

# Applying Faceted Classification to the Personal Organization of Electronic Records: Insights into the User Experience



SABINE MAS, DOMINIQUE MAUREL, and INGE ALBERTS

**RÉSUMÉ** Malgré les progrès technologiques récents, la plupart des systèmes d'exploitation continuent de classer les documents numériques selon une structure hiérarchique ou arborescente. Bien qu'elles puissent avoir l'air intuitives, ces méthodes de classement peuvent être limitantes puisque les documents sont classés dans un endroit donné selon la logique d'un seul individu, ce qui rend la compréhension difficile pour les autres personnes. De plus, les frontières inhérentes à ce type de structure exigent que les gens prennent des décisions arbitraires à partir des critères spécifiques sur lesquels le classement sera effectué (par exemple : la fonction ou l'activité administrative, le type de document, ou son sujet), même si un document peut avoir plusieurs attributs et pourrait être classé à divers endroits. Dans ce texte, les auteurs proposent une approche originale à facettes pour l'organisation des documents numériques, dans le but d'éliminer certaines de ces contraintes bien connues. Un aperçu de l'expérience des utilisateurs s'avère maintenant nécessaire afin d'améliorer à la fois les modèles théoriques et l'interface de navigation à facettes proposé. À cette fin, les auteurs présentent une étude pilote visant à examiner la pertinence d'ajouter aux schémas de classification traditionnels une classification à facettes pour le classement des documents dans un environnement numérique décentralisé.

**ABSTRACT** Despite recent technological advances, most operating systems still support the organization of electronic records according to hierarchical structures or directory trees. Although seemingly intuitive, these organization methods can be limiting as records are classified in one place according to one person's logic, thereby making it difficult for other users to relate. Furthermore, the inherent boundaries of such a structure necessitate making arbitrary decisions on the specific criteria upon which the classification will be based (e.g., the administrative function or activity, the document type, or its subject) even though a document may have several attributes and require classification in several classes. In this article, an original faceted approach to support electronic records organization is proposed in order to eliminate some of these well-known constraints. Insights into the user experience are now needed to improve the theoretical models and enhance the proposed faceted navigation interface. To that end, this paper discusses a pilot study investigating the pertinence of augmenting traditional hierarchical classification schemes with a faceted classification for records organization in a decentralized, digital environment.

## Introduction

Within an organization, documents are created and received in the context of a business function in order to perform an administrative or operational activity. This notion of context is crucial in an organizational environment where a document can be seen as the cause, or result of, a business process or activity. In order to account for this activity while providing better interpretation, authenticity, and probative value of the information contained in these documents, archivists and records managers advocate using document and folder classification based on the functions and activities of the organization.<sup>1</sup>

Although institutional classification schemes are intended to facilitate the organization of records for the short or long term as well as the localization and retrieval of these records by all users, employees rarely use this type of tool to organize and retrieve the electronic records that are saved on their workstations. To manage their electronic records, employees tend to prefer “personal” classification schemes that appear to be more adapted to their needs in fulfilling their daily activities, rather than functional classification schemes that reflect the institutional vision.<sup>2</sup> Individualism, which prevails most frequently in this context, is fostered by factors such as: the lack of leadership, resources, or institutional motivations; the dearth of policies, standards, methods, and processes for managing electronic documents; the lack of employee training and support; and the belief that electronic documents are not “official” documents of archival value.<sup>3</sup> To these factors, one should add the complexity inherent in the use of institutional classification schemes. Institutional classification schemes, originally designed for the organization of paper documents, are difficult to apply in the digital environment because of several pre-existing constraints. First, the often too generic terminology used in institutional classification schemes does not always correspond to what is used by the employees in their work. Employees name their folders and documents according to personal semantics often related to their domain of expertise. As well, the rigidity of institutional classification schemes and their irregular – or even non-existent – maintenance discourage individuals from using them in the electronic environment.<sup>4</sup> Finally, the cognitive effort involved in the act of classifying, and the lack of time generally, result in individuals using

- 1 International Organization for Standardization (ISO), *ISO 15489-1: Information and documentation - Records management. Part 1: General* (Geneva, 2001).
- 2 Sabine Mas, “Schémas de classification et repérage des documents administratifs électroniques dans un contexte de gestion décentralisée des ressources informationnelles” (PhD diss., Université de Montréal, 2007).
- 3 J. Timothy Sprehe, Charles R. McClure, and Philip Zellner, “The Role of Situational Factors in Managing U.S. Federal Recordkeeping,” *Government Information Quarterly*, vol. 19, no. 3 (2002), pp. 289–305.
- 4 William Saffady, *Managing Electronic Records* (Prairie Village, KS, 2002).

personal classification schemes to organize their electronic documents. In doing so, individuals no longer have to negotiate or discuss the rules of logical division or classificatory conventions they use; they just remember their own classificatory rules, changes, and exceptions.<sup>5</sup>

While individuals seem to be satisfied using their own classification schemes, one may question the ability of third parties to find documents using these personal schemes with which they are not familiar.<sup>6</sup> In the absence of a systematic and integrated approach to organize business documents, it becomes increasingly difficult for large organizations to manage their documentary heritage so as to comply with access to information laws, ensure decision-making transparency, and defend their rights using authentic records of their activities. The need to reconcile both individual and collective information practices that sometimes conflict with one another cannot be ignored, since user dissatisfaction with methods or systems judged to be inaccurate can lead to inefficient workarounds. Thus, the analysis of individual information practices and methods has to be considered in accordance with the type of information governance planned by the organization, as well as the informational culture already in place; taken together, these factors will affect the management of organizational memory.<sup>7</sup>

To overcome the difficulties associated with document organization, proposed solutions are emerging from different research areas. For instance, in information science, solutions such as replacing traditional hierarchical classification systems with a faceted infrastructure for indexing and searching through a collection of Web-based documentary resources, individual personal spaces, or entire organizations are proposed.<sup>8</sup> The demonstrated posi-

5 Barbara H. Kwasnik, "How a Personal Document's Intended Use or Purpose Affects Its Classification in an Office," *Proceedings of the 12th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Cambridge, Massachusetts* (1989), pp. 207–210.

6 Mas, "Schémas de classification et repérage."

7 Dominique Maurel and Pierrette Bergeron, "Quel rôle pour les archivistes dans la gestion de la mémoire organisationnelle?" *Archives*, vol. 40, no. 2 (2009), pp. 27–44; Dominique Maurel and Aïda Chebbi, "Adéquation des processus et systèmes documentaires aux impératifs des processus d'affaires: pour une gestion stratégique de l'information," *Revue du management électronique* 24 (June 2009), <http://www.cyber-gestion.com/prog/lettre revue/lettre24.htm> (accessed on 23 June 2011).

