciales telles que colloques, congrès, voyages.
- J'accomplis temporairement les tâches d'un poste connexe ou inférieur lorsque requis.
- Autres tâches (merci de préciser) :

SECTION 2. Compétences en informatique

Q 2.1 Depuis combien d'années utilisez-vous un ordinateur ?

< 1 an

1 - 5 ans

6 - 10 ans

> 10 ans

Q 2.2 Quel système d'exploitation utilisez-vous principalement dans votre milieu de travail ?

Windows

Mac OS

Unix

Autre système d'exploitation (merci de préciser) :

Je ne sais pas

Q 2.3 Parmi les descriptions des niveaux de compétence en informatique qui suivent, cochez l'énoncé qui correspond le mieux à votre niveau.

Novice : J'essaie de me servir de l'informatique mais j'ai souvent besoin d'aide.

Intermédiaire : Je peux exécuter des tâches simples avec un nombre limité de logiciels.

Expert : Je peux utiliser d'une façon compétente une grande variété de technologies informatiques.

SECTION 3. Formation en gestion des documents administratifs

La gestion des documents administratifs désigne l'ensemble des activités et méthodes relatives à la mise en place de systèmes d'administration des documents permettant d'organiser, de décrire, de conserver et de retrouver, sûrement et à tout moment, tout document produit ou reçu dans le cadre des activités d'un organisme.

Q 3.1 Avez-vous déjà reçu une formation en gestion des documents administratifs ?

Oui

Non

Si vous avez répondu "Non", merci d'aller à la [section 4](#), sinon merci de continuer.

Q 3.2 Si la réponse à la question 3.1 est « oui », veuillez indiquer les circonstances dans lesquelles vous avez reçu une formation (merci de cocher tous les énoncés qui s'appliquent)

Programme d'étude collégial

Nom du programme :

Nom du collège :

Dernière année de formation :

Programme d'étude universitaire

Nom du programme :

Nom de l'Université :

Dernière année de formation :

Formation donnée par la Division des archives de l'Université de Montréal

Dernière année de formation :

Autre (par exemple, cours, séminaires, etc.) (merci de préciser) :

Dernière année de formation :

Q 3.3 Si la réponse à la question 3.1 est « oui », est-ce que cette formation comprenait la classification des documents ?

- Oui
 Non
 Je ne m'en souviens pas

Q 3.4 Si la réponse à la question 3.3 est « oui », est-ce que cette formation comprenait la classification des "documents électroniques" ?

- Oui
 Non
 Je ne m'en souviens pas

SECTION 4. Organisation des documents électroniques

Dans le cadre de cette recherche nous étudions l'organisation des documents électroniques qui sont conservés à l'intérieur du répertoire « Mes Documents » ou son équivalent (par exemple, un espace personnel sur un serveur). Dans le répertoire « Mes Documents » (ou son équivalent) les documents électroniques sont le plus souvent organisés à l'intérieur de répertoires (ou dossiers) qui s'emboîtent les uns dans les autres pour former une **structure arborescente de répertoires**. Cette structure de répertoires est visible lorsque l'on ouvre la fenêtre de l'*Explorateur Windows*.

Certain(e)s employé(e)s de l'Université de Montréal utilisent le [plan de classification institutionnel](#) proposé par la Division des archives de l'Université de Montréal pour concevoir cette structure de répertoires. D'autres employé(e)s conçoivent une structure de répertoires plus personnalisée. Les questions qui suivent visent à mieux connaître votre propre structure de répertoires. Vos réponses et vos opinions nous sont précieuses afin de mieux comprendre les raisons de l'utilisation et de la non-utilisation du plan de classification institutionnel.

4.1 Utilisation du plan de classification institutionnel

Q 4.1.1 Utilisez-vous le plan de classification institutionnel conçu par la Division des archives de l'Université de Montréal pour organiser et nommer votre structure de répertoires se trouvant sur votre espace personnel de travail ?

- Oui, j'utilise **intégralement** le plan de classification institutionnel pour organiser mes documents électroniques
- Oui, j'utilise **en partie** le plan de classification institutionnel pour organiser mes documents électroniques
- Non, je n'utilise **pas du tout** le plan de classification institutionnel pour organiser mes documents électroniques
- Je ne sais pas

Si vous avez répondu "Non", merci d'aller à la [section 4.2](#), sinon merci de continuer.

Q 4.1.2 Pourquoi utilisez-vous le plan de classification institutionnel ? (merci de cocher tous les énoncés qui s'appliquent)

- Parce que le plan de classification était déjà installé sur l'ordinateur qui m'a été alloué
- Parce que l'utilisation du plan de classification a été rendue obligatoire par mon ou ma supérieure
- Pour faciliter l'organisation de mes documents électroniques
- Pour faciliter le repérage de mes documents électroniques
- Comme support à la gestion de mes dossiers
- Pour d'autre(s) raison(s) (merci de préciser) :

Q 4.1.3 Quelles catégories du plan de classification institutionnel utilisez-vous ? (merci de cocher autant de cases que nécessaire)

- Toutes les catégories
1. Comptabilité
2. Gestion et conservation des documents administratifs
3. Organisme décisionnel ou consultatif (U. de M.)
4. Courrier et télécommunications
5. Affiliation
6. Animalerie
7. Coopération et échange
8. Organisme externe
9. Bourse et prix
10. Activité sportive
11. Fournitures
12. Colloque, conférence et événement spécial
13. Organisation de la recherche

- 14. Financement
- 15. Association
- 16. Correspondance générale et communications diverses
- 17. Personnel enseignant
- 18. Activité communautaire
- 19. Textes et dossiers juridiques
- 20. Assurance
- 21. Construction et aménagement
- 22. Gestion des immeubles, terrains et véhicules
- 23. Personnel
- 24. Ressources informatiques
- 25. Mobilier, appareils et outillage
- 26. Publications et relations publiques
- 27. Cliniques et services professionnels
- 28. Organisation de l'enseignement
- 29. Vie étudiante
- 30. Ressources audiovisuelles
- 31. Étudiants - affaires scolaires
- 32. Organisation et régie interne
- 33. Projet de recherche
- 34. Statistique
- 35. Budget
- 36. Programme d'étude
- 37. Ressources documentaires et collections
- 38. Laboratoire
- 39. Organisation et politiques générales (U. De M.)

Commentaires détaillés et suggestions au sujet du plan de classification institutionnel

Q 4.1.4 Qu'aimez-vous dans le plan de classification institutionnel ?

Q 4.1.5 Quelles modifications aimeriez-vous voir apporter au plan de classification institutionnel ?

Q 4.1.6 En quoi ces modifications vous seraient-elles utiles ?

4.2 Questions UNIQUEMENT POUR CEUX QUI N'UTILISENT PAS le plan de classification institutionnel ou QUI L'UTILISENT EN PARTIE (si vous utilisez INTÉGRALEMENT le plan de classification institutionnel pour organiser vos documents électroniques merci d'aller à la [section 4.3](#))

Q 4.2.1 Qu'est-ce qui vous retient d'utiliser le plan de classification institutionnel ? (merci de cocher tous les énoncés qui s'appliquent)

- Je n'ai pas eu connaissance de cet outil
- Il est trop complexe/difficile à utiliser
- Il n'est pas assez précis
- Il est trop détaillé
- J'utilise déjà mon propre plan de classification
- Autre (merci de préciser) :

Q 4.2.2 Qu'est-ce qui vous motiverait à utiliser le plan de classification institutionnel ?

4.3 Stratégies personnelles d'organisation des documents électroniques

Dans cette sous-section, nous aimerions mieux connaître les stratégies que vous avez testées – certaines avec succès, d'autres non – pour organiser vos documents électroniques. Par exemple, vous avez peut-être essayé une nouvelle structure de répertoires. Peut-être que grâce à l'utilisation d'un moteur de recherche, vous avez été tenté(e) de laisser tous vos documents électroniques directement au premier niveau, c'est-à-dire directement classés sous le répertoire « Mes Documents » (ou son équivalent), etc.

Q 4.3.1 Décrivez une stratégie que vous avez essayée et qui **A BIEN FONCTIONNÉ** pour organiser vos documents électroniques.

Q 4.3.2 Décrivez une stratégie que vous avez essayée et qui **N'A PAS FONCTIONNÉ** pour organiser vos documents électroniques.

Q 4.3.3 Dans l'ensemble, comment qualifieriez-vous l'organisation de vos documents électroniques ?

Documents
extrêmement
désorganisés

Documents
très
désorganisés

Documents
moyennement
organisés

Documents
très bien
organisés

Documents
extrêmement
bien organisés

SECTION 5. Repérage des documents électroniques

Q 5.1 À quelle fréquence vous arrive-t-il de NE PAS être capable de retrouver un document électronique ?

- Jamais
- 1 - 2 fois par année
- Au moins 1 fois par session
- 1 - 2 fois par mois
- 1 - 2 fois par semaine
- Au moins une fois par jour
- Je ne sais pas

Q 5.2 À quelle fréquence vous arrive-t-il de remarquer qu'un document électronique a été placé dans le mauvais répertoire ?

- Jamais
- 1 - 2 fois par année
- Au moins 1 fois par session
- 1 - 2 fois par mois
- 1 - 2 fois par semaine
- Au moins une fois par jour
- Je ne sais pas

5.3 Stratégies personnelles de repérage des documents électroniques

Dans cette sous-section, nous aimerions mieux connaître les stratégies que vous avez testées – certaines avec succès, d'autres non – pour repérer vos documents électroniques. Par exemple, vous avez peut-être essayé une nouvelle façon de nommer vos répertoires. Peut-être que pour faciliter le repérage, vous avez été tenté(e) de classer tous vos documents électroniques utilisés couramment directement dans un répertoire particulier, etc.

Q 5.3.1 Décrivez une stratégie que vous avez essayée et qui A BIEN FONCTIONNÉ pour repérer vos documents électroniques.

Q 5.3.2 Décrivez une stratégie que vous avez essayée et qui N'A PAS FONCTIONNÉ pour repérer vos documents électroniques.

Q 5.3.3 Dans l'ensemble, comment qualifieriez-vous le repérage de vos documents électroniques ?

Repérage
extrêmement
difficile

Repérage
très
difficile

Repérage
moyennement
facile

Repérage
très
facile

Repérage
extrêmement
facile

Q 5.3.4 Identifiez ci-dessous 3 documents que vous jugez FACILES À REPÉRER par autrui

Nom du **premier document** (Ex.: TREMBLAY, Lise.doc) :

Description du document (Ex.: rapport de stage de Lise TREMBLAY réalisé à l'automne 2004) :

Localisation ou chemin d'accès du document (Ex.: Rapports de stages/A2004/TREMBLAY, Lise.doc) :

Nom du **deuxième document** :

Description du document :

Localisation ou chemin d'accès du document :

Nom du **troisième document** :

Description du document :

Localisation ou chemin d'accès du document :

Q 5.3.5 Identifiez ci-dessous 3 documents que vous jugez **DIFFICILES À REPÉRER** par autrui et expliquez brièvement pourquoi vous jugez chacun de ces documents difficiles à repérer par quelqu'un d'autre que vous-même

Nom du **premier document** :

Description du document :

Localisation ou chemin d'accès du document :

Raison(s) de la difficulté du repérage :

Nom du **deuxième document** :

Description du document :

Localisation ou chemin d'accès du document :

Raison(s) de la difficulté du repérage :

Nom du **troisième document** :

Description du document :

Localisation ou chemin d'accès du document :

Raison(s) de la difficulté du repérage :

5.4 Participation volontaire à une entrevue

Nous sommes à la recherche de techniciens ou techniciennes en coordination de travail de bureau de l'Université de Montréal qui accepteraient, dans le cadre d'une entrevue avec la chercheure, de décrire oralement la structure arborescente des répertoires ou dossiers électroniques qui est utilisée pour organiser leurs documents électroniques et de répondre à diverses questions relatives à leurs pratiques d'organisation (i.e., mise à jour de la structure) et d'utilisation (i.e., fréquence d'utilisation) de ces répertoires électroniques. La chercheure effectuera des impressions écrans de la structure de répertoires électroniques créée dans le répertoire « Mes documents ». Une application sera également utilisée pour colliger des données statistiques relatives à cette structure (i.e., nombre total de fichiers et de répertoires, nombre de sous-répertoires par répertoires, etc.). Grâce à ces entrevues, nous souhaitons mieux connaître les pratiques, les méthodes ainsi que les besoins des techniciens ou techniciennes en coordination de travail de bureau en matière d'organisation des documents électroniques afin de les aider, dans la mesure du possible, à faciliter la gestion de leurs documents électroniques.

Pour vraiment comprendre les pratiques d'organisation des documents électroniques, il est essentiel pour nous que ceux qui n'utilisent pas le plan de classification de l'Université de Montréal aussi bien que ceux qui l'utilisent participent à ces entrevues. Cette participation nous est précieuse en vue de nous aider à formuler des recommandations appropriées en ce qui concerne l'organisation des documents électroniques par les individus dans un contexte de travail.

Q 5.4.1 Acceptez-vous de nous accorder une entrevue ?

Non

Oui (merci d'indiquer les informations suivantes afin de nous permettre de communiquer avec vous) :

Numéro de téléphone (Travail) : (514) - poste

Adresse courriel :

S.V.P. Veuillez nous retourner ce formulaire AVANT LE 17 FÉVRIER 2006 par *courriel* [redacted] ou par *courrier interne* à l'adresse mentionnée ci-dessous :

Sabine Mas
École de bibliothéconomie et des sciences de l'information
Pavillon Lionel-Groulx
Local C-2004

Adresse postale :
C.P. 6128, succursale Centre-ville
Montréal QC
H3C 3J7

Merci pour votre aimable participation !

Annexe 4. Questionnaire adressé aux coordonnateurs de stages

Note: Ce document a été envoyé par courriel aux coordonnateurs de stages avant les entrevues

Organisation et repérage des documents électroniques en milieu de travail

Enquête sur les pratiques de certaines catégories d'employés de l'Université de Montréal

Ce questionnaire fait partie d'une recherche doctorale sur l'organisation des documents électroniques en milieu de travail. Un des buts visés par cette recherche est d'étudier l'influence des structures de répertoires électroniques mises en place par différents groupes d'employés sur le repérage des documents électroniques.

À cet effet, nous avons préparé ce questionnaire afin d'explorer les pratiques d'organisation et de repérage des documents électroniques par certaines catégories d'employés et d'identifier parmi les répondants les personnes qui accepteraient de participer à une entrevue d'une heure et demie environ en vue de colliger des données descriptives sur la structure de répertoires mise en place. Si vous acceptez de nous accorder cette entrevue nous vous demandons de bien vouloir indiquer votre adresse de courriel à la fin du questionnaire afin de pouvoir communiquer avec vous. Votre participation est essentielle si nous voulons mieux comprendre la relation entre les structures de répertoires existantes et le repérage des documents électroniques. Ce que nous apprendrons nous aidera à formuler des recommandations pour l'organisation des documents électroniques tout en tenant compte à la fois des besoins institutionnels et individuels de différents groupes d'employés.


Nous vous serons très reconnaissants si vous prenez le temps (20 à 30 minutes) de répondre à ce questionnaire. Soyez assuré(e) que vos réponses demeureront strictement confidentielles.

Sabine Mas,
Candidate au doctorat
Membre de la Commission des archives de l'Université de Montréal

Louise Gagnon-Arguin, Ph.D. (directrice de recherche)
Professeure associée

Hudon Michèle, Ph.D. (co-directrice de recherche)
Professeure agrégée

École de bibliothéconomie et des sciences de l'information
Faculté des arts et des sciences
Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal, QC
H3C 3J7
Téléphone : (514) 343-2245
Télécopieur : (514) 343-5753



SECTION 1. Informations générales

Q 1.1 Vous êtes :

Une femme

Un homme

Q 1.2 Quel est votre âge ?

18 - 25

26 - 34

35 - 49

50 et +

Q 1.3 Depuis combien d'années travaillez-vous à l'Université de Montréal ?

< 1 an

1 à 5 ans

6 à 10 ans

11 à 15 ans

16 ans et +

Q 1.4 Depuis combien d'années travaillez-vous dans ce département/faculté/école ?

< 1 an

1 à 5 ans

6 à 10 ans

11 à 15 ans

16 ans et +

Q 1.5 Quel est le titre du poste actuellement occupé ?

Q 1.6 Depuis combien d'années occupez-vous ce poste ?

< 1 an

1 à 5 ans

6 à 10 ans

11 à 15 ans

16 ans et +

Q 1.7 **Quels sont les énoncés qui correspondent à vos tâches ? (merci de cocher tous les énoncés qui s'appliquent)**

- J'informe les étudiant(e)s des objectifs et du déroulement des stages
- J'effectue ou je collabore à l'approbation des projets de stage
- J'effectue ou je collabore à l'évaluation des rapports de stage
- J'assure la coordination générale du stage
- J'assure l'encadrement de la démarche d'apprentissage du stagiaire et je rencontre chaque étudiant(e) individuellement, au besoin, afin de vérifier l'accomplissement des activités d'apprentissage et le progrès de l'étudiant(e)
- Je participe à l'évaluation finale du stage
- Je gère la banque de stages
- Je participe à la planification et à l'organisation matérielle d'activités spéciales telles que journée de l'emploi, salon des études
- Je participe à diverses activités de relations publiques
- Je coordonne ou participe à diverses publications (bulletins, dépliants, rapports, etc.)
- J'agis à titre de personne-ressource en matière d'emploi auprès des étudiant(e)s et du milieu professionnel en alimentant un babillard d'offres d'emploi spécialisé et une banque de CV d'étudiants ou de finissants
- Je collabore à la bonne marche de mon secteur en colligeant de l'information, en présentant sous les formes appropriées les données relatives à mon domaine et en suggérant des modifications susceptibles d'améliorer le fonctionnement des activités de mon secteur
- Autres tâches, merci de préciser

SECTION 2. Compétences en informatique

Q 2.1 Depuis combien d'années utilisez-vous un ordinateur ?

< 1 an

1 - 5 ans

6 - 10 ans

> 10 ans

Q 2.2 Quel système d'exploitation utilisez-vous principalement dans votre milieu de travail ?

Windows

Mac OS

Unix

Autre système d'exploitation (merci de préciser) :

Je ne sais pas

Q 2.3 Parmi les descriptions des niveaux de compétence en informatique qui suivent, cochez l'énoncé qui correspond le mieux à votre niveau.

Novice : J'essaie de me servir de l'informatique mais j'ai souvent besoin d'aide.

Intermédiaire : Je peux exécuter des tâches simples avec un nombre limité de logiciels.

Expert : Je peux utiliser d'une façon compétente une grande variété de technologies informatiques.

SECTION 3. Formation en gestion des documents administratifs

La gestion des documents administratifs désigne l'ensemble des activités et méthodes relatives à la mise en place de systèmes d'administration des documents permettant d'organiser, de décrire, de conserver et de retrouver, sûrement et à tout moment, tout document produit ou reçu dans le cadre des activités d'un organisme.

Q 3.1 Avez-vous déjà reçu une formation en gestion des documents administratifs ?

Oui

Non

Si vous avez répondu "Non", merci d'aller à la [section 4](#), sinon merci de continuer.

Q 3.2 Si la réponse à la question 3.1 est « oui », veuillez indiquer les circonstances dans lesquelles vous avez reçu une formation (merci de cocher tous les énoncés qui s'appliquent)

Programme d'étude collégial

Nom du programme :

Nom du collège :

Dernière année de formation :

Programme d'étude universitaire

Nom du programme :

Nom de l'Université :

Dernière année de formation :

Formation donnée par la Division des archives de l'Université de Montréal

Dernière année de formation :

Autre (par exemple, cours, séminaires, etc.) (merci de préciser) :

Dernière année de formation :

Q 3.3 Si la réponse à la question 3.1 est « oui », est-ce que cette formation comprenait la classification des documents ?

- Oui
- Non
- Je ne m'en souviens pas

Q 3.4 Si la réponse à la question 3.3 est « oui », est-ce que cette formation comprenait la classification des "documents électroniques" ?

- Oui
- Non
- Je ne m'en souviens pas

SECTION 4. Organisation des documents électroniques

Dans le cadre de cette recherche nous étudions l'organisation des documents électroniques qui sont conservés à l'intérieur du répertoire « Mes Documents » ou son équivalent (par exemple, un espace personnel sur un serveur). Dans le répertoire « Mes Documents » (ou son équivalent) les documents électroniques sont le plus souvent organisés à l'intérieur de répertoires (ou dossiers) qui s'emboîtent les uns dans les autres pour former une **structure arborescente de répertoires**. Cette structure de répertoires est visible lorsque l'on ouvre la fenêtre de l'*Explorateur Windows*.

Certain(e)s employé(e)s de l'Université de Montréal utilisent le [plan de classification institutionnel](#) proposé par la Division des archives de l'Université de Montréal pour concevoir cette structure de répertoires. D'autres employé(e)s conçoivent une structure de répertoires plus personnalisée. Les questions qui suivent visent à mieux connaître votre propre structure de répertoires. Vos réponses et vos opinions nous sont précieuses afin de mieux comprendre les raisons de l'utilisation et de la non-utilisation du plan de classification institutionnel.

4.1 Utilisation du plan de classification institutionnel

Q 4.1.1 Utilisez-vous le plan de classification institutionnel conçu par la Division des archives de l'Université de Montréal pour organiser et nommer votre structure de répertoires se trouvant sur votre espace personnel de travail ?

- Oui, j'utilise **intégralement** le plan de classification institutionnel pour organiser mes documents électroniques
- Oui, j'utilise **en partie** le plan de classification institutionnel pour organiser mes documents électroniques
- Non, je n'utilise **pas du tout** le plan de classification institutionnel pour organiser mes documents électroniques
- Je ne sais pas

Si vous avez répondu "Non", merci d'aller à la [section 4.2](#), sinon merci de continuer.

Q 4.1.2 Pourquoi utilisez-vous le plan de classification institutionnel ? (merci de cocher tous les énoncés qui s'appliquent)

- Parce que le plan de classification était déjà installé sur l'ordinateur qui m'a été alloué
- Parce que l'utilisation du plan de classification a été rendue obligatoire par mon ou ma supérieure
- Pour faciliter l'organisation de mes documents électroniques
- Pour faciliter le repérage de mes documents électroniques
- Comme support à la gestion de mes dossiers

Pour d'autre(s) raison(s) (merci de préciser) :

Q 4.1.3 **Quelles catégories du plan de classification institutionnel utilisez-vous ? (merci de cocher autant de cases que nécessaire)**

- Toutes les catégories
- 1. Comptabilité
- 2. Gestion et conservation des documents administratifs
- 3. Organisme décisionnel ou consultatif (U. de M.)
- 4. Courrier et télécommunications
- 5. Affiliation
- 6. Animalerie
- 7. Coopération et échange
- 8. Organisme externe
- 9. Bourse et prix
- 10. Activité sportive
- 11. Fournitures
- 12. Colloque, conférence et événement spécial
- 13. Organisation de la recherche
- 14. Financement
- 15. Association
- 16. Correspondance générale et communications diverses
- 17. Personnel enseignant
- 18. Activité communautaire
- 19. Textes et dossiers juridiques
- 20. Assurance
- 21. Construction et aménagement
- 22. Gestion des immeubles, terrains et véhicules
- 23. Personnel
- 24. Ressources informatiques
- 25. Mobilier, appareils et outillage
- 26. Publications et relations publiques

- 27. Cliniques et services professionnels
- 28. Organisation de l'enseignement
- 29. Vie étudiante
- 30. Ressources audiovisuelles
- 31. Étudiants - affaires scolaires
- 32. Organisation et régie interne
- 33. Projet de recherche
- 34. Statistique
- 35. Budget
- 36. Programme d'étude
- 37. Ressources documentaires et collections
- 38. Laboratoire
- 39. Organisation et politiques générales (U. De M.)

Commentaires détaillés et suggestions au sujet du plan de classification institutionnel

Q 4.1.4 Qu'aimez-vous dans le plan de classification institutionnel ?

Q 4.1.5 Quelles modifications aimeriez-vous voir apporter au plan de classification institutionnel ?

Q 4.1.6 En quoi ces modifications vous seraient-elles utiles ?

4.2 Questions UNIQUEMENT POUR CEUX QUI N'UTILISENT PAS le plan de classification institutionnel ou QUI L'UTILISENT EN PARTIE (si vous utilisez INTÉGRALEMENT le plan de classification institutionnel pour organiser vos documents électroniques merci d'aller à la [section 4.3](#))

Q 4.2.1 Qu'est-ce qui vous retient d'utiliser le plan de classification institutionnel ? (merci de cocher tous les énoncés qui s'appliquent)

- Je n'ai pas eu connaissance de cet outil
- Il est trop complexe/difficile à utiliser
- Il n'est pas assez précis
- Il est trop détaillé
- J'utilise déjà mon propre plan de classification
- Autre (merci de préciser) :

Q 4.2.2 Qu'est-ce qui vous motiverait à utiliser le plan de classification institutionnel ?

4.3 Stratégies personnelles d'organisation des documents électroniques

Dans cette sous-section, nous aimerions mieux connaître les stratégies que vous avez testées – certaines avec succès, d'autres non – pour organiser vos documents électroniques. Par exemple, vous avez peut-être essayé une nouvelle structure de répertoires. Peut-être que grâce à l'utilisation d'un moteur de recherche, vous avez été tenté(e) de laisser tous vos documents électroniques directement au premier niveau, c'est-à-dire directement classés sous le répertoire « Mes Documents » (ou son équivalent), etc.

Q 4.3.1 Décrivez une stratégie que vous avez essayée et qui **A BIEN FONCTIONNÉ** pour organiser vos documents électroniques.

Q 4.3.2 Décrivez une stratégie que vous avez essayée et qui **N'A PAS FONCTIONNÉ** pour organiser vos documents électroniques.

Q 4.3.3 Dans l'ensemble, comment qualifieriez-vous l'organisation de vos documents électroniques ?