8 Chaim Zins, "Models for Classifying Internet Resources," *Knowledge Organization*, vol. 29, no. 1 (2002), pp. 20–28; Sarah Henderson, "Genre, Task, Topic and Time: Facets of Personal Digital Document Management," *Proceedings of the 6th ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction: Making CHI Natural* (New York, 2005), <http://portal.acm.org/citation.cfm?doid=1073943.1073957> (accessed on 24 June 2011); Michèle Hudon, "Proposal for an Alternate Structure to Organize and Access Resources in a Virtual Library in Education," *Proceedings of the 35th Conference of the Canadian Association for Information Science, May 10–12, 2007, McGill University, Montreal, Canada* (2007), [http://www.cais-acsi.ca/proceedings/2007/hudon\\_2007.pdf](http://www.cais-acsi.ca/proceedings/2007/hudon_2007.pdf)

tive impact of faceted classification approaches for organizing information in a networked environment<sup>9</sup> offers great promise for the virtual organization of electronic records.<sup>10</sup> These studies suggest that the flexibility, expressiveness, and simplicity of a faceted model – linking the content of the information object (closer to the individual needs) to the context in which this object was created (a requirement for records management) – would overcome the difficulties of implementing unidimensional, institutional, hierarchical classification schemes. It would also improve search precision, while promoting greater automation of the classification process so as to release users from a task they rarely accomplish with great success.<sup>11</sup>

Faceted classification as it applies to the virtual organization of electronic records in a business context has been little studied from the perspective of the user, even if such a classification approach has been identified as a promising alternative to a strictly hierarchical classification structure.<sup>12</sup> To improve the theoretical model and enhance the proposed faceted navigation interface, insights into the user experience are required.

This paper describes how a faceted classification approach was applied in a pilot project that aimed to design, use, and evaluate two specific faceted classification schemes for the daily creation of different virtual views of the documents that have been classified in a real work setting. The faceted classification approach is introduced in the literature review. We will then present and discuss the findings of this experiment. In conclusion, we identify areas of research related to the design of faceted classification schemes for organizing records in a decentralized digital environment. The project explored the following research questions:

1. What are the difficulties and sources of dissatisfaction experienced by

---

(accessed on 24 June 2011).

- 9 Winfried Gödert, "Facet Classification in Online Retrieval," *International Classification*, vol. 18, no. 2 (1991), pp. 98–109; Peter Ingwersen and Irene Wormell, "Ranganathan in the Perspective of Advanced Information Retrieval," *Libri*, vol. 42, no. 3 (1992), pp. 184–201; Marc Maisonneuve and Cécile Touitou, "Une nouvelle famille d'OPAC: Navigation à facettes et nuages de mots," *Bulletin des bibliothèques de France*, vol. 52, no. 6 (2007), pp. 12–19.
- 10 Henderson, "Genre, Task, Topic and Time."
- 11 Mas, "Schémas de classification et repérage."
- 12 Henderson, "Genre, Task, Topic and Time"; Sabine Mas, L'Hédi Zaher, and Manuel Zacklad, "Design and Evaluation of Multi-Viewed Knowledge System for Administrative Electronic Document Organization," *Proceedings of the Tenth International ISKO Conference: Culture and Identity in Knowledge Organization, August 5–8, 2008, Montreal, Canada* (2008), pp. 137–43; Sabine Mas and Yves Marleau, "Proposition of a Faceted Classification Model to Support Corporate Information Organization and Digital Records Management," *Proceedings of the 42nd Hawaii International Conference on System Sciences (HICSS), January 5–9, 2009, Waikoloa, Big Island, Hawaii* (2009), <http://doi.ieeecomputersociety.org/10.1109/HICSS.2009.874> (accessed on 24 June 2011).

users while organizing their electronic records using a faceted classification scheme, and what are the suggested changes?

2. What are the facilitating factors and the respondents' perceived sources of satisfaction while using a faceted classification scheme?
3. What are the respondents' perceptions of the impact of using a faceted classification scheme on their work performance?
4. What type of records organizing method (hierarchical or faceted) do the respondents prefer?

## Literature Review

In library and information science (LIS), "classification" refers to the act of organizing documents into classes based on common attributes or division criteria.<sup>13</sup> The first level of division is the most important since it describes and organizes the document holdings into fundamental logical classes.<sup>14</sup> However, even the simplest documents have several attributes (such as type or topic) that can be used to group records according to different classes (the possible combinations being endless since the same documents can be classified in different ways).<sup>15</sup>

The concept of "classification scheme" refers to the documentary classification tool. It is generally a hierarchical structure of descriptive categories used to facilitate the organization, retrieval, and use of a collection. This structure has a dual function since it should allow "both the classification of a field of knowledge and the classification of physical documents, depending on the nature of their subject."<sup>16</sup> Consequently, the structure must be both stable and sufficiently flexible to adapt to inevitable changes in a field of knowledge. These functions and constraints explain the limited number of model structures for classification schemes available in the literature; the hierarchical model being the most common model for the classification of administrative documents, in a paper or digital format.

Institutional hierarchical classification schemes are generally enumerative. On one hand, they must define in advance and include all topics relevant to an existing domain (e.g., the "human resources management" domain), while

13 Jacques Maniez, *Les langages documentaires et classificatoires: conception, construction et utilisation dans les systèmes documentaires* (Paris, 1987); Eric Hunter, "Do We Still Need Classification?" in *The Future of Classification*, eds. Rita Marcella and Arthur Maltby (London, 2000), pp. 1–17.

14 Barbara H. Kwasnik, "The Role of Classification in Knowledge Representation and Discovery," *Library Trends*, vol. 48, no. 1 (1999), pp. 22–47.

15 Derek Wilton Langridge, *Classification: Its Kinds, Elements, Systems, and Applications* (London, 1992); Jens-Erik Mai, "Likeness: A Pragmatic Approach," *Proceedings of the 6th International ISKO Conference, July 10–13, 2000, Toronto, Canada* (2000), pp. 23–27.

16 Maniez, *Les langages documentaires et classificatoires*, p. 43.

they must also predict the emergence of new topics (to ensure the classification coverage is accurate for each document) (See Figure 1).

**3100 Staffing of Human Resources**

3110 Staffing of regular employees

3120 Staffing of contract employees

3130 Staffing of occasional employees

**3200 Evaluation of Human Resources**

3210 Evaluation of regular employees

3220 Evaluation of contract employees

3230 Evaluation of occasional employees

**Figure 1: Excerpt from an Enumerative Hierarchical Classification Scheme.**<sup>17</sup>

In hierarchical classification schemes, there is an inevitable redundancy of concepts, exemplified in Figure 1 by the listing of the different categories of employees being the object of the evaluation and staffing activities. This type of classification system can be considered rigid since each document can only be associated to a single class, which is predetermined and fixed.

The concept of “personal” or “decentralized” organization of digital administrative records is distinct from the concept of “institutional” or “central” organization of documents. The personal organization of documents is defined as a process established by an individual to facilitate the organization and retrieval of documents that are under their direct control in order to meet personal needs.<sup>18</sup> Document organization is being used to refer to both the intellectual organization of documents (classification, indexing, etc.) and their physical organization (arrangement, storage) within a folder or directory structure. In the context of centralized management, it is an archivist or a records manager who designs and controls a system for organizing documents for a variety of individuals with different needs within either an organization or a single work unit.

Employees in organizations have been using hierarchical interfaces for several decades to organize the electronic records located on their workstations. On the one hand, these hierarchical interfaces have several cognitive

17 Example taken from Sabine Mas, “La classification à facettes: théorie et application pour l’organisation des documents institutionnels,” *Actes du 33e Congrès de l’Association des archivistes du Québec, Sainte-Adèle, Québec (2004)*, pp. 29–50.

18 Xia Lin and Lois Mai Chan, “Personalized Knowledge Organization and Access for the Web,” *Library & Information Science Research*, vol. 21, no. 2 (1999), pp. 153–72; Ofer Bergman, Ruth Beyth-Marom, and Rafi Nachmias, “The User-Subjective Approach to Personal Information Management Systems,” *Journal of the American Society for Information Science and Technology*, vol. 54, no. 9 (2003), pp. 872–78.

benefits for the user; for example, they enable the user to locate a document through local browsing of a tree-structure directory representing the spatial location where the document is classified.<sup>19</sup> On the other hand, the inherent limitations of such structures, already noted by several researchers,<sup>20</sup> are forcing individuals to classify documents according to only one personal criterion at each subdivision level, for instance, either the theme of the document, the type of document, the activity within which a document is used, or any other relevant classes, despite the fact that the document could simultaneously be classified in several different ways.

Some researchers have already observed that documents saved on individuals' workstations are classified according to different division criteria for which the application rules are not always comprehensible to other employees. For instance, the results of an analysis of twenty-one hierarchical classification schemes created and used by employees in a Canadian university have showed many differences in the logical division between personal classification and institutional classification schemes.<sup>21</sup> While institutional classification schemes are primarily designed to organize documents according to the administrative and operational functions and activities in which they were created or received, personal classification schemes are used by individuals to organize documents by topic (28% of top-level classes), format, or by a combination of classes representing several concepts (e.g., "Student Records F-2003").

The great variability observed in the folder names is consistent with the results of previous studies. Richard Boardman and Angela Sasse's conceptual analysis of 31 folder structures indicates that the most common types of file folder were *project* (e.g., "UCL presentation") (34%); *document class* (e.g., "Letters") (17%); and *role* (long-term activities, e.g., "Teaching") (9%).<sup>22</sup> A similar investigation of electronic documents found that the concepts most commonly used in folder names were: *genre* (e.g., "Lecture notes") (24%); *task* (e.g., "Evaluation") (15%); *topic* (e.g., "Database architecture") (11%); *time* (e.g., "2005") (8%); and *multiple concepts* (e.g., "Recruiting 2003") (8%), among others concepts.<sup>23</sup> The practice of mixing multiple division

19 Deborah K. Barreau and Bonnie A. Nardi, "Finding and Reminding: File Organization from the Desktop," *SIGCHI Bulletin*, vol. 27, no. 3 (1995), pp. 39–43.

20 Ibid.; Paul Dourish, W. Keith Edwards, Anthony LaMarca, John Lamping, Karin Petersen, Michael Salisbury, Douglas B. Terry, and James Thornton, "Extending Document Management Systems with User-Specific Active Properties," *ACM Transaction on Information Systems*, vol. 18, no. 2 (2000), pp. 140–70.

21 Mas, "Schémas de classification et repérage."

22 Richard Boardman and M. Angela Sasse, "Stuff Goes into the Computer and Doesn't Come Out: A Cross-Tool Study of Personal Information Management," *Proceedings of the 2004 Conference on Human Factors in Computing Systems, April 24–29, 2004, Vienna, Austria* (New York, 2004), pp. 583–90.

23 Henderson, "Genre, Task, Topic and Time."

concepts or principles in a developing hierarchy is contrary to the principles of classification because it creates classes of documents that are not mutually exclusive,<sup>24</sup> thus “causing uncertainty for the browser when he has to select a category.”<sup>25</sup> Such a mix is found at the first two levels in all personal classification schemes examined, making the structure less efficient for organization, retrieval, and update.<sup>26</sup> The inherent and ongoing process of updating personal classification schemes is one of the major problems raised by employees during interviews.<sup>27</sup> The results of this research confirm the relevance of exploring the potential of using faceted classification as a multi-dimensional means of organizing digital administrative documents.

Shiyali Ramamrita Ranganathan, an Indian mathematician and librarian of the first half of the twentieth century, introduced the concept of faceted classification for document description and management, and was the first to classify books in libraries. Ranganathan’s theory of facets provides syntax for expressing and “manually” representing the subject of a document. As in traditional classification approaches, Ranganathan proposed a first level of division between the major areas of knowledge, under the designation of “main facets.” These facets are associated with five “fundamental facets,” which he considered to be the basic categories needed to compose a subject: Personality, Matter, Energy, Space, and Time (PMEST). *Personality* is the object referred to in the document; *Matter* is the substance, property, or quality; *Energy* is the main operation or the action described in relation to the object; *Space* is the geographical location; and *Time* is the time period.<sup>28</sup>

Using the example provided in Figure 1, we observe that the human resources management process has at least two different types of concepts: an action (such as staffing or evaluation), and an object for this action (employee). It becomes possible to divide the previous example into two distinct facets: *Action* and *Object* (Figure 2).

<b>Action</b>	<b>Object</b>
Staffing	Regular employees
Evaluation	Contract employees
	Occasional employees

**Figure 2: Example of Faceted Classification.**

- 24 Kwasnik, “The Role of Classification in Knowledge Representation and Discovery.”
- 25 Martin S. Van der Walt, “The Structures of Classification Schemes Used in Internet Search Engines,” *Proceedings of the Fifth International ISKO Conference: Structures and Relations in Knowledge Organization, 25–29 August 1998, Lille, France*, (1998), p. 382.
- 26 Pat Molholt, “Qualities of Classification Schemes for the Information Superhighway,” *Cataloging & Classification Quarterly*, vol. 21, no. 2 (1995), pp. 19–22.
- 27 Mas, “Schémas de classification et repérages.”
- 28 Shiyali Ramamrita Ranganathan, *The Colon Classification* (New Brunswick, NJ, 1965).

The concepts, represented by terms or values, are listed within each facet and combined in a proposition when necessary: "Rather than attempting to enumerate all possible subjects, a faceted classification system allows a classifier to choose appropriate facets for expressing the document subject and to synthesize a notation to make the subject explicit."<sup>29</sup>

When saving an electronic document, the employee can, by tagging, choose between one or many pre-established values within one or many facets with the help of a drop-down menu; "tagging" the document allows it to be associated with one or many classes within the faceted classification scheme. As a result, rather than having a single organization of electronic documents (as in a hierarchical, enumerative classification structure), users can create several different "virtual" views of the documents that have been classified based on the faceted classification scheme. The effectiveness of faceted classifications lies in their ability to integrate the analysis of different information object dimensions, facilitating the characterization and the access to this information through multiple perspectives.