Documents
extrêmement
désorganisés

Documents
très
désorganisés

Documents
moyennement
organisés

Documents
très bien
organisés

Documents
extrêmement
bien organisés

SECTION 5. Repérage des documents électroniques

Q 5.1 À quelle fréquence vous arrive-t-il de NE PAS être capable de retrouver un document électronique ?

- Jamais
- 1 - 2 fois par année
- Au moins 1 fois par session
- 1 - 2 fois par mois
- 1 - 2 fois par semaine
- Au moins une fois par jour
- Je ne sais pas

Q 5.2 À quelle fréquence vous arrive-t-il de remarquer qu'un document électronique a été placé dans le mauvais répertoire ?

- Jamais
- 1 - 2 fois par année
- Au moins 1 fois par session
- 1 - 2 fois par mois
- 1 - 2 fois par semaine
- Au moins une fois par jour
- Je ne sais pas

5.3 Stratégies personnelles de repérage des documents électroniques

Dans cette sous-section, nous aimerions mieux connaître les stratégies que vous avez testées – certaines avec succès, d'autres non – pour repérer vos documents électroniques. Par exemple, vous avez peut-être essayé une nouvelle façon de nommer vos répertoires. Peut-être que pour faciliter le repérage, vous avez été tenté(e) de classer tous vos documents électroniques utilisés couramment directement dans un répertoire particulier, etc.

Q 5.3.1 Décrivez une stratégie que vous avez essayée et qui **A BIEN FONCTIONNÉ** pour repérer vos documents électroniques.

Q 5.3.2 Décrivez une stratégie que vous avez essayée et qui **N'A PAS FONCTIONNÉ** pour repérer vos documents électroniques.

Q 5.3.3 Dans l'ensemble, comment qualifieriez-vous le repérage de vos documents électroniques ?

Repérage
extrêmement
difficile

Repérage
très
difficile

Repérage
moyennement
facile

Repérage
très
facile

Repérage
extrêmement
facile

Q 5.3.4 Identifiez ci-dessous 3 documents que vous jugez FACILES À REPÉRER par autrui

Nom du **premier document** (Ex.: TREMBLAY, Lise.doc) :

Description du document (Ex.: rapport de stage de Lise TREMBLAY réalisé à l'automne 2004) :

Localisation ou chemin d'accès du document (Ex.: Rapports de stages/A2004/TREMBLAY, Lise.doc) :

Nom du **deuxième document** :

Description du document :

Localisation ou chemin d'accès du document :

Nom du **troisième document** :

Description du document :

Localisation ou chemin d'accès du document :

Q 5.3.5 Identifiez ci-dessous 3 documents que vous jugez **DIFFICILES À REPÉRER** par autrui et expliquez brièvement pourquoi vous jugez chacun de ces documents difficiles à repérer par quelqu'un d'autre que vous-même

Nom du **premier document** :

Description du document :

Localisation ou chemin d'accès du document :

Raison(s) de la difficulté du repérage :

Nom du **deuxième document** :

Description du document :

Localisation ou chemin d'accès du document :

Raison(s) de la difficulté du repérage :

Nom du **troisième document** :

Description du document :

Localisation ou chemin d'accès du document :

Raison(s) de la difficulté du repérage :

5.4 Participation volontaire à une entrevue

Nous sommes à la recherche de coordonnateurs ou coordonnatrices de stages de l'Université de Montréal qui accepteraient, dans le cadre d'une entrevue avec la chercheuse, de décrire oralement la structure arborescente des répertoires ou dossiers électroniques qui est utilisée pour organiser leurs documents électroniques et de répondre à diverses questions relatives à leurs pratiques d'organisation (i.e., mise à jour de la structure) et d'utilisation (i.e., fréquence d'utilisation) de ces répertoires électroniques. La chercheuse effectuera des impressions écrans de la structure de répertoires électroniques créée dans le répertoire « Mes documents ». Une application sera également utilisée pour colliger des données statistiques relatives à cette structure (i.e., nombre total de fichiers et de répertoires, nombre de sous-répertoires par répertoires, etc.). Grâce à ces entrevues, nous souhaitons mieux connaître les pratiques, les méthodes ainsi que les besoins des coordonnateurs ou coordonnatrices de stages en matière d'organisation des documents électroniques afin de les aider, dans la mesure du possible, à faciliter la gestion de leurs documents électroniques.

Pour vraiment comprendre les pratiques d'organisation des documents électroniques, il est essentiel pour nous que ceux qui n'utilisent pas le plan de classification de l'Université de Montréal aussi bien que ceux qui l'utilisent participent à ces entrevues. Cette participation nous est précieuse en vue de nous aider à formuler des recommandations appropriées en ce qui concerne l'organisation des documents électroniques par les individus dans un contexte de travail.

Q 5.4.1 Acceptez-vous de nous accorder une entrevue ?

Non

Oui (merci d'indiquer les informations suivantes afin de nous permettre de communiquer avec vous) :

Numéro de téléphone (Travail) : (514) - poste

Adresse courriel :

S.V.P. Veuillez nous retourner ce formulaire AVANT LE 17 FÉVRIER 2006 par *courriel* [REDACTED] [REDACTED] ou par *courrier interne* à l'adresse mentionnée ci-dessous :

Sabine Mas
École de bibliothéconomie et des sciences de l'information
Faculté des arts et des sciences
Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal QC
H3C 3J7

Merci pour votre aimable participation !

Annexe 5. Classes du Système officiel de classification de l'Université de Montréal utilisées par quatre techniciennes

	TCTB02	TCTB03	TCTB05	TCTB10	Total
1.Comptabilité	x				1
2.Gestionetconservationdesdocumentsadministratifs	x	x			2
3.Organismedécisionnelouconsultatif(U.deM.)	x	x		x	3
4.Courrieretélécommunications	x	x		x	3
5.Affiliation		x		x	2
6.Animalerie					0
7.Coopérationetéchange		x		x	2
8.Organismeexterne	x			x	2
9.Bourseetprix	x	x		x	3
10.Activitésportive					0
11.Fournitures		x			1
12.Colloque,conférenceetévenementspécial	x	x	x	x	4
13.Organisationdelarecherche	x	x		x	3
14.Financement		x		x	2
15.Association				x	1
16.Correspondancegénéraleetcommunicationsdiverses		x	x	x	3
17.Personnelenseignant	x	x	x	x	4
18.Activitécommunautaire					0
19.Textesetdossiersjuridiques	x				1
20.Assurance	x				1
21.Constructionetaménagement	x				1
22.Gestiondesimmeubles,terrainsetvéhicules	x				1
23.Personnel	x	x	x	x	4
24.Ressourcesinformatiques	x	x		x	3
25.Mobilier,appareilsetoutillage	x	x			2
26.Publicationsetrelationspubliques				x	1
27.Cliniquesetservicesprofessionnels		x			1
28.Organisationdel'enseignement	x			x	2
29.Vieétudiante				x	1
30.Ressourcesaudiovisuelles	x	x			2
31.Étudiants-affairescolaires				x	1
32.Organisationetrégieinterne	x	x		x	3
33.Projetderecherche				x	1
34.Statistique					0
35.Budget	x	x	x	x	4
36.Programmed'étude				x	1
37.Ressourcesdocumentairesetcollections		x			1
38.Laboratoire	x			x	2
39.Organisationetpolitiquesgénérales(U.deM.)					0
Total des classes principales utilisées	21	20	5	23	

Annexe 6. Protocole d'entrevue

Objectif de cet outil : Ce guide a été conçu afin de rappeler à la chercheure quelle tâche elle doit être amenée à faire et quelles questions doivent être posées à un moment ou l'autre au cours de l'entrevue.

1. DOCUMENTS & MATÉRIEL À APPORTER

- 1 protocole d'entrevue
- 2 formulaires de consentement (photocopiés recto/verso)
- 1 feuille de sélection aléatoire des documents
- 1 table des nombres aléatoires
- 1 enregistreur numérique + piles
- clé USB contenant un répertoire portant le code de répondant (Ex.: CS01, TCTB02). Ce répertoire contient:
 1. Un fichier Word vide pour les captures d'écran dont le nom possède la syntaxe suivante : «PTCS#_ImpressionEcran.doc»
En-tête: code de répondant – Dated'entrevue, Pied de page: numérotation des pages
 2. Le logiciel PDS

2. FORMULAIRE DE CONSENTEMENT

- Décrire le projet et le déroulement de l'entrevue
- Lire ensemble le formulaire de consentement
- Questions?
- Signer les deux exemplaires du formulaire (l'employé en conserve une copie)

3. INSTALLATION DU MATÉRIEL

- Installer l'enregistreur numérique
- Tester l'enregistreur numérique (Note: mettre le volume suffisamment fort, rapprocher le micro du répondant)
- Mettre en marche l'enregistreur
- Insérer la clé USB dans l'ordinateur de l'employé

4. DESCRIPTION DES TÂCHES

- Demander à l'employé s'il possède une description écrite de ses tâches et s'il peut en laisser une copie
- Inviter l'employé à résumer ses principales fonctions et tâches

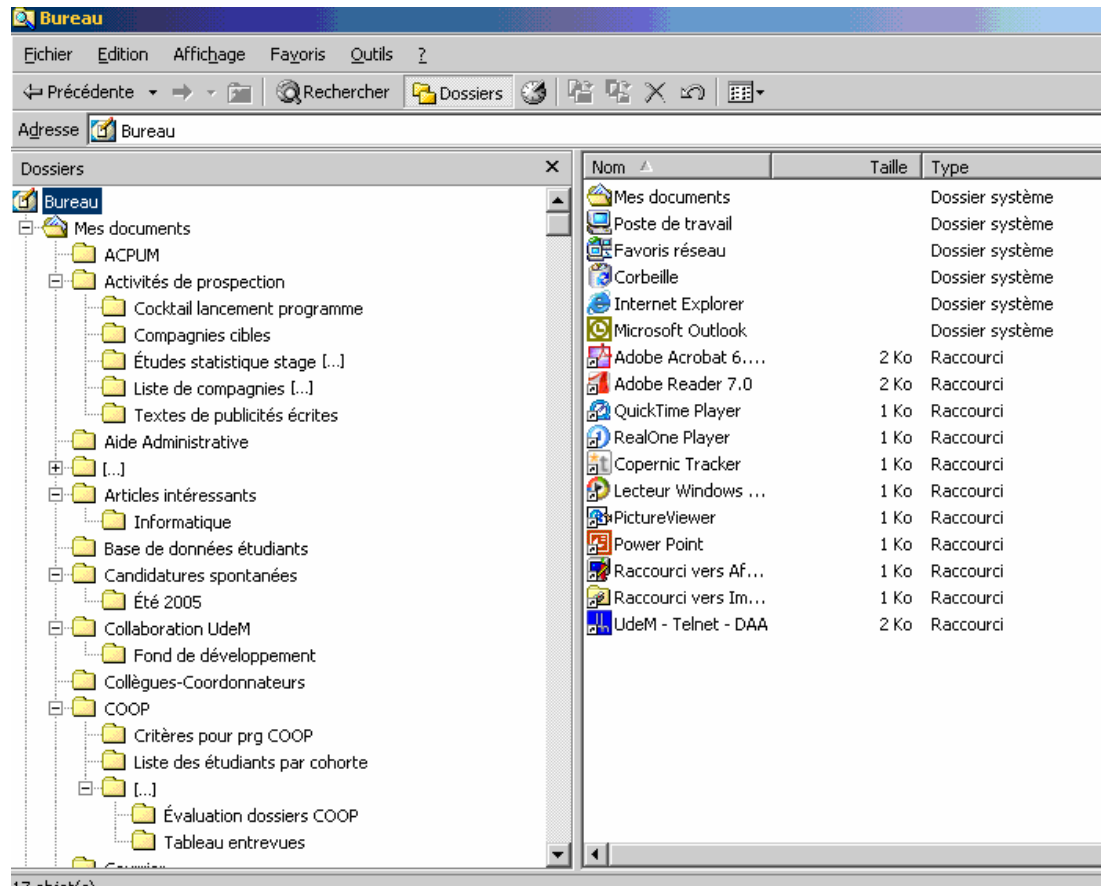
7. TÂCHES À EFFECTUER PAR L'EMPLOYÉ

1. À partir de l'Explorateur Windows, demander à l'employé de décrire brièvement le contenu de chaque répertoire
2. Sélectionner 20 documents aléatoirement dans l'arborescence des répertoires. L'employé doit décrire chacun de ces documents.