Some exploratory studies have investigated the possibility of replacing institutional hierarchical classification with faceted classification. A first study was conducted in 2001 to support the work of the Quebec government's Document Classification and Indexing Task Force.<sup>30</sup> The main objective of this analysis was to better characterize the nature of the facets suggested by the task force (e.g., "Business Process" and "Document Type") and to clarify their definition. The goal was also to verify that these facets included all the important aspects of the institutional hierarchical classification scheme; using the set of suggested facets, it was essential to be able to preserve the document creation context by identifying the administrative activity in which the document was produced or received. The analysis was conducted using a sample of six government departments and agencies' classification schemes pre-selected by the task force. A comparison within and between classification schemes was performed to assess the higher-level facets (such as "Business Process" and "Document Type"). This sample, however, was not sufficient to assess the lower level facets, namely the ones that do not normally appear in the classification scheme but rather in the document content (e.g., "Time," among others).

Following this research, an exploratory study was then conducted to investigate the possibility of replacing the hierarchical classification structures of electronic documents used by an internship coordinator at the University of Montreal, with a faceted classification scheme that would not only represent

29 Clare Beghtol, "Mapping Sentences and Classification Schedules as Methods of Displaying Facets," *Proceedings for the 6th ASIS/SIG. Workshop in Classification Research, Medford, NJ* (1995), p. 8.

30 Michèle Hudon and Sabine Mas, *Analyse des facettes pour la classification des documents institutionnels au Gouvernement du Québec* (Québec, 2001).

the context of production and use of these documents, but also their topics (or aboutness).<sup>31</sup> Based on the analysis of the documented activities and concepts used by the employee to group his documents, a faceted classification scheme was developed by the researchers to provide the respondent with the ability to classify and retrieve his documents according to six facets: 1) Activity; 2) Document Type; 3) Academic Semester; 4) Origin of the Document; 5) Document Recipient; and 6) Status of the Student. The assessment and use of this faceted classification scheme by the respondent has produced very encouraging results. The experimental design was used to validate with the employee the relevance and completeness of the facets identified for the organization of his documents, while facilitating the further maintenance of the classification structure by his active participation in the design process.

Even though there is empirical evidence suggesting better document organization and retrieval using faceted queries,<sup>32</sup> the question of how employees might interact with the faceted classification to create multiple views of their electronic records on a daily basis remains. Increasing information overload, cognitive effort, and work constraints are making document organization increasingly difficult. As already mentioned in previous knowledge organization research, it is a challenge to get users to manually enter metadata.<sup>33</sup> Employees are busy getting on with their work and are not usually fully committed to managing their documents beyond the minimum required to ensure further retrieval.

These considerations led us to experiment with ISIS (Integrated Semantic Information Services), a Web 2.0 software solution developed in partnership with the Cogniva Information Science Research Institute.<sup>34</sup> ISIS is an information infrastructure based on a faceted classification model, whose goal is to consistently create different “virtual” views of electronic records. The objectives of ISIS include providing a semantic infrastructure to fulfill records, information, and knowledge management needs for organizations while facilitating the information tasks for users. This is achieved by allowing a better description of the content and context of all types of information objects by supporting some classification tasks for the end-users, linking the information objects to the business processes where they are acquired or created, and by using a flexible, yet controlled vocabulary, through term localization.<sup>35</sup>

31 Mas, Zaher, and Zacklad, “Design and Evaluation of Multi-Viewed Knowledge System.”

32 Pertti Vakkari and Martti Kuokkanen, “Theory Growth in Information Science: Applications of the Theory of Science to a Theory of Information Seeking,” *Journal of Documentation*, vol. 53, no. 5 (1997), pp. 487–519.

33 Mas, “Schémas de classification et repérage.”

34 See <http://www.cognivaresearch.org/>.

35 Mas and Marleau, “Proposition of a Faceted Classification Model.”

## Methodology

A study was conducted to explore the use of a faceted classification scheme that would represent a document's context of production, use, and subject conveyed. The purpose of the study was to test the pertinence and usability of the faceted classification itself rather than investigate the usability of a browsable and co-buildable tool for the faceted classification of electronic records.

Two ISIS faceted classification models pertaining to different professional domains were developed for this experiment. University professors developed the first model for the purpose of course delivery. A small, private company that delivers information management (IM) consultancy services developed the second model. These models were derived from the analysis of two distinct business processes. In this study, a "business process" is defined as a sequence of tasks performed according to a pre-defined order. A process map identifies who is responsible for each task, the tasks being accomplished, and the informational inputs and outputs (e.g., documents).<sup>36</sup> Business processes may be operational, based on the mandate of the organization (production, customer service, etc.) or administrative, supporting these operational processes (human resources, financial, etc.). By their nature, some processes are strictly functional, i.e., corresponding only to the work of one administrative unit, while others are inter-functional, i.e., corresponding to the work common to more than one unit.<sup>37</sup> Documents created as part of business processes are considered "business objects" since they constitute the record of actions and decisions.

The two processes used to develop the ISIS faceted classification models were both operational and inter-functional. Although the subject domains are very different, the methodology followed in building these models was the same. The business analysis resulted in two process maps from which facets (their organization as well as semantic relationships) were derived. To provide a concrete example while avoiding redundancy, the ISIS model development process for the university domain is further detailed below.

### *Faceted Classification Model Development*

One of the processes analyzed for the pilot study was the design and deliv-

36 Inge Alberts, Jennifer Schellinck, Craig Eby, and Yves Marleau, "Bridging Functions and Processes for Records Management," *Canadian Journal of Information and Library Science*, vol. 34, no. 4 (2010), pp. 365–90.

37 Dominique Maurel and Dany Bouchard, "Repenser les fonctions documentaires selon les trajectoires des documents corporatifs," *Actes de la troisième conférence Document numérique et Société: Enjeux politiques du document numérique, 15–16 novembre 2010, Aix-en-Provence, France* (Paris, 2010), pp. 153–73.

ery of a course by university teachers. The resulting faceted classification model was developed by two of the authors of this paper, who are professors at the university in archival and information studies. The model elaboration was based on the combination of a top-down (knowledge of administrative functions and activities) and a bottom-up (knowledge of the records to be arranged) approach, a method of analysis commonly employed when designing hierarchical classification schemes. On one hand, it consisted of subdividing the conceptual domain structure according to rules of logical division, regardless of the content of the documents. The deductive approach (top-down) ensured consistency and flexibility for the future evolution of the classification scheme.<sup>38</sup> On the other hand, literary warrant, which based the development of the classification scheme on the documents available in the collection, ensured the development of a scheme tailored to user needs.<sup>39</sup>

Analysis of the activities/tasks performed by the professors during the planning and delivery of a 15-week course was first carried out. This analysis also involved the identification of the organizational actors (professors, students, teaching assistants, invited speakers, librarians, etc.). These players all contributed in varying degrees to achieving the activities and/or the creation of documents relevant to the management of the course. Finally, for each activity, the professors listed all the documents received (inputs) or created (outputs) during the process.

Process mapping was an essential component of the methodology to identify facets that would be used to classify information from an institutional perspective (such as linking documents to their functions), while fulfilling specific user task requirements. For instance, the professors were involved in different administrative, scientific, pedagogical, and professional activities associated with their positions. They not only were responsible for the tasks related to the management of a course (e.g., organizing meetings with students, planning, development and updating of teaching materials, planning lectures and course work, test scoring, and evaluation of students), they also played the role of administrator, researcher, lecturer, and author, to name but a few.