8. FIN DE L'ENTREVUE

- Remercier le répondant
- Faire une copie de sécurité des fichiers numériques

Annexe 7. Exemple de saisie d'écran



Annexe 8. Extrait d'un rapport produit par le logiciel *PDS*

```
#####
                          Analysing Locus 1
#####

#####
                          TOTAL LOCUS STATS
#####

Number of Files: 1376
Number of subDirs: 196
Branching Factor: 2.614
Tree Skewness: 6.901
Branching Deciles: [1, 1, 1, 1, 2, 2, 3, 5, 8, 13, 42]

Number of Files by Type:
  2002-2003 - msc - coord - 1 (0.07%)
  11 nov 2004doc - 1 (0.07%)
  2004-2005doc - 1 (0.07%)
  doc - 931 (67.66%)
  htm - 117 (8.50%)
  jpg - 4 (0.29%)
  pdf - 102 (7.41%)
  pps - 1 (0.07%)
  ppt - 43 (3.13%)
  pst - 1 (0.07%)
  rtf - 6 (0.44%)
  tmp - 7 (0.51%)
  txt - 11 (0.80%)
  xls - 150 (10.90%)

Number of Files by Class:
  Unknown - 4 (0.29%)
  image - 4 (0.29%)
```

Annexe 9. Caractéristiques des schémas de classification personnels observés

Tableau 9.1 Caractéristiques structurelles des schémas de classification des techniciennes à la coordination de travail de bureau

	<i>TCTB01</i>	<i>TCTB02p</i>	<i>TCTB03p</i>	<i>TCTB04p</i>	<i>TCTB05</i>	<i>TCTB06</i>	<i>TCTB07p</i>	<i>TCTB08</i>	<i>TCTB09</i>	<i>TCTB10</i>	<i>TCTB11p</i>	<i>Moyenne</i>	<i>Écart-type</i>
Nombretotaldefichiers	1617	1101	1962	2355	1696	1324	2680	5739	n/a	20857	2529	4186	6001.43
Nombretotaldeclasses	185	214	208	384	121	201	181	235	106	3541	387	523.91	1004.70
Nombredeclasses principales	22	21	105	16	29	11	29	7	41	42	14	30.64	27.16
Nombredeniveaux maximum	6	6	5	5	3	4	4	5	3	6	3	4.55	1.21
Nombredeniveaux minimum	1	1	1	1	1	1	1	1	1	1	1	1.00	0.00
Nombremoyendeniveaux	2.50	2.19	1.27	1.81	1.39	1.69	1.76	1.83	1.48	1.54	1.58	1.73	0.36
Nombremoyendeclassesà chaqueniveau	1.89	2.66	2.64	3.03	4.57	1.62	3.36	1.78	4.35	2.33	1.92	2.74	1.01
Écart-typedelamoyennedunombredeclassesàchaque niveau	4.05	3.20	16.22	9.12	8.43	3.35	4.44	9.90	7.15	8.37	5.09	7.21	3.85

p : Sc hémadeclassificationpersonnelpartagéavecdescollèguesdetravail.

Tableau 9.2 Caractéristiques structurelles des schémas de classification des coordonnateurs de stage

	<i>CS01</i>	<i>CS02</i>	<i>CS03</i>	<i>CS03p</i>	<i>CS04</i>	<i>CS05</i>	<i>CS06</i>	<i>CS07</i>	<i>CS07p</i>	<i>CS08</i>	<i>Moyenne</i>	<i>Écart-type</i>
Nombretotaldefichiers	1289	n/a	1376	661	261	976	880	2891	n/a	741	1134.38	792.77
Nombretotaldeclasses	167	13	196	131	20	210	164	618	71	103	169.30	171.91
Nombredeclasses principales	56	13	42	11	20	6	18	54	16	18	25.40	18.24
Nombredeniveaux maximum	4	1	5	6	1	6	6	6	4	3	4.20	1.99
Nombredeniveaux minimum	1	1	1	1	1	2	1	1	1	1	1.10	0.32
Nombremoyendeniveaux	1.29	1	1.56	1.54	1	2.86	1.94	1.59	2.06	1.25	1.61	0.56
Nombremoyendeclassesà chaqueniveau	3.29	n/a	2.61	2.01	n/a	2.2	2.1	2.7	2.58	4.31	2.73	0.76
Écart-typedelamoyenne dunombredeclassesà chaqueniveau	10.5	n/a	6.9	2.27	n/a	3.04	3.04	11.46	2.87	6.15	5.78	3.62

p : Schémadeclassificationpersonnelpartagéavecdescollèguesdetravail.

Tableau 9.3 Synthèse des caractéristiques structurelles des schémas de classification personnels observés

	<i>Moyenne</i>	<i>Écart-type</i>
Nombretotaldefichiers	2829.72	4664.90
Nombretotalde classes	355.05	742.25
Nombredeclasses principales	28.14	22.93
Nombredeniveauxmaximum	4.38	1.60
Nombredeniveauxminimum	1.05	0.22
Nombremoyendeniveaux	1.67	0.46
Nombremoyendeclasses àchaqueniveau	2.73	0.89
Écart-typedelamoyennedunombre declassesàchaque niveau	6.61	3.72

Tableau 9.4 Caractéristiques logiques des schémas de classification des techniciennes à la coordination de travail de bureau

	<i>TCTB01</i>	<i>TCTB02p</i>	<i>TCTB03p</i>	<i>TCTB04p</i>	<i>TCTB05</i>	<i>TCTB06</i>	<i>TCTB07p</i>	<i>TCTB08</i>	<i>TCTB09</i>	<i>TCTB10</i>	<i>TCTB11p</i>
1 ^{er} critère de division (1)	Type de contenu (32%) (n=7)	Thème (52%) (n=11)	Thème (32%) (n=34)	Type de contenu (44%) (n=7)	Thème (31%) (n=9)	Mélange (36%) (n=4)	Type de contenu (36%) (n=10)	Type de contenu (71%) (n=5)	Mélange (40%) (n=14)	Thème (50%) (n=21)	Thème (43%) (n=6)
2 ^{ème} critère de division (2)	Thème (29%) (n=4)	Thème (36%) (n=24)	Temps (32%) (n=18)	Temps (27%) (n=12)	Contact (61%) (n=43)	Mélange (38%) (n=10)	Temps (50%) (n=27)	Mélange (34%) (n=38)	Temps (45%) (n=22)	Contact Thème (33%) (n=27)	Mélange (45%) (n=9)
Ordre de succession des classes (1)	Alphabétique (95%) (n=21)	Numérique (100%) (n=21)	Numérique (100%) (n=105)	Alphabétique (100%) (n=16)	Alphabétique (100%) (n=29)	Alphabétique (100%) (n=11)	Alphabétique (100%) (n=28)	Alphabétique (100%) (n=7)	Alphabétique (100%) (n=35)	Numérique (95%) (n=40)	Alphabétique (93%) (n=13)
Ordre de Succession des classes (2)	Alphabétique (100%) (n=14)	Alphabétique (100%) (n=66)	Alphabétique (86%) (n=49)	Alphabétique (71%) (n=31)	Alphabétique (96%) (n=68)	Alphabétique (100%) (n=26)	Alphabétique (91%) (n=49)	Alphabétique (100%) (n=112)	Alphabétique (55%) (n=27)	Alphabétique (96%) (n=79)	Alphabétique (90%) (n=18)

(1): Proportion(%) et nombre(n) de classes principales (1^{er} niveau hiérarchique) de chacun des schémas de classification.

(2): Proportion(%) et nombre(n) de classes subordonnées (2^{ème} niveau hiérarchique) de chacun des schémas de classification.

p: Schéma de classification personnel par catégorie avec des collègues de travail.

Tableau 9.5 Caractéristiques logiques des schémas de classification des coordonnateurs de stage

	<i>CS01</i>	<i>CS02</i>	<i>CS03</i>	<i>CS03p</i>	<i>CS04</i>	<i>CS05</i>	<i>CS06</i>	<i>CS07</i>	<i>CS07p</i>	<i>CS08</i>
1 ^{er} critère de division(1)	Thème Type de contenu (27%) (n=13)	Thème (50%) (n=6)	Thème (40%) (n=17)	Mélange (27%) (n=3)	Thème (45%) (n=9)	Activité (67%) (n=4)	Thème (55%) (n=6)	Mélange (31%) (n=15)	Mélange (87%) (n=14)	Mélange (61%) (n=11)
2 ^{ème} critère de division(2)	Thème (30%) (n=13)	n/a	Mélange (46%) (n=33)	Mélange (70%) (n=7)	n/a	Activité (27%) (n=8)	Thème (35%) (n=16)	Contact (48%) (n=99)	Contact Thème (42%) (n=10)	Temps (58%) (n=11)
Ordre de succession des classes(1)	Alphabétique (100%) (n=49)	Alphabétique (100%) (n=12)	Alphabétique (100%) (n=42)	Alphabétique (82%) (n=9)	Alphabétique (100%) (n=20)	Alphabétique (100%) (n=6)	Numérique (100%) (n=11)	Alphabétique (100%) (n=49)	Alphabétique (100%) (n=16)	Alphabétique (100%) (n=18)
Ordre de succession des classes(2)	Alphabétique (100%) (n=44)	n/a	Alphabétique (92%) (n=66)	Alphabétique (100%) (n=10)	n/a	Alphabétique (100%) (n=30)	Numérique (89%) (n=41)	Alphabétique (100%) (n=205)	Alphabétique (92%) (n=22)	Alphabétique (84%) (n=16)

(1): Proportion(%) et nombre(n) de classes principales(1^{er} niveau hiérarchique) de chacun des schémas de classification.

(2): Proportion(%) et nombre(n) de classes subordonnées(2^{ème} niveau hiérarchique) de chacun des schémas de classification.

p : Schéma de classification personnel par tagé avec des collègues de travail.

Tableau 9.6 Synthèse des caractéristiques logiques des schémas de classification personnels observés

	<i>Moyenne</i>
1 ^{er} critère de division(a)	Thème (47.61%) (n=10)
2 ^{ème} critère de division(b)	Aucun critère ne prédomine
Ordre de succession des classes(a)	Alphabétique (80.95%) (n=17)
Ordre de succession des classes(b)	Alphabétique (85.71%) (n=18)

(a): Proportion(%) et nombre(n) de schémas de classification qui privilégient ce critère de division/ordre de succession des

classes au 1^{er} niveau hiérarchique.

(b): Proportion(%) et nombre(n) de schémas de classification qui privilégient ce critère de division/ordre de succession des

classes au 2^{ème} niveau hiérarchique.

Tableau 9.7 Caractéristiques sémantiques des schémas de classification des techniciennes à la coordination de travail de bureau

	<i>TCTB01</i>	<i>TCTB02p</i>	<i>TCTB03p</i>	<i>TCTB04p</i>	<i>TCTB05</i>	<i>TCTB06</i>	<i>TCTB07p</i>	<i>TCTB08</i>	<i>TCTB09</i>	<i>TCTB10</i>	<i>TCTB11p</i>
Longueur moyenne (en nombre de caractères) des noms de classe	13.29	22.41	21.75	10.50	12.70	19.05	17.25	13.05	11.66	17.90	16.32
Proportion de classes abrégées	41	28	27	55	27	56	56	34	34	39	49
Stratégie d'abréviation privilegiée (tous niveaux confondus)	Mélange (15%) (n=19)	Acronyme (18%) (n=36)	Acronyme (16%) (n=34)	Code (40%) (n=108)	Troncature (11%) (n=13)	Troncature (41%) (n=24)	Troncature (41%) (n=37)	Acronyme (12%) (n=23)	Acronyme (13%) (n=13)	Acronyme (22%) (n=41)	Acronyme (32%) (n=15)
Stratégie d'abréviation privilegiée (1 ^{er} niveau) (1)	Troncature (9%) (n=2)	n/a	Acronyme (21%) (n=22)	Acronyme (6%) (n=1)	Acronyme (7%) (n=2)	Troncature (55%) (n=6)	Acronyme (25%) (n=7)	Juxtaposition Mélange (14%)(n=1)	Contraction (17%) (n=6)	Mélange (26%) (n=11)	Acronyme Code Mélange (7%)(n=1)
Stratégie d'abréviation privilegiée (2 ^{ème} niveau) (2)	Troncature (21%) (n=3)	Troncature (18%) (n=12)	Acronyme (7%) (n=4)	Acronyme Code Troncature (2%)(n=1)	Troncature (18%) (n=13)	Acronyme (8%) (n=2)	Acronyme (24%) (n=13)	Acronyme (12%) (n=13)	Acronyme (16%) (n=8)	Acronyme (33%) (n=27)	Troncature (20%) (n=4)
Proportion (%) et nombre (n) de classes redondantes (tous niveaux confondus)	8.53% (n=11)	32.82% (n=64)	25.12% (n=52)	5.56% (n=15)	12.50% (n=15)	13.56% (n=8)	61.45% (n=110)	37.70% (n=72)	16% (n=16)	7.41% (n=14)	38.30% (n=18)
Proportion (%) et nombre (n) de classes redondantes (1 ^{er} niveau)	0%	42.86% (n=9)	27.62% (n=29)	12.50% (n=2)	13.79% (n=4)	0%	0%	0%	25.71% (n=9)	0%	0%

(1): Proportion (%) et nombre (n) de classes principales (1^{er} niveau hiérarchique) de chacun des schémas de classification.

(2): Proportion (%) et nombre (n) de classes subordonnées (2^{ème} niveau hiérarchique) de chacun des schémas de classification.

p : Schéma de classification personnel par tagé avec des collègues de travail.

Tableau 9.8 Caractéristiques sémantiques des schémas de classification des coordonnateurs de stage

	<i>CS01</i>	<i>CS02</i>	<i>CS03</i>	<i>CS03p</i>	<i>CS04</i>	<i>CS05</i>	<i>CS06</i>	<i>CS07</i>	<i>CS07p</i>	<i>CS08</i>
Longueurmoyenne (ennombredecaractères) desnomsdeclasse	17.69	15.33	19.12	23.55	13.83	13.10	15.81	13.26	13.79	17.13
Proportiondeclasses abrégées	19	42	32	46	85	12	18	21	17	40
Stratégied'abréviation privilégiée(tousniveaux confondus)	Troncature (5%) (n=6)	Acronyme (17%) (n=2)	Acronyme (10%) (n=20)	Code (26%) (n=27)	Code Mélange (40%)(n=8)	Contraction (5%) (n=11)	Mélange (7%) (n=11)	Acronyme (12%) (n=44)	Code (10%) (n=7)	Mélange (18%) (n=19)
Stratégied'abréviation privilégiée(1 ^{er} niveau)(1)	Acronyme (4%) (n=2)	Acronyme (17%) (n=2)	Acronyme (12%) (n=5)	Code (9%) (n=1)	Code Mélange (40%)(n=8)	n/a	n/a	Acronyme (14%) (n=7)	Troncature (6%) (n=1)	Mélange (50%) (n=9)
Stratégied'abréviation privilégiée(2 ^{ème} niveau)(2)	Troncature (11%) (n=5)	n/a	Acronyme (7%) (n=5)	Contraction (30%) (n=3)	n/a	Acronyme Troncature Mélange (3%) (n=1)	Mélange (22%) (n=10)	Acronyme (12%) (n=25)	Code (25%) (n=6)	Mélange (42%) (n=8)
Proportion(%)etnombre (n)declassesredondantes (tousniveauxconfondus)	34.51% (n=39)	0%	35.05% (n=68)	40.95% (n=43)	65% (n=13)	25.12% (n=51)	19.08% (n=29)	12.47% (n=47)	21.13% (n=15)	25.24% (n=26)
Proportion(%)etnombre (n)declassesredondantes (1 ^{er} niveau)	16.33% (n=8)	0%	9.52% (n=4)	0%	65% (n=13)	0%	0%	20.41% (n=10)	68.75% (n=11)	50% (n=9)

(1):Proportion(%)etnombre(n)declassesprincipales(1^{er}niveauhiérarchique)dechacundesschémasdeclassification.

(2):Proportion(%)etnombre(n)declassesubordonnées(2^{ème}niveauhiérarchique)dechacundesschémasdeclassification.

p :Schémadeclassificationpersonnelpartagéavec descollèguesdetravail.

Tableau 9.9 Synthèse des caractéristiques sémantiques des schémas de classification personnels observés

	<i>Techniciennes</i>		<i>Coordonneurs de stage</i>		<i>Ensemble des employés</i>	
	<i>Moyenne</i>	<i>Écart-type</i>	<i>Moyenne</i>	<i>Écart-type</i>	<i>Moyenne</i>	<i>Écart-type</i>
Longueurmoyenne (ennombredes caractères) desnomsde classe	15.99	4.06	16.26	3.27	16.12	0.19
Proportionde classes abrégées	40.55%	11.71	33.20	21.77	37.05	17.20
Stratégie d'abréviation privilegiée (tousniveauxconfondus)	Acronyme (54.54%)(n=6)		Acronyme(30%)(n=3) Code(30%)(n=3)		Acronyme 42.85%(n=9)	
Stratégie d'abréviation privilegiée(1 ^{er} niveau)	Acronyme (45.46%)(n=5)		Acronyme (40%)(n=4)		Acronyme 42.85%(n=9)	
Stratégie d'abréviation privilegiée(2 ^{ème} niveau)	Acronyme (54.54%)(n=6)		Acronyme (30%)(n=3)		Acronyme 42.85%(n=9)	
Classes redondantes (1 ^{er} niveau hiérarchique)(a)	45.46% (n=5)		60% (n=6)		52.38% (n=11)	

(a): Proportion(%) et nombre(n) des schémas de classification dotés de classes principales redondantes.

Annexe 10. Lettre d'appel à participation (Étudiants et diplômés)

[REDACTED]

Envoyé : [Date] [Heure]

À : [Listes]

Objet : Recherche de volontaires pour une expérimentation dans le cadre d'une recherche doctorale en sciences de l'information à l'EBSI

Bonjour,

Je suis actuellement à la recherche d'une trentaine de volontaires pour participer à une expérimentation reliée à l'organisation des documents administratifs électroniques. Cette expérience est réalisée dans le cadre de mon projet de recherche doctoral en sciences de l'information à l'EBSI.

Objectifs du projet : Cette recherche doctorale a pour but de vérifier l'existence d'un lien entre les différents schémas de classification (ou structures de répertoires) des documents électroniques produits ou reçus par le personnel d'un organisme dans le cadre de ses activités administratives et le repérage de documents électroniques. Plus précisément, cette recherche vise à identifier certains facteurs ou caractéristiques déterminantes à l'efficacité relative du repérage d'un sous-ensemble donné de documents électroniques. Par votre participation, vous contribuerez à une meilleure compréhension de la relation entre les schémas de classification et le repérage des documents administratifs électroniques.

Tâches à exécuter : Durant une séance d'une durée d'une heure environ, vous utiliserez un ordinateur pour simuler le repérage d'un sous-ensemble de vingt documents cibles préalablement décrits et organisés selon cinq structures de répertoires électroniques sensiblement différentes. Les mesures suivantes seront colligées : durée du repérage, nombre d'essais, taux de succès. Je colligerai votre appréciation générale de chaque structure de répertoires utilisée pour le repérage. Je vous demanderai également de répondre à un court questionnaire en vue de colliger des données de nature démographique. Aucune connaissance particulière n'est exigée outre l'habilité à naviguer à travers une structure arborescente de répertoires électroniques (i.e. l'Explorateur Windows) à l'aide d'une souris.

Caractère confidentiel des informations : Cette recherche a été validée par le Comité d'éthique de la Faculté des arts et des sciences de l'Université de Montréal. Les données individuelles demeureront anonymes et confidentielles. Un numéro de code vous sera attribué. Seuls les résultats de groupes pourront être rendus publics.

Bénéfices prévus pour les sujets : Sur simple demande, je vous transmettrai les résultats de cette recherche, une fois l'étude terminée. Chaque participant(e) sera également admissible à un tirage au sort d'un certificat-cadeau de 50\$ chez Renaud-Bray (fin juin).

Modalités : L'expérience se déroulera au bureau adjacent (C-2052) au laboratoire d'archivistique de l'EBSI. La collecte des données sera réalisée jusqu'au 30 juin 2006. L'expérience pourra être réalisée dans la journée, en soirée ou la fin de semaine selon votre convenance et vos disponibilités.

Pour participer à l'étude ou pour toute information complémentaire :

[REDACTED]

Je vous remercie d'avance pour votre aimable participation à cette recherche et pour votre contribution à la formation des étudiants au doctorat en sciences de l'information de l'Université de Montréal.

Sabine Mas - Étudiante - Doctorat en sciences de l'information
École de bibliothéconomie et des sciences de l'information - Université de Montréal

Annexe 11. Exemple de grille d'observation (Groupe 1)

Code du répondant : E____
Date de la simulation : ____/____/____

Heure : _____ à _____
Durée : ____h____

Description du document		Localisation (3 essais)	Temps
Structure 111			
1.	Procès-verbal de la réunion du 30 septembre 2003 du Comité de gestion Comité de gestion	1.	
		2.	
		3.	
2.	Logos du département Communications\LOGO	1.	
		2.	
		3.	
3.	Fichier Excel comportant les résultats des évaluations des stagiaires à l'automne 2005 Stage\2005-2006\NotationA05	1.	
		2.	
		3.	
4.	Liste des superviseurs de stages qui ont reçu leur attestation en avril 2004 AttestationSS\EnvoiAvril2004	1.	
		2.	
		3.	
Appréciation générale du répondant :			

Description du document		Localisation (3 essais)	Temps
Structure 311			
5.	Travail pratique réalisé dans le cadre d'un cours de français pris à l'Université Cours de français	1.	
		2.	
		3.	
6.	Résultats de l'assignation des étudiants en stage en milieu communautaire pour l'année 2003-2004 Stages-2003-2004-Assignations 2003-2004-AssignationsCO	1.	
		2.	
		3.	
7.	Cahier d'accompagnement à l'intention du stagiaire pour le 1er stage (1234) Comité des stages-1234	1.	
		2.	
		3.	
8.	Lettre de remerciement adressée aux superviseurs de stage qui ont encadré un étudiant en stage durant l'année 2003-2004 Stage-2003-2004-Lettres de remerciement 2003-2004	1.	
		2.	
		3.	
Appréciation générale du répondant :			

Description du document		Localisation (3 essais)	Temps
Structure 331			
9.	Correspondancesurlecas problématique d'un étudiant durant son stage à l'automne 2005 01Stages\2005-2006\05Casprob	1.	
		2.	
		3.	
10	Fichierhtmlénumérantdes ressourcesdocumentairesen pédagogieà l'intentiondes professeursetdes superviseursdestage 06Documentationde référence\Pédagogie	1.	
		2.	
		3.	
11	Calendrierdescours offerts auxsuperviseursdestageen 2005-2006pourqu'ils reçoiventleurattestationleur permettantdesuperviserdes étudiants 03S.S\01AttestationSS\01 Cours\2005-2006	1.	
		2.	
		3.	
12	Formulairestandardqueles étudiantsdoiventcompléter avantleurstageauNouveau- Brunswick 01Stages\NB	1.	
		2.	
		3.	
Appréciation générale du répondant :			

Description du document		Localisation (3 essais)	Temps
Structure 113			
13	Sondage relatif au programme d'agrément à la maîtrise	1.	
		2.	
		3.	
MSc\Agrément			
14	Informations sur les chambres à louer à Chicoutimi à l'intention des stagiaires	1.	
		2.	
		3.	
Logement			
15	Disponibilité reçue d'un milieu (établissement de santé) pour recevoir des étudiants en stage en 2004-2005	1.	
		2.	
		3.	
Milieux\ES			
16	Compte rendu de la réunion d'équipe du 15 février 2005 sur la journée pré-stage	1.	
		2.	
		3.	
Stage\Réunions d'équipe			
Appréciation générale du répondant :			

Description du document		Localisation (3 essais)	Temps
Structure 333			
17	Lettre envoyée aux étudiants en juin 2005 relative à la nouvelle recommandation pour la vaccination contre l'hépatite B 02 Gestion du programme de vaccination\2004-2005\01 Documentation envoyée aux étudiants\Hépatite B	1.	
		2.	
		3.	
18	Document envoyé aux milieux pour connaître leurs disponibilités avant d'assigner les étudiants en stage en 2005-2006 01 Gestion courante des stages (annuelle)\2005-2006\01 Relation avec les milieux\01 Disponibilité des milieux	1.	
		2.	
		3.	
19	Liste d'adresses courrielles des superviseurs par milieu 03 Superviseurs de stage\03 Liste courrielles des superviseurs de stage par milieu	1.	
		2.	
		3.	
20	Compte-rendu de la rencontre annuelle du 16 mars 2004 avec les superviseurs-chefs 01 Gestion courante des stages (annuelle)\2003-2004\01 Relation avec les milieux\03 Rencontres annuelles	1.	
		2.	
		3.	
Appréciation générale du répondant :			

Données démographiques :

Âge : _____

Sexe : _____

Programme d'études : _____

Année d'étude : ____

Annexe 12. Directives communiquées aux répondants au début de la séance de simulation

DIRECTIVES

Dans le cadre de cette expérience, vous utiliserez un ordinateur portable pour repérer un sous-ensemble de vingt documents préalablement décrits et organisés selon cinq structures de répertoires électroniques différentes.

Les structures de répertoires que vous aurez à utiliser pour localiser les documents sont basées sur une structure réellement créée par une personne en vue d'organiser et de repérer les documents électroniques utilisés dans le cadre de ses fonctions. Cette personne, nous l'appellerons *Julie*, est responsable de la coordination des stages des étudiants d'un département de l'Université de Montréal dans le domaine des sciences de la santé. Elle réalise, entre autres, les activités suivantes :

- Informe les étudiants des objectifs et du déroulement des stages
- Assure la coordination générale du stage
- Coordonne ou participe à diverses publications (bulletins, dépliants, rapports, etc.)
- Collabore à la bonne marche de son secteur en colligeant de l'information, en présentant sous les formes appropriées les données relatives à son domaine et en suggérant des modifications susceptibles d'améliorer le fonctionnement des activités de son secteur.
- Gère le programme de vaccination des étudiants
- Agit comme personne ressource pour les étudiants et les partenaires qui supervisent les stagiaires
- Effectue des suivis pendant les stages auprès des partenaires pour évaluer la progression de l'étudiant et surtout régler les cas problèmes
- Assure le bon déroulement des stages
- Assure les communications avec les stagiaires et les partenaires
- Gère l'assignation des stages et le processus d'assignation
- Planifie, organise et met en œuvre les activités de promotion et d'information concernant le programme de stage (par exemple, les journées pré-stage)
- Organise et participe à l'animation des rencontres avec les milieux de stage
- Tient à jour les informations pertinentes concernant les organismes et les milieux professionnels appropriés

Dans le cadre de cette recherche, nous vous demandons de naviguer (en utilisant l'Explorateur Windows) à travers différentes structures de répertoires successives afin de localiser le répertoire (ou sous-répertoire) dans lequel vous pensez que le document décrit se trouve classé. Pour des raisons de confidentialité, les différentes structures que vous utiliserez ne contiennent aucun document. C'est la chercheuse qui vous informera si vous avez localisé le « bon » répertoire, c'est-à-dire celui qui est réellement utilisé par la personne pour classer le document recherché. Vous avez la possibilité d'effectuer trois tentatives pour identifier le répertoire ou sous-répertoire. Pour chacune des tentatives, nous mesurons le temps écoulé pour localiser le répertoire ou sous-répertoire. Nous colligerons votre appréciation générale pour chacune des structures arborescentes de répertoires utilisée pour le repérage des documents. Pour terminer, nous vous demanderons de bien vouloir répondre à un court questionnaire en vue de colliger des données de nature démographique.