The nature of the tasks suggests that a professor may need more frequent and faster access to specific categories of documents as part of their activities. Professors participating in the study were interacting with the same types

38 Maniez, *Les langages documentaires et classificatoires*; Don Fallis and Kay Mathiesen, "Consistency Rules for Classification Schemes (or How to Organize your Beanie Babies)," *Proceedings of the Sixth International ISKO Conference: Dynamism and Stability in Knowledge Organization, July 10–13, 2000, Toronto, Canada* (Würzburg, 2000), pp. 339–44.

39 André Canonne, *Vocabulaire élémentaire des classifications* (Liège, 1993); Langridge; Hemalata Iyer, *Classificatory Structures Concepts, Relations and Representation* (Frankfurt, 1995).

of documents on a daily basis. They also had to repetitively manage (every year or academic term) the same documentation and correspondence, lists of students, lesson plans, class notes, documents used as visual aids during classroom presentations, exams and assignments, application forms for teaching assistants, bibliographic database, as well as numerous reports generated or received. Moreover, if professors work on a recurring basis with models or types of generic documents (e.g., lesson plans, forms, memos sent to students) and reference material on various topics (e.g., plagiarism), they are also called to deal with more specific documents (e.g., letters of invitation and thanks to guest speakers, lists of students, etc.).

Through our analysis, we found out that the personal classification schemes examined presented various approaches to organizing documents. An analysis of the concepts represented in the sample of folder names revealed seven main division criteria: 1) long-term *activity* (e.g., “Exam evaluation”); 2) *document type* (e.g., “Lesson plan”); 3) *document sender* (e.g., “Student”); 4) *document recipient* (e.g., “Director”); 5) *document version* (e.g., “Final”); 6) *university semester* (e.g., “Fall 2010”); and 7) *course number* (e.g., “SCIXXXX”).

The model concepts were also validated using the professors’ folder names to assess what terminology was used to explicitly refer to course-related concepts. This analysis was used as a validation method regarding the coverage of the faceted classification. The broad use of some classes representing multiple concepts – each expressing two or multiple document dimensions (such as a course code followed by a date) – corroborated the fact that using a single hierarchical, one-dimensional classification structure would not be sufficient to adequately support user needs.

Starting with the list of concepts identified through the business process and folder names analysis, we extracted recurrent values, that is, keywords and terms used in the folder names (e.g., “Lesson Plan,” “Director”) and clustered them into a limited number of classes already used by the professors (e.g., *Document Type*) or into new classes (e.g., *Document Recipient*) revealed by the folder names. Facets and related values were derived from these classes.

During the faceted classification model development, Ranganathan’s principles of facet analysis, refined by Louise Spiteri’s “Simplified Model for Facet Analysis,”<sup>40</sup> were used to determine the criteria for selecting and naming facets: 1) *differentiation* (criteria of organization between top-level facets and their sub-facets are distinctive and logical); 2) *relevance* (facets are expected to adequately reflect the purpose, subject, and scope of the classification system); 3) *ascertainability* (facet names are simple and circumscribed); 4) *permanence* (facets represent permanent qualities of the item

40 Louise Spiteri, “A Simplified Model for Facet Analysis,” *Canadian Journal of Information and Library Science*, vol. 23, nos. 1–2 (1998), pp. 1–30.

being divided); 5) *homogeneity* (top-level facets are homogeneous [i.e., they are at the same level of granularity]); and 6) *mutual exclusivity* (characteristics of division between facets are mutually exclusive [i.e., each facet describes one single aspect of a document]).

Following the analysis of the activities performed by the two professors and the main concepts they used to classify documents related to the design and delivery of their courses, a classification model was developed to offer them the possibility of classifying documents according to several facets: generic facets (relevant to all documents) and specific facets (relevant to documents in a specific context) were identified (see Table 1).

**Table 1: Generic and Specific Facets for Course Design and Delivery**

<b>Generic Facets</b>	<b>Specific Facets</b>
<b>Actor</b> (Employee name, surname)	<b>Document version</b> (e.g., preliminary, final)
<b>Position</b> (e.g., professor)	<b>University Semester</b> (e.g., Fall 2010, Winter 2011)
<b>Department</b> (e.g., School of Library and Information Science)	<b>Course Number</b> (e.g., ARVXXXX, SCIXXXX)
<b>Activity/Task</b> (e.g., development of a course, exam evaluation)	<b>Document Recipient</b> (e.g., student, director)
<b>Document Type</b> (e.g., class notes, lesson plan)	<b>Document Sender</b> (e.g., teacher, student)

The identification of these facets and their related values had also required the normalization of the vocabulary associated to each facet. For the design and delivery of a course, questions such as the following were examined: 1) For the activity of correcting student work and examinations, should we focus on “correcting” or “evaluating?”; 2) For documents that are given to students in the case of assignments and examinations, are these statements or guidelines?; 3) For communication with lecturers, teaching assistants, or students by email, should we talk of correspondence, letters, emails, or messages? Such questions, highlighted during the design of faceted classification schemes, reveal the implication of each individual perspective in the development of a classification system.

Once the vocabulary was standardized, semantic relationships were added to the ISIS faceted model. Based on the different components of the process map, semantic relationships are persistent associations between facet values and/or a set of facets. Following a user’s selection, or based on the identity of the user, the system will restrict the display of the facet and facet values available for tagging to each user. For instance, the organizational structure can be

deduced from the identity of the user, and the roles and tasks presented can be limited to those that he/she actually performs. This significantly simplifies the complexity of selecting values during tagging.<sup>41</sup>

The process of developing an ISIS faceted classification model for the small private IM consultancy company was very similar. This second model was designed by one of the authors of this paper who is also a consultant for the private organization that participated in the study. An analysis of the activities/tasks performed during consultancy with clients was carried out during one focus group involving three consultants (including the author). A detailed process map was developed to identify the facets and their related values. Although the analysis focused on the process of delivering consulting services, a higher business model was also created to identify the potential processes to be developed for this company; the resulting model comprised the facets found in Table 2.

**Table 2: Generic and Specific Facets for Delivering Consulting Services**

<b>Generic Facets</b>	<b>Specific Facets</b>
<b>Actor</b> (Employee name, surname)	<b>Document version</b> (e.g., preliminary, final)
<b>Position</b> (e.g., business analyst, consultant)	<b>Topic</b> (e.g., information architecture, file plan development)
<b>Branch</b> (e.g., business development branch, consulting services)	<b>Project Title</b> (e.g., Development of an information architecture for client X)
<b>Activity/Task</b> (e.g., receive and analyze requirements grid from consultants)	<b>Project Start Date</b> (e.g., March-2010)
<b>Content Type</b> (e.g., deliverable, project plan, request for proposal)	<b>Project End Date</b> (e.g., June-2011)
<b>Client</b> (e.g., Company X)	<b>Project ID</b> (e.g., PXXXXX)

While the business areas associated with a university and a private consulting firm differ, it is interesting to note the similarities between the two models. In fact, it appears that most of the generic facets that can be used to describe a document are similar: actor; position; reference to the organizational structure (department or branch); activity/task; and document/content type. This finding

41 Alberts, Schellinck, Eby, and Marleau, "Bridging Functions and Processes for Records Management!"

is consistent with the ISIS approach, which proposes as a guideline a *meta-model*: an explicit model of the core facets and the semantic relationships needed to build a specific ISIS model in any organization. Following this perspective, each organization has its own ISIS model, even if the meta-model remains very similar.

Prior to the phase of the study involving the respondents, the researchers performed a final validation of both the course design and delivery model, and the consulting services model to ensure the consistency of the vocabulary and the validity of the semantic relationships.

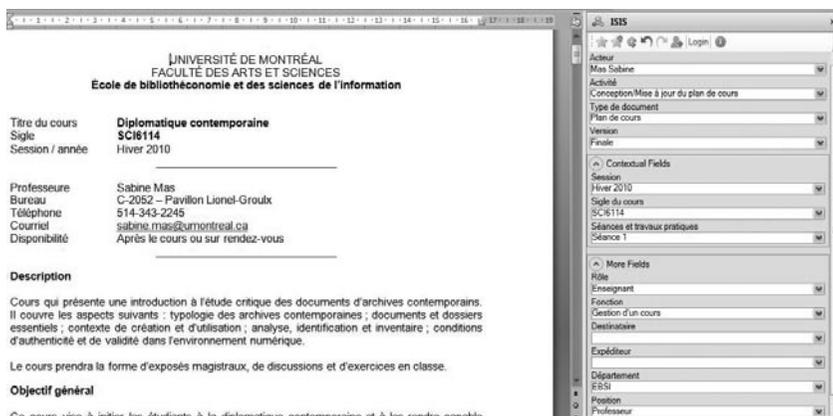
### ***Data Collection Methodology***

The study adopted a qualitative and naturalistic approach to investigate users' experience with faceted classification.<sup>42</sup> The sample of eight respondents consisted of two university professors and six information professionals, among whom three co-authored this paper.<sup>43</sup> Because the pilot project was conducted within the organizations that employ the researchers and authors of this paper, the respondents were willing to test the prototype and follow the protocol for data collection. For logistical reasons, the pilot involving the professors was conducted over three weeks while the pilot for the information professionals lasted one week. The respondents were free to withdraw at any time during the experiment. In order to maintain the confidentiality of respondents, we refer to them using alphanumeric codes when we quote their words.

During the testing period, respondents were asked to organize (create different virtual views of) within the faceted classification scheme any new electronic records produced or received in the course of their work. For the professors, these documents were restricted to the business process of course development and delivery; for the information professionals, these documents were restricted to the business process of delivering consulting services. ISIS, a Web 2.0 software solution, was used to represent, validate, and co-build the newly developed faceted models. The models were displayed to the respondents using the ISIS Classification Module. Integrated as a side panel within MS Office Suite applications, this module allows users to classify documents by selecting the appropriate facet values (see Figure 3). In case of any terms missing in the model, the respondents were invited to make structural and conceptual changes from their own perspective using the ISIS software "Suggest a New Term" functionality.

42 Norman K. Denzin and Yvonna S. Lincoln, eds., *The Landscape of Qualitative Research: Theories and Issues* (Thousand Oaks, CA, 2003), pp. 4–5, 29–30; Michael Quinn Patton, *Qualitative Research and Evaluation Methods* (Thousand Oaks, CA, 2002), pp. 39–43.

43 Two authors were in the Professors group; the third author was in the Information Professionals group.



**Figure 3: ISIS Classification Module.** Credit: used with the permission of Cogniva.

Tables 3 and 4 provide an example of how a document was tagged and thus organized, using the facets available in the classification panel.

**Table 3: Example of Document Tagged Using the University Faceted Classification**

<b>Description:</b>
<i>PowerPoint presentation on the lecture related to the archival discipline provided in the course ARVXXXX during Winter 2011</i>
<b>File Title:</b>
ARVXXXX_W11_03_[Archival Discipline]_students.pptx
<b>Facets and Values:</b>
ACTIVITY: Course Management DOCUMENT TYPE: Lecture Notes ACADEMIC TERM: Winter 2011 RECIPIENT OF THE DOCUMENT: Students COURSE CODE: ARVXXXX SESSION: Lecture #3

**Table 4: Example of Document Tagged Using the Consulting Company Faceted Classification**

<b>Description:</b>
<i>Statement of work received from [Name of the Client] for the development of their file plan.</i>
<b>Facets and Values:</b>
ACTIVITY: Project Planning CONTENT TYPE: Statement of Work CLIENT: [Name of the Client] TOPIC: File Plan Development PROJECT TITLE: Development of a File Plan for [Name of the Client]

The diary method, combined with a questionnaire, was used for data collection. Diaries allowed respondents, at the end of each day, to document the actions they performed in relation to the phenomenon under study, and to express their thoughts on these actions.<sup>44</sup> This method had the potential to “capture” the events and actions as they occurred, and to keep track of them during a continuous period of time. In order to standardize the types of data to be collected, the diary was structured as a questionnaire comprising both close-ended and open-ended questions with predetermined form.<sup>45</sup> Close-ended questions were answered using a selection from multiple choices (e.g., How many documents have you filed today using the faceted classification scheme?), or a simple piece of information (e.g., Would you prefer using folders or facets for classifying your documents?). Open-ended questions were used to obtain more information on the experience of classification for each respondent: reflections on the components of the faceted classification model, the impact of classification on the job performance, perceptions of difficulties while using the faceted classification, desired improvements, etc.

Respondents were asked to fill out the daily questionnaire at the end of each work day in which they were involved in the organization (by tagging) of documents newly created or received. At the end of the pilot, a total of 46 daily questionnaires were completed (20 questionnaires by the two professors and 26 questionnaires by the six information professionals). The unit of analysis was the completed daily questionnaire, each showing a daily experience of classification.

At the end of the study, a final questionnaire, consisting mainly of open-ended questions, was completed by each respondent so as to obtain a broader

44 Louise Corti, “Using Diaries in Social Research,” *Social Research Update* 2 (1993), <http://sru.soc.surrey.ac.uk/SRU2.html> (accessed on 25 June 2011).

45 Patton, pp. 342–48.

perspective on their classification experience. For reasons of availability, only five respondents (two professors and three information professionals) were able to fill out the final questionnaire. Because of the geographical distribution of respondents between two organizations, the daily and the final questionnaires were available online (using Survey Monkey) and were self-administered. Respondents could choose not to respond to one or more questions. Both questionnaires were available in French and English, given that four respondents were French-speaking (two professors and two information professionals) and four were English-speaking (four information professionals).