Nous vous remercions pour votre aimable participation.

Annexe 13. Description des documents à repérer

1. Procès-verbal de la réunion du 30 septembre 2003 du Comité de gestion
2. Logo du département
3. Fichier Excel comportant les résultats des évaluations des stagiaires à l'automne 2005
4. Liste des superviseurs de stages qui ont reçu leur attestation en avril 2004
5. Travail pratique réalisé dans le cadre d'un cours de français pris à l'Université
6. Résultats de l'assignation des étudiants en stage en milieu communautaire pour l'année 2003-2004
7. Cahier d'accompagnement à l'intention du stagiaire pour le 1er stage (1234)
8. Lettre de remerciement adressée aux superviseurs de stage qui ont encadré un étudiant en stage durant l'année 2003-2004
9. Correspondance sur le cas problématique d'un étudiant durant son stage à l'automne 2005
10. Fichier html énumérant des ressources documentaires en pédagogie à l'intention des professeurs et des superviseurs de stage

11. Calendrier des cours offerts aux superviseurs de stage en 2005-2006 pour qu'ils reçoivent leur attestation leur permettant de superviser des étudiants
12. Formulaire standard que les étudiants doivent compléter avant leur stage au Nouveau-Brunswick
13. Sondage relatif au programme d'agrément à la maîtrise
14. Information sur les chambres à louer à Chicoutimi à l'intention des stagiaires
15. Disponibilité reçue d'un milieu (établissement de santé) pour recevoir des étudiants en stage en 2004-2005
16. Compte rendu de la réunion d'équipe du 15 février 2005 sur la journée pré-stage
17. Lettre envoyée aux étudiants en juin 2005 relative à la nouvelle recommandation pour la vaccination contre l'hépatite B
18. Document envoyé aux milieux pour connaître leurs disponibilités avant d'assigner les étudiants en stage en 2005-2006
19. Liste d'adresses de courriel des superviseurs par milieu
20. Compte-rendu de la rencontre annuelle du 16 mars 2004 avec les superviseurs-chefs

Annexe 14. Préparation des données pour le SCMEQ (Extrait)

id sujets	S111d1Temps	S111d1Essais	S111d1Succès	S111d2Temps	S111d2Essais
E01	8,65	1	1	63,98	1
E02	9,64	1	1	66,87	1
E03	3,98	1	1	69,68	2
E04	16,22	1	1	58,22	1
E05	29,48	1	1	42,14	1
E06	33,16	1	1	86,13	3
E07	5,4	1	1	34,65	2
E08	20,43	1	1	26,06	1
E09	20,6	1	1	29,65	1
E10	16,25	1	1	68,59	2
E11	2,16	1	1	22,18	1
E12	9,12	1	1	54,6	2
E13	24,36	2	1	18,31	1
E14	90,31	2	1	105,81	2
E15	5,51	1	1	11,59	1
E16	21,61	1	1	40,16	1
E17	56,09	2	1	50,65	1
E18	9,92	1	1	8,37	1
E19	17,8	1	1	13,78	1
E20	179,12	3	0	291,8	1
E21	45,14	1	1	32,81	1
E22	48,06	2	1	283,93	2
E23	212,79	2	1	7,23	1
E24	62,39	1	1	51,13	1
E25	21,64	1	1	50,16	1
E26	17,02	1	1	8,26	1
E27	64,4	1	1	11,21	1
E28	28,82	1	1	35,28	1
E29	65,08	1	1	21,41	1
E30	14,99	1	1	48,06	1
E31	15,52	1	1	5,72	1
E32	133	1	1	66,62	1
E33	29,69	1	1	63,18	3
E34	34,08	1	1	11,62	1
Etc.	:	:	:	:	:

Annexe 15. Analyse de la variable « Succès »

Tableaux croisés : schéma versus succès, échecs et abandons

Tableau 15.1 Récapitulatif du traitement des observations

	Observations					
	Valide		Manquante		Total	
	N	Pourcent	N	Pourcent	N	Pourcent
Schéma * Succès/échec/abandon	1400	100.0%	0	.0%	1400	100.0%

Tableau 15.2 Tableau croisé Schéma * Succès/échec/abandon

			Succès/échec/abandon			Total
			échec (3 essais)	succès	abandon	
Schéma	A	Effectif	27	233	20	280
		% dans Schéma	9.6%	83.2%	7.1%	100.0%
		% dans Succès/échec/abandon	19.6%	19.9%	22.5%	20.0%
		% du total	1.9%	16.6%	1.4%	20.0%
	B	Effectif	24	239	17	280
		% dans Schéma	8.6%	85.4%	6.1%	100.0%
		% dans Succès/échec/abandon	17.4%	20.4%	19.1%	20.0%
		% du total	1.7%	17.1%	1.2%	20.0%
	C	Effectif	40	220	20	280
		% dans Schéma	14.3%	78.6%	7.1%	100.0%
		% dans Succès/échec/abandon	29.0%	18.8%	22.5%	20.0%
		% du total	2.9%	15.7%	1.4%	20.0%
D	Effectif	33	226	21	280	
	% dans Schéma	11.8%	80.7%	7.5%	100.0%	
	% dans Succès/échec/abandon	23.9%	19.3%	23.6%	20.0%	
	% du total	2.4%	16.1%	1.5%	20.0%	
E	Effectif	14	255	11	280	
	% dans Schéma	5.0%	91.1%	3.9%	100.0%	
	% dans Succès/échec/abandon	10.1%	21.7%	12.4%	20.0%	
	% du total	1.0%	18.2%	.8%	20.0%	
Total	Effectif	138	1173	89	1400	
	% dans Schéma	9.9%	83.8%	6.4%	100.0%	
	% dans Succès/échec/abandon	100.0%	100.0%	100.0%	100.0%	
	% du total	9.9%	83.8%	6.4%	100.0%	

Tableaux croisés : succès versus document

Tableau 15.3 Récapitulatif du traitement des observations

	Observations					
	Valide		Manquante		Total	
	N	Pourcent	N	Pourcent	N	Pourcent
Document * Succès	1400	100.0%	0	.0%	1400	100.0%

Tableau 15.4 Tableau croisé Document * Succès

			Succès		Total
			échec	succès	
Document	1	Effectif	15	335	350
		% dans Document	4.3%	95.7%	100.0%
		% dans Succès	6.6%	28.6%	25.0%
		% du total	1.1%	23.9%	25.0%
	2	Effectif	44	306	350
		% dans Document	12.6%	87.4%	100.0%
		% dans Succès	19.4%	26.1%	25.0%
		% du total	3.1%	21.9%	25.0%
	3	Effectif	96	254	350
		% dans Document	27.4%	72.6%	100.0%
		% dans Succès	42.3%	21.7%	25.0%
		% du total	6.9%	18.1%	25.0%
4	Effectif	72	278	350	
	% dans Document	20.6%	79.4%	100.0%	
	% dans Succès	31.7%	23.7%	25.0%	
	% du total	5.1%	19.9%	25.0%	
Total	Effectif	227	1173	1400	
	% dans Document	16.2%	83.8%	100.0%	
	% dans Succès	100.0%	100.0%	100.0%	
	% du total	16.2%	83.8%	100.0%	

Tableaux croisés : succès versus nombre d'essais

Tableau 15.5 Récapitulatif du traitement des observations

	Observations					
	Valide		Manquante		Total	
	N	Pourcent	N	Pourcent	N	Pourcent
Essais (nombre) * Succès	1400	100.0%	0	.0%	1400	100.0%

Tableau 15.6 Tableau croisé Essais (nombre) * Succès

			Succès		Total
			échec	succès	
Essais (nombre)	1	Effectif	26	930	956
		% dans Essais (nombre)	2.7%	97.3%	100.0%
		% dans Succès	11.5%	79.3%	68.3%
		% du total	1.9%	66.4%	68.3%
	2	Effectif	63	177	240
		% dans Essais (nombre)	26.3%	73.8%	100.0%
		% dans Succès	27.8%	15.1%	17.1%
		% du total	4.5%	12.6%	17.1%
	3	Effectif	138	66	204
		% dans Essais (nombre)	67.6%	32.4%	100.0%
		% dans Succès	60.8%	5.6%	14.6%
		% du total	9.9%	4.7%	14.6%
Total	Effectif	227	1173	1400	
	% dans Essais (nombre)	16.2%	83.8%	100.0%	
	% dans Succès	100.0%	100.0%	100.0%	
	% du total	16.2%	83.8%	100.0%	

Tableaux croisés : succès versus essais pour chacun des schémas

1) Schéma A

Tableau 15.7 Récapitulatif du traitement des observations (a)

	Observations					
	Valide		Manquante		Total	
	N	Pourcent	N	Pourcent	N	Pourcent
Essais (nombre) * Succès	280	100.0%	0	.0%	280	100.0%

a Schéma = A

Tableau 15.8 Tableau croisé Essais (nombre) * Succès (a)

			Succès		Total
			échec	succès	
Essais (nombre)	1	Effectif	8	174	182
		% dans Essais (nombre)	4.4%	95.6%	100.0%
		% dans Succès	17.0%	74.7%	65.0%
		% du total	2.9%	62.1%	65.0%
	2	Effectif	12	42	54
		% dans Essais (nombre)	22.2%	77.8%	100.0%
		% dans Succès	25.5%	18.0%	19.3%
		% du total	4.3%	15.0%	19.3%
	3	Effectif	27	17	44
		% dans Essais (nombre)	61.4%	38.6%	100.0%
		% dans Succès	57.4%	7.3%	15.7%
		% du total	9.6%	6.1%	15.7%
Total		Effectif	47	233	280
		% dans Essais (nombre)	16.8%	83.2%	100.0%
		% dans Succès	100.0%	100.0%	100.0%
		% du total	16.8%	83.2%	100.0%

a Schéma = A

2) Schéma B

Tableau 15.9 Récapitulatif du traitement des observations (a)

	Observations					
	Valide		Manquante		Total	
	N	Pourcent	N	Pourcent	N	Pourcent
Essais (nombre) * Succès	280	100.0%	0	.0%	280	100.0%

a Schéma = B

Tableau 15.10 Tableau croisé Essais (nombre) * Succès (a)

			Succès		Total
			échec	succès	
Essais (nombre)	1	Effectif	4	185	189
		% dans Essais (nombre)	2.1%	97.9%	100.0%
		% dans Succès	9.8%	77.4%	67.5%
		% du total	1.4%	66.1%	67.5%
	2	Effectif	13	40	53
		% dans Essais (nombre)	24.5%	75.5%	100.0%
		% dans Succès	31.7%	16.7%	18.9%
		% du total	4.6%	14.3%	18.9%
	3	Effectif	24	14	38
		% dans Essais (nombre)	63.2%	36.8%	100.0%
		% dans Succès	58.5%	5.9%	13.6%
		% du total	8.6%	5.0%	13.6%
Total		Effectif	41	239	280
		% dans Essais (nombre)	14.6%	85.4%	100.0%
		% dans Succès	100.0%	100.0%	100.0%
		% du total	14.6%	85.4%	100.0%

a Schéma = B

3) Schéma C

Tableau 15.11 Récapitulatif du traitement des observations (a)

	Observations					
	Valide		Manquante		Total	
	N	Pourcent	N	Pourcent	N	Pourcent
Essais (nombre) * Succès	280	100.0%	0	.0%	280	100.0%

a Schéma = C

Tableau 15.12 Tableau croisé Essais (nombre) * Succès (a)

			Succès		Total
			échec	succès	
Essais (nombre)	1	Effectif	3	173	176
		% dans Essais (nombre)	1.7%	98.3%	100.0%
		% dans Succès	5.0%	78.6%	62.9%
	2	Effectif	17	33	50
		% dans Essais (nombre)	34.0%	66.0%	100.0%
		% dans Succès	28.3%	15.0%	17.9%
	3	Effectif	40	14	54
		% dans Essais (nombre)	74.1%	25.9%	100.0%
		% dans Succès	66.7%	6.4%	19.3%
Total	Effectif	60	220	280	
	% dans Essais (nombre)	21.4%	78.6%	100.0%	
	% dans Succès	100.0%	100.0%	100.0%	
	% du total	21.4%	78.6%	100.0%	

a Schéma = C

4) Schéma D

Tableau 15.13 Récapitulatif du traitement des observations (a)

	Observations					
	Valide		Manquante		Total	
	N	Pourcent	N	Pourcent	N	Pourcent
Essais (nombre) * Succès	280	100.0%	0	.0%	280	100.0%

a Schéma = D

Tableau 15.14 Tableau croisé Essais (nombre) * Succès (a)

			Succès		Total
			échec	succès	
Essais (nombre)	1	Effectif	7	178	185
		% dans Essais (nombre)	3.8%	96.2%	100.0%
		% dans Succès	13.0%	78.8%	66.1%
	2	Effectif	14	34	48
		% dans Essais (nombre)	29.2%	70.8%	100.0%
		% dans Succès	25.9%	15.0%	17.1%
	3	Effectif	33	14	47
		% dans Essais (nombre)	70.2%	29.8%	100.0%
		% dans Succès	61.1%	6.2%	16.8%
Total	Effectif	54	226	280	
	% dans Essais (nombre)	19.3%	80.7%	100.0%	
	% dans Succès	100.0%	100.0%	100.0%	
	% du total	19.3%	80.7%	100.0%	

a Schéma = D

5) Schéma E

Tableau 15.15 Récapitulatif du traitement des observations (a)

	Observations					
	Valide		Manquante		Total	
	N	Pourcent	N	Pourcent	N	Pourcent
Essais (nombre) * Succès	280	100.0%	0	.0%	280	100.0%

a Schéma = E

Tableau 15.16 Tableau croisé Essais (nombre) * Succès (a)

			Succès		Total
			échec	succès	
Essais (nombre)	1	Effectif	4	220	224
		% dans Essais (nombre)	1.8%	98.2%	100.0%
		% dans Succès	16.0%	86.3%	80.0%
	2	Effectif	7	28	35
		% dans Essais (nombre)	20.0%	80.0%	100.0%
		% dans Succès	28.0%	11.0%	12.5%
	3	Effectif	14	7	21
		% dans Essais (nombre)	66.7%	33.3%	100.0%
		% dans Succès	56.0%	2.7%	7.5%
Total	Effectif	25	255	280	
	% dans Essais (nombre)	8.9%	91.1%	100.0%	
	% dans Succès	100.0%	100.0%	100.0%	
	% du total	8.9%	91.1%	100.0%	

a Schéma = E

Fréquences pour le nombre de succès total par répondant

Tableau 15.17 Nombre de succès total par répondant

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	11.00	2	2.9	2.9	2.9
	13.00	3	4.3	4.3	7.1
	14.00	4	5.7	5.7	12.9
	15.00	8	11.4	11.4	24.3
	16.00	10	14.3	14.3	38.6
	17.00	15	21.4	21.4	60.0
	18.00	14	20.0	20.0	80.0
	19.00	11	15.7	15.7	95.7
	20.00	3	4.3	4.3	100.0
	Total	70	100.0	100.0	

Fréquences de la proportion de succès par schéma

Tableau 15.18 Statistiques

		Proportion de documents repérés-Schéma A	Proportion de documents repérés-Schéma B	Proportion de documents repérés-Schéma C	Proportion de documents repérés-Schéma D	Proportion de documents repérés-Schéma E
N	Valide	70	70	70	70	70
	Manquante	0	0	0	0	0

Tableau de fréquences

Tableau 15.19 Proportion de documents repérés-Schéma A

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.25	2	2.9	2.9	2.9
	.50	7	10.0	10.0	12.9
	.75	27	38.6	38.6	51.4
	1.00	34	48.6	48.6	100.0
	Total	70	100.0	100.0	

Tableau 15.20 Proportion de documents repérés-Schéma B

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.50	8	11.4	11.4	11.4
	.75	25	35.7	35.7	47.1
	1.00	37	52.9	52.9	100.0
	Total	70	100.0	100.0	

Tableau 15.21 Proportion de documents repérés-Schéma C

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.25	2	2.9	2.9	2.9
	.50	9	12.9	12.9	15.7
	.75	36	51.4	51.4	67.1
	1.00	23	32.9	32.9	100.0
	Total	70	100.0	100.0	

Tableau 15.22 Proportion de documents repérés-Schéma D

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.25	1	1.4	1.4	1.4
	.50	11	15.7	15.7	17.1
	.75	29	41.4	41.4	58.6
	1.00	29	41.4	41.4	100.0
	Total	70	100.0	100.0	

Tableau 15.23 Proportion de documents repérés-Schéma E

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.50	4	5.7	5.7	5.7
	.75	17	24.3	24.3	30.0
	1.00	49	70.0	70.0	100.0
	Total	70	100.0	100.0	

ANOVA à mesures répétées pour la proportion de succès par schéma

Tableau 15.24 Facteurs intra-répondants

struct	Variable dépendante
1	propsuccesA
2	propsuccesB
3	propsuccesC
4	propsuccesD
5	propsuccesE

Tableau 15.25 Statistiques descriptives

	Moyenne	Ecart-type	N
Proportion de documents repérés-Schéma A	.8321	.19375	70
Proportion de documents repérés-Schéma B	.8536	.17285	70
Proportion de documents repérés-Schéma C	.7857	.18692	70
Proportion de documents repérés-Schéma D	.8071	.19116	70
Proportion de documents repérés-Schéma E	.9107	.14766	70

Tableau 15.26 Tests des effets intra-répondants

Source		Somme des carrés de type III	ddl	Moyenne des carrés	F	Signification
struct	Sphéricité supposée	.648	4	.162	5.877	.000
	Greenhouse-Geisser	.648	3.725	.174	5.877	.000
	Huynh-Feldt	.648	3.964	.163	5.877	.000
	Borne inférieure	.648	1.000	.648	5.877	.018
Erreur(struct)	Sphéricité supposée	7.603	276	.028		
	Greenhouse-Geisser	7.603	257.005	.030		
	Huynh-Feldt	7.603	273.494	.028		
	Borne inférieure	7.603	69.000	.110		

Tableau 15.27 Comparaisons par paires de schémas de classification

(I) struct	(J) struct	Différence des moyennes (I-J)	Erreur standard	Signification (a)	Intervalle de confiance de la différence à 95%(a)	
					Borne inférieure	Limite supérieure
1	2	-.021	.026	1.000	-.096	.053
	3	.046	.028	.963	-.033	.126
	4	.025	.030	1.000	-.062	.112
	5	-.079(*)	.026	.032	-.153	-.004
2	1	.021	.026	1.000	-.053	.096
	3	.068	.029	.230	-.017	.153
	4	.046	.028	.963	-.033	.126
	5	-.057	.027	.379	-.135	.021
3	1	-.046	.028	.963	-.126	.033
	2	-.068	.029	.230	-.153	.017
	4	-.021	.032	1.000	-.114	.072
	5	-.125(*)	.025	.000	-.197	-.053
4	1	-.025	.030	1.000	-.112	.062
	2	-.046	.028	.963	-.126	.033
	3	.021	.032	1.000	-.072	.114
	5	-.104(*)	.030	.009	-.190	-.017
5	1	.079(*)	.026	.032	.004	.153
	2	.057	.027	.379	-.021	.135
	3	.125(*)	.025	.000	.053	.197
	4	.104(*)	.030	.009	.017	.190

Basée sur les moyennes marginales estimées

* La différence de moyennes est significative au niveau .05

a Ajustement des comparaisons multiples : Bonferroni.

Annexe 16. Analyse de la variable « Temps »

ANOVA à mesures répétées pour le temps moyen des réussites pour ceux ayant repéré un minimum de deux documents

Tableau 16.1 Facteurs intra-répondants

structur	Variable dépendante
1	SATempsmoy2
2	SBTempsmoy2
3	SCTempsmoy2
4	SDTempsmoy2
5	SETempsmoy2

Tableau 16.2 Statistiques descriptives

	Moyenne	Ecart-type	N
Temps moyen-Schéma A (2)	70.3840	47.10418	66
Temps moyen-Schéma B (2)	79.5943	50.32045	66
Temps moyen-Schéma C (2)	64.7149	38.44922	66
Temps moyen-Schéma D (2)	55.3266	41.07871	66
Temps moyen-Schéma E (2)	47.4734	36.77577	66

Tableau 16.3 Tests des effets intra-répondants

Source		Somme des carrés de type III	ddl	Moyenne des carrés	F	Signification
structur	Sphéricité supposée	41682.227	4	10420.557	7.605	.000
	Greenhouse-Geisser	41682.227	3.563	11697.983	7.605	.000
	Huynh-Feldt	41682.227	3.795	10982.593	7.605	.000
	Borne inférieure	41682.227	1.000	41682.227	7.605	.008
Erreur(structur)	Sphéricité supposée	356249.929	260	1370.192		
	Greenhouse-Geisser	356249.929	231.608	1538.160		
	Huynh-Feldt	356249.929	246.694	1444.094		
	Borne inférieure	356249.929	65.000	5480.768		

Tableau 16.4 Comparaisons par paires de schémas de classification

(I) structur	(J) structur	Différence des moyennes (I-J)	Erreur standard	Signification (a)	Intervalle de confiance de la différence à 95%(a)	
					Borne inférieure	Limite supérieure
1	2	-9.210	7.620	1.000	-31.355	12.935
	3	5.669	6.448	1.000	-13.069	24.408
	4	15.057	6.422	.221	-3.604	33.719
	5	22.911(*)	5.759	.002	6.175	39.647
2	1	9.210	7.620	1.000	-12.935	31.355
	3	14.879	6.134	.181	-2.947	32.706
	4	24.268(*)	7.568	.021	2.275	46.261
	5	32.121(*)	6.642	.000	12.819	51.422
3	1	-5.669	6.448	1.000	-24.408	13.069
	2	-14.879	6.134	.181	-32.706	2.947
	4	9.388	6.127	1.000	-8.415	27.192
	5	17.241(*)	5.726	.037	.602	33.881
4	1	-15.057	6.422	.221	-33.719	3.604
	2	-24.268(*)	7.568	.021	-46.261	-2.275
	3	-9.388	6.127	1.000	-27.192	8.415
	5	7.853	5.640	1.000	-8.537	24.243
5	1	-22.911(*)	5.759	.002	-39.647	-6.175
	2	-32.121(*)	6.642	.000	-51.422	-12.819
	3	-17.241(*)	5.726	.037	-33.881	-.602
	4	-7.853	5.640	1.000	-24.243	8.537

Basée sur les moyennes marginales estimées

* La différence de moyennes est significative au niveau .05

a Ajustement des comparaisons multiples : Bonferroni.

ANOVA à mesures répétées pour le temps moyen des réussites pour ceux ayant repéré un minimum de trois documents

Tableau 16.5 Facteurs intra-répondants

structur	Variable dépendante
1	SATempsmoy3
2	SBTempsmoy3
3	SCTempsmoy3
4	SDTempsmoy3
5	SETempsmoy3

Tableau 16.6 Statistiques descriptives

	Moyenne	Ecart-type	N
Temps moyen-Schéma A (3)	67.6909	40.90768	38
Temps moyen-Schéma B (3)	81.1393	52.51817	38
Temps moyen-Schéma C (3)	67.8499	38.48468	38
Temps moyen-Schéma D (3)	59.5896	37.98877	38
Temps moyen-Schéma E (3)	44.6657	23.65979	38

Tableau 16.7 Tests des effets intra-répondants

Source		Somme des carrés de type III	ddl	Moyenne des carrés	F	Signification
structur	Sphéricité supposée	27181.045	4	6795.261	4.984	.001
	Greenhouse-Geisser	27181.045	3.160	8601.742	4.984	.002
	Huynh-Feldt	27181.045	3.489	7789.836	4.984	.002
	Borne inférieure	27181.045	1.000	27181.045	4.984	.032
Erreur(structur)	Sphéricité supposée	201789.412	148	1363.442		
	Greenhouse-Geisser	201789.412	116.918	1725.905		
	Huynh-Feldt	201789.412	129.104	1562.999		
	Borne inférieure	201789.412	37.000	5453.768		

Tableau 16.8 Comparaisons par paires de schémas de classification

(I) structur	(J) structur	Différence des moyennes (I-J)	Erreur standard	Signification (a)	Intervalle de confiance de la différence à 95%(a)	
					Borne inférieure	Limite supérieure
1	2	-13.448	10.274	1.000	-44.120	17.223
	3	-.159	7.043	1.000	-21.183	20.865
	4	8.101	9.295	1.000	-19.648	35.850
	5	23.025(*)	6.675	.014	3.098	42.952
2	1	13.448	10.274	1.000	-17.223	44.120
	3	13.289	9.012	1.000	-13.613	40.192
	4	21.550	10.648	.502	-10.236	53.336
	5	36.474(*)	9.336	.004	8.604	64.343
3	1	.159	7.043	1.000	-20.865	21.183
	2	-13.289	9.012	1.000	-40.192	13.613
	4	8.260	7.531	1.000	-14.221	30.742
	5	23.184(*)	6.794	.016	2.903	43.465
4	1	-8.101	9.295	1.000	-35.850	19.648
	2	-21.550	10.648	.502	-53.336	10.236
	3	-8.260	7.531	1.000	-30.742	14.221
	5	14.924	6.846	.357	-5.513	35.360
5	1	-23.025(*)	6.675	.014	-42.952	-3.098
	2	-36.474(*)	9.336	.004	-64.343	-8.604
	3	-23.184(*)	6.794	.016	-43.465	-2.903
	4	-14.924	6.846	.357	-35.360	5.513

Basée sur les moyennes marginales estimées

* La différence de moyennes est significative au niveau .05

a Ajustement des comparaisons multiples : Bonferroni.