### *Limits of the Study*

Before analyzing the findings, it is important to note the limitations and biases of this study. First, there were few respondents, and the time of the experiment ranged from one week (for the information professionals) to three weeks (for the professors). Since respondents were working in two different professional backgrounds, they have used two different faceted classification schemes corresponding to two distinct business processes. Comparing two different business processes in two different settings may constitute a limitation of the study. However, studies of information behaviour often relate situations in which respondents describe behaviours in settings (work or everyday life) that vary from one respondent to another. Additionally, the fact that the experiment was conducted by different teams with an unequal number of respondents (two professors on the one hand and six information professionals on the other) using different models, corresponds to the reality of an organization where various processes and sub-processes are represented in different branches of a faceted classification scheme.

During the period of the pilot, some documents might have been used several times by a respondent, even if they were created and classified only once. For example, a document such as “Course Notes” was revised many times during the week, even though it was tagged once using the faceted classification. As part of their daily activities, respondents were also working on other business processes outside the study. Consequently, their focus was not solely on the pilot project. Also, in the information professional group, a respondent was not able to participate in the study until the end because of other priorities. This explains that in total, forty-six daily questionnaires were collected from eight respondents for the duration of the pilot. A further limitation is that the final questionnaire was completed by only five respondents. Nevertheless, conducting the experiment with a small sample permitted the inclusion (in the daily and final questionnaires) of a number of open-ended questions that provided rich, multiple perspectives. This additional qualitative data provided important insights into respondents’ perceptions and use of the

faceted classification scheme. However, these factors limit both the transferability and generalizability of the results of the study.

The final source of bias is the respondents' knowledge of the principles inherent to ISIS and the faceted classification approach. The three researchers not only contributed to the mapping process and the development of the faceted classification model, but also participated in the study as respondents. These factors, which are *a priori* a source of bias, have contributed to reducing the technical and ergonomic problems associated with the use of a specific tool, which was not within the scope of this pilot. We therefore believe that respondents were able to concentrate more on the conceptual dimensions of the faceted model. The quality of the results obtained during the data collection process confirms this assumption.

### **Findings<sup>46</sup>**

The results are the product of descriptive statistics (for close-ended questions) and content analysis (for open-ended questions). The percentages presented in the results are calculated from the close-ended questions collected in the 46 daily questionnaires completed by respondents: 20 daily questionnaires completed by the two professors, and 26 daily questionnaires completed by the six information professionals. Content analysis was done on the open-ended questions collected in both the daily and the final questionnaires.

To answer the research questions, we present the respondents' comments on the following issues: 1) the difficulties and dissatisfactions perceived while organizing their electronic documents with the faceted classification scheme, and the suggested changes to improve this scheme; 2) the elements which, according to respondents, facilitate document organization and are considered as sources of satisfaction; and 3) the respondents' perceptions of the usefulness of the faceted classification scheme as it relates to both their work performance and productivity, and their preferred method for organizing records (hierarchical or faceted).

#### ***1) Difficulties and Dissatisfactions Perceived while Organizing Documents, and Suggested Changes***

Of the 46 responses obtained ( $n = 46$ ) on the number of documents that were classified daily during the study, the majority of the responses shows that respondents classified between 1 and 5 documents (71.74%) and between 6 and 15 documents (17.39%) using the faceted classification scheme. In 5 cases

46 In reporting the findings, the authors have translated into English any comments from respondents that were originally in French.

(10.87%), respondents explicitly stated they had not classified any new document at the end of their working day. The results are similar for both organizations under study (see Table 5).

**Table 5: Number of Documents Tagged Using the Faceted Classification**

<b>Number of Documents Filed Daily Using the Faceted Classification</b>	<b>Professors</b>	<b>Information Professionals</b>	<b>Total Responses</b>
<b>0 Documents</b>	2 (4.35%)	3 (6.52%)	5 (10.87%)
<b>1–5 Documents</b>	17 (36.96%)	16 (34.78%)	33 (71.74%)
<b>6–15 Documents</b>	1 (2.17%)	7 (15.22%)	8 (17.39%)
<b>16+ Documents</b>	0	0	0
<b>Total Responses</b>	20 (43.48%)	26 (56.52%)	46 (100%)

For the duration of the pilot study, the 46 responses also show that respondents estimated that tagging their documents was extremely easy (15.22%), fairly easy (41.30%), neither easy/difficult (19.57%), somewhat difficult (13.04%), or have not been able to decide (10.87%) (see Table 6). These results are relatively similar for the respondents of both organizations.

**Table 6: Perceived Ease/Difficulty Using the Faceted Classification**

<b>Perceived Ease/Difficulty Using the Faceted Classification</b>	<b>Professors</b>	<b>Information Professionals</b>	<b>Total Responses</b>
<b>Extremely Easy</b>	2 (4.35%)	5 (10.87%)	7 (15.22%)
<b>Fairly Easy</b>	10 (21.74%)	9 (19.57%)	19 (41.30%)
<b>Neither Easy/Difficult</b>	3 (6.52%)	6 (13.04%)	9 (19.57%)
<b>Somewhat Difficult</b>	3 (6.52%)	3 (6.52%)	6 (13.04%)
<b>Extremely Difficult</b>	0	0	0
<b>Do Not Know</b>	2 (4.35%)	3 (6.52%)	5 (10.87%)
<b>Total Responses</b>	20 (43.48%)	26 (56.52%)	46 (100%)

Notwithstanding the evolution in the learning curve and increased overall familiarity with the new classification scheme, 13.04% of the respondents still considered, at the end, that classifying documents with the faceted classification is rather difficult, and 10.87% did not know how to categorize the level of ease or difficulty in doing so. It would be interesting to further examine the user learning process while interacting with faceted classification schemes for a longer period of time.

Among the main challenges expressed by the respondents, the 21 comments (8 for the professors and 13 for the information professionals) are related to: 1) the relevance of the categories and the facet naming; 2) the semantic relationships between the facet values; and 3) the fact that users had to select many different values to tag a document. The same types of comments were made by both the professors and the information professionals, indicating similar attitudes and perceptions for both groups of respondents. These challenges are further detailed below.

#### *Relevance of the Categories and Facet Naming*

- For some facets, the values offered were not adequate or sufficient to classify certain documents. For instance, in the Document Type facet, a respondent indicated the need to distinguish between “Lecture

Notes” and “PowerPoint Presentation,” which are two different, though complementary, documents in the course development task. Two respondents also indicated that due to values that seemed rather inadequate, they sometimes felt they were making subjective choices.

- The comments of 15 respondents focused on the need to use values that would be more relevant or “best suited” to their work, as well as the need for facet names that would be more specific and meaningful to describe the content of a document. For example, content related values should include more specific elements relating to the subject taught during a lecture (e.g., Archival History) rather than a lecture number, minimally meaningful to describe the content of the material taught during a course session (e.g., Lecture #3). The main facets for which respondents would like values more suited to their work are related to the facets Activity, Sender, Recipient, and Type of Document.
- One respondent mentioned that it would be convenient to have a glossary to define the values of the facets whose names were not easily understood. In the same vein, a respondent revealed the need to develop a thesaurus to be able to use “equivalent” terms for the same value. For example, the respondent mentioned the difficulty of distinguishing between the Role (e.g., teaching) and the Position (e.g., professor) facets. Another example mentioned focused on the nuances between the values “evaluation” and “correction” in the Activity facet.