Annexe 17. Analyse de la variable « Essai »

Fréquences pour la proportion de documents trouvés en un essai (pour un minimum de deux documents repérés)

Tableaux des fréquences pour la proportion des documents repérés en un essai

Tableau 17.1 Fréquences pour la proportion de documents repérés en un essai - Schéma A (2)

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.25	6	8.6	8.8	8.8
	.50	25	35.7	36.8	45.6
	.75	31	44.3	45.6	91.2
	1.00	6	8.6	8.8	100.0
	Total	68	97.1	100.0	
Manquante	Système manquant	2	2.9		
Total		70	100.0		

Tableau 17.2 Fréquences pour la proportion de documents repérés en un essai - Schéma B (2)

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.25	3	4.3	4.3	4.3
	.50	26	37.1	37.1	41.4
	.75	34	48.6	48.6	90.0
	1.00	7	10.0	10.0	100.0
	Total	70	100.0	100.0	

Tableau 17.3 Fréquences pour la proportion de documents repérés en un essai - Schéma C (2)

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.00	2	2.9	2.9	2.9
	.25	8	11.4	11.8	14.7
	.50	20	28.6	29.4	44.1
	.75	29	41.4	42.6	86.8
	1.00	9	12.9	13.2	100.0
	Total	68	97.1	100.0	
Manquante	Système manquant	2	2.9		
Total		70	100.0		

Tableau 17.4 Fréquences pour la proportion de documents repérés en un essai - Schéma D (2)

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.25	5	7.1	7.2	7.2
	.50	29	41.4	42.0	49.3
	.75	26	37.1	37.7	87.0
	1.00	9	12.9	13.0	100.0
	Total	69	98.6	100.0	
Manquante	Système manquant	1	1.4		
Total		70	100.0		

Tableau 17.5 Fréquences pour la proportion de documents repérés en un essai - Schéma E (2)

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.25	2	2.9	2.9	2.9
	.50	14	20.0	20.0	22.9
	.75	26	37.1	37.1	60.0
	1.00	28	40.0	40.0	100.0
	Total	70	100.0	100.0	

ANOVA à mesures répétées pour la proportion de documents trouvés en un essai (pour un minimum de deux documents repérés)

Modèle linéaire général

Tableau 17.6 Facteurs intra-répondants

proessai	Variable dépendante
1	SAprpessai2
2	SBpropessai2
3	SCpropessai2
4	SDpropessai2
5	SEpropessai2

Tableau 17.7 Statistiques descriptives

	Moyenne	Ecart-type	N
Taux de réussite en un essai - Schéma A (2)	.6402	.19675	66
Taux de réussite en un essai - Schéma B (2)	.6705	.17628	66
Taux de réussite en un essai - Schéma C (2)	.6250	.24515	66
Taux de réussite en un essai - Schéma D (2)	.6439	.20597	66
Taux de réussite en un essai - Schéma E (2)	.7955	.20526	66

Tableau 17.8 Tests des effets intra-répondants

Source		Somme des carrés de type III	ddl	Moyenne des carrés	F	Signification
proessai	Sphéricité supposée	1.268	4	.317	8.255	.000
	Greenhouse-Geisser	1.268	3.673	.345	8.255	.000
	Huynh-Feldt	1.268	3.920	.323	8.255	.000
	Borne inférieure	1.268	1.000	1.268	8.255	.005
Erreur(proessai)	Sphéricité supposée	9.982	260	.038		
	Greenhouse-Geisser	9.982	238.741	.042		
	Huynh-Feldt	9.982	254.812	.039		
	Borne inférieure	9.982	65.000	.154		

Tableau 17.9 Comparaisons par paires de schémas de classification

(I) proessai	(J) proessai	Différence des moyennes (I-J)	Erreur standard	Signification (a)	Intervalle de confiance de la différence à 95%(a)	
					Borne inférieure	Limite supérieure
1	2	-.030	.030	1.000	-.117	.056
	3	.015	.035	1.000	-.088	.118
	4	-.004	.033	1.000	-.100	.092
	5	-.155(*)	.030	.000	-.244	-.067
2	1	.030	.030	1.000	-.056	.117
	3	.045	.036	1.000	-.058	.149
	4	.027	.029	1.000	-.057	.110
	5	-.125(*)	.034	.005	-.225	-.025
3	1	-.015	.035	1.000	-.118	.088
	2	-.045	.036	1.000	-.149	.058
	4	-.019	.037	1.000	-.127	.089
	5	-.170(*)	.039	.001	-.285	-.056
4	1	.004	.033	1.000	-.092	.100
	2	-.027	.029	1.000	-.110	.057
	3	.019	.037	1.000	-.089	.127
	5	-.152(*)	.036	.001	-.255	-.048
5	1	.155(*)	.030	.000	.067	.244
	2	.125(*)	.034	.005	.025	.225
	3	.170(*)	.039	.001	.056	.285
	4	.152(*)	.036	.001	.048	.255

Basée sur les moyennes marginales estimées

* La différence de moyennes est significative au niveau .05

a Ajustement des comparaisons multiples : Bonferroni.

**Fréquences pour la proportion de documents trouvés en un essai
(pour un minimum de trois documents repérés)**

Tableaux des fréquences pour la proportion des documents repérés en un essai

Tableau 17.10 Fréquences pour la proportion de documents repérés en un essai - Schéma A (3)

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.25	5	7.1	8.2	8.2
	.50	19	27.1	31.1	39.3
	.75	31	44.3	50.8	90.2
	1.00	6	8.6	9.8	100.0
	Total	61	87.1	100.0	
Manquante	Système manquant	9	12.9		
Total		70	100.0		

Tableau 17.11 Fréquences pour la proportion de documents repérés en un essai - Schéma B (3)

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.25	1	1.4	1.6	1.6
	.50	20	28.6	32.3	33.9
	.75	34	48.6	54.8	88.7
	1.00	7	10.0	11.3	100.0
	Total	62	88.6	100.0	
Manquante	Système manquant	8	11.4		
Total		70	100.0		

Tableau 17.12 Fréquences pour la proportion de documents repérés en un essai - Schéma C (3)

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.00	2	2.9	3.4	3.4
	.25	6	8.6	10.2	13.6
	.50	13	18.6	22.0	35.6
	.75	29	41.4	49.2	84.7
	1.00	9	12.9	15.3	100.0
	Total	59	84.3	100.0	
Manquante	Système manquant	11	15.7		
Total		70	100.0		

Tableau 17.13 Fréquences pour la proportion de documents repérés en un essai - Schéma D (3)

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.25	4	5.7	6.9	6.9
	.50	19	27.1	32.8	39.7
	.75	26	37.1	44.8	84.5
	1.00	9	12.9	15.5	100.0
	Total	58	82.9	100.0	
Manquante	Système manquant	12	17.1		
Total		70	100.0		

Tableau 17.14 Fréquences pour la proportion de documents repérés en un essai - Schéma E (3)

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.50	12	17.1	18.2	18.2
	.75	26	37.1	39.4	57.6
	1.00	28	40.0	42.4	100.0
	Total	66	94.3	100.0	
Manquante	Système manquant	4	5.7		
Total		70	100.0		

ANOVA à mesures répétées pour la proportion de documents trouvés en un essai pour un minimum de trois documents repérés)

Modèle linéaire général

Tableau 17.15 Facteurs intra-répondants

proessai	Variable dépendante
1	SApropessai3
2	SBpropessai3
3	SCpropessai3
4	SDpropessai3
5	SEpropessai3

Tableau 17.16 Statistiques descriptives

	Moyenne	Ecart-type	N
Taux de réussite en un essai - Schéma A (3)	.6513	.19745	38
Taux de réussite en un essai - Schéma B (3)	.6776	.16345	38
Taux de réussite en un essai - Schéma C (3)	.6447	.25089	38
Taux de réussite en un essai - Schéma D (3)	.6842	.20701	38
Taux de réussite en un essai - Schéma E (3)	.8355	.17703	38

Tableau 17.17 Tests des effets intra-répondants

Source		Somme des carrés de type III	ddl	Moyenne des carrés	F	Signification
proessai	Sphéricité supposée	.932	4	.233	6.251	.000
	Greenhouse-Geisser	.932	3.401	.274	6.251	.000
	Huynh-Feldt	.932	3.787	.246	6.251	.000
	Borne inférieure	.932	1.000	.932	6.251	.017
Erreur(proessai)	Sphéricité supposée	5.518	148	.037		
	Greenhouse-Geisser	5.518	125.842	.044		
	Huynh-Feldt	5.518	140.124	.039		
	Borne inférieure	5.518	37.000	.149		

Tableau 17.18 Comparaisons par paires

(I) proessai	(J) proessai	Différence des moyennes (I-J)	Erreur standard	Signification (a)	Intervalle de confiance de la différence à 95%(a)	
					Borne inférieure	Limite supérieure
1	2	-.026	.040	1.000	-.145	.092
	3	.007	.046	1.000	-.130	.143
	4	-.033	.043	1.000	-.162	.097
	5	-.184(*)	.037	.000	-.296	-.073
2	1	.026	.040	1.000	-.092	.145
	3	.033	.049	1.000	-.114	.180
	4	-.007	.033	1.000	-.106	.093
	5	-.158(*)	.043	.007	-.285	-.031
3	1	-.007	.046	1.000	-.143	.130
	2	-.033	.049	1.000	-.180	.114
	4	-.039	.049	1.000	-.184	.105
	5	-.191(*)	.054	.011	-.351	-.030
4	1	.033	.043	1.000	-.097	.162
	2	.007	.033	1.000	-.093	.106
	3	.039	.049	1.000	-.105	.184
	5	-.151(*)	.046	.021	-.288	-.015
5	1	.184(*)	.037	.000	.073	.296
	2	.158(*)	.043	.007	.031	.285
	3	.191(*)	.054	.011	.030	.351
	4	.151(*)	.046	.021	.015	.288

Basée sur les moyennes marginales estimées

* La différence de moyennes est significative au niveau .05

a Ajustement des comparaisons multiples : Bonferroni.

Annexe 18. Modèle d'analyse des schémas de classification

INDICATEURS DE MESURES	VALEURS
Code de l'employé	
Répertoire racine (analysé par le logiciel PDS)	
Créateur de la structure (O/N)	
Structure partagée avec des collègues (O/N)	
Dimension structurelle	
Étendue de la structure	
Nombre de classes principales (1 ^{er} niveau)	
Nombre moyen de classes à chaque niveau	
Profondeur de la structure	
Nombre maximum de niveaux observés	
Nombre moyen de niveaux par classe principale	
Équilibre de la structure	
Écart-type de la moyenne de classes à chaque niveau	
Complexité de la structure	
Nombre total de classes	
Dimension logique	
Critère de division logique (proportion des classes)	
1 ^{er} critère de division (classes principales)	
	Activité
	Corps
	Projet
	Temps
	Thème
	Type de contenu
	Mélange
2 ^{ème} critère de division logique (classes subordonnées)	
	Activité
	Corps
	Projet
	Temps
	Thème
	Type de contenu
	Mélange
Ordre de succession des classes	
1 ^{er} niveau	
2 ^{ème} niveau	
Dimension sémantique	
Longueur moyenne des intitulés de l'ensemble des classes observées (en nombre de caractères)	
Nombre moyen de caractères (classes de niveau 1)	
Nombre moyen de caractères (classes de niveau 2)	
Nombre moyen de caractères (classes de niveau 3)	
Nombre moyen de caractères (toutes les classes)	
Stratégies d'abréviation (proportion des classes)	
1 ^{er} niveau	

	Acronyme	
	Juxtaposition	
	Code	
	Contraction	
	Troncature	
	Mélange	
	Aucune abréviation	
2 ^{eme} niveau		
	Acronyme	
	Juxtaposition	
	Code	
	Contraction	
	Troncature	
	Mélange	
	Aucune abréviation	
Tous niveaux		
	Acronyme	
	Juxtaposition	
	Code	
	Contraction	
	Troncature	
	Mélange	
	Aucune abréviation	
Degré de redondance structurelle et conceptuelle (proportion des classes)		
1 ^{er} niveau		
2 ^{eme} niveau		
Tous niveaux		

**Annexe 19. Modélisation de la relation entre les variables
Schéma de classification et Repérage**

FACTEURS INSTITUTIONNELS

Politiques/Directives

Motivation/Leadership

Ressources



FACTEURS HUMAINS

Style de classement

Formation

Expérience

Rôle
professionnel



Schéma de classification

Repérage



Format technologique

Quantité de documents

Valeur/Fonction



FACTEURS DOCUMENTAIRES

Annexe 20. Synthèse des caractéristiques des schémas de classification personnels observés et du schéma A