### *Semantic Relationships between the Facet Values*

- When a value was deemed inadequate or missing, the respondents were able to override the semantic relationships to allow the display of all the values available for a facet, or to suggest a new value (a new term) to be associated with a facet. On this matter, one respondent mentioned the difficulties encountered while adding values to certain facets. Faced with these difficulties, the respondent was no longer motivated and as a result, stopped classifying his documents. Another respondent also mentioned that the ability to override the system adds extra work; he/she recommended developing more comprehensive semantic relationships between the facets to “avoid clicks” and simplify the classification of documents, since values that are associated can automatically be selected by the system.
- Within ISIS, a value pertaining to a specific facet may appear by default when selecting a value pertaining to a more generic facet. For example, after selecting the value “Documentation” in the generic facet “Document Type” (a thematic report the professor wrote in another context and from which content is reused to develop a course), the value “Student” is displayed by default in a facet called “Recipient.” Two

respondents mentioned that some of the semantic relationships linking facets, as well as the values associated with these facets, should further be refined.

### *Selection of Many Different Values to Tag a Document*

- Two comments related to the difficulty in identifying the relevant values in some of the facets due to the long list of available terms. One respondent noted that it took considerable time to find the most relevant value for his work while he initially wanted to send a very quick email. In order to avoid scanning long lists of values, another respondent suggested that combining similar values (for instance, by grouping all the forms under the same heading), would be beneficial.
- From the respondents' comments, it appears that many documents could, theoretically, be classified according to two or three additional values in a particular facet. The comments show that choosing which one would best apply was a major challenge. For these documents, respondents said they would like to be able to select all values that apply instead of having to make a single choice. Using the "Activity" facet for instance, the statement of course work, the list of students participating in a field visit, or an email sent to students as a reminder about this visit, could all be classified with the same three values: design of a course session, design of a course requirement, and visit management. In reality, all these values are accurate since the visit takes place within the context of a course session and will result in an assignment.

### **2) *Elements that Facilitate Respondents' Document Organization and are Considered as Sources of Satisfaction***

As demonstrated in Table 6, the respondents expressed some difficulty in classifying their documents according to the faceted classification. Nonetheless over half of the 46 responses showed that 56.52% of respondents found the experience of classification "extremely easy" (15.22%) or "fairly easy" (41.30%). These results are most likely due in part to the necessary learning curve, and acceptance and use of the facets, as indicated by the comments of three respondents (e.g., "The level of familiarity with the classification plan increases over time" [A-1]). Thus, the difficulties were gradually reduced during the pilot project. Once again, it would be relevant to further examine the respondents' learning process over a longer period of time.

Respondents were asked to name the elements that facilitate the tagging of their documents with the faceted classification scheme. One of the most salient results is that some facilitating factors and sources of satisfaction

identified by certain respondents were considered to be sources of difficulties for others. We see this as a limitation of our study, which is based on a small sample and therefore does not permit generalization, but rather identifies these facts as avenues for future research. On the whole, comments pertained to the flexibility of facets and their related values as well as the use of semantic relationships. These are further described below.

First, although some respondents report difficulties and dissatisfactions in the use of the faceted classification attributed to the inadequate or insufficient values, or to the presence of lists of values that are too long, five comments show that the respondents particularly enjoyed the customization of values available, which reflected their own documents and tasks, the use of familiar terminology tailored to their work, and the low numbers of values and facets allowing for easy scanning and identification of terms while tagging a document. One respondent affirmed that one of the tool's strengths relied on the ability to more easily exploit the content of documents for classification and retrieval of documents.

Also, four comments provide evidence that the respondents particularly appreciated the flexibility of the faceted classification scheme in terms of the ability to add, under each facet, values that were more "refined" and more adapted to their reality in carrying out business processes. This flexibility – which is not desired by some respondents who want to describe their documents as quickly as possible by selecting a value from a predefined list – allows others to customize the classification system to better meet their specific needs. Furthermore, other facilitating factors identified by the respondents pertained to the interface used to view the faceted classification scheme and classify documents. For instance, two comments mentioned the drop-down menus that "facilitate viewing the facets and the choices of values available in every facet" [A-1], as well as the possibility to include default values, and save favourites. Two additional comments relate to the usefulness of favourites, i.e., the ability to save a set of facet values and apply them all at once. Finally, two comments relate to the usefulness of semantic relationships between facets or facet values. For instance, one respondent appreciated that these relationships indicate which facets are essential to tag a document: "The links between who I was, my Roles, Tasks, and content Types limited the values to those I needed" [B-5].

### **3) *Perceptions of the Faceted Classification Scheme Utility, and Preferred Records Organizing Method***

Another set of questions related to the overall perceptions of the faceted classification scheme (e.g., ease of use, utility, and helpfulness) (see Table 7).

**Table 7: Overall Perceptions of the Faceted Classification Scheme**

<b>In the Context of your Work Activities and Thinking of the Documents you Have Tagged (Filed) Today, you Find that the Faceted Classification Scheme:</b>	<b>Professors</b>	<b>Information Professionals</b>	<b>Total Responses*</b>
<b>Is Easy to Use in a Work Context</b>	7 (12.07%)	12 (20.7%)	19 (32.77%)
<b>Improves your Performance at Work</b>	-	6 (10.34%)	6 (10.34%)
<b>Helps to Better Manage your Documents</b>	-	10 (17.24%)	10 (17.24%)
<b>Should be Improved</b>	7 (12.07%)	6 (10.34%)	13 (22.41%)
<b>Do Not Know</b>	6 (10.34%)	4 (6.9%)	10 (17.24%)
<b>Total Responses</b>	20 (34.48%)	38 (65.52%)	58 (100%)

\*Note: Respondents could select more than one answer to this question; this explains why  $n = 58$ .

Overall, in the context of their daily activities, respondents considered that the faceted classification scheme is easy to use (32.77% of the responses). However, several respondents argued that some aspects of the faceted classification should still be improved (22.41% of the responses). This is not surprising, considering that users needed to become familiar with, and adapt to, the new model (e.g., by searching for values and adding them in some facets), as demonstrated by this comment: “I found that using the faceted classification system got much easier over the course of the time period, as I got better at remembering that I could classify documents, and also because by the end I had entered values into the system that I knew about and could look for easily. The general idea of using facets to classify documents was straightforward” [B-6].

Some respondents expressed doubts relative to the improvement of their job performance (10.34% of responses) using faceted classification. In the case of respondents who wanted to send emails as quickly as possible due to time constraints, two indicated that classifying emails using facets slowed them down in their work. In some cases, they had already sent the emails before

they remembered that they should have added the metadata: “Today I was working under a lot of time pressure, updating report deliverables and sending numerous emails. I only remembered that I could classify them after I had sent the report and the emails, so I didn’t go back and classify them” [B-6]. One respondent also suggested using the software to incorporate a tagging reminder function or a warning that would be triggered upon closing a document without having added the required facet values. Another respondent revealed that work performance decreases while tagging: “... my performance at work is slow because I have to classify a document. ... I am interrupted in my work when I need to focus on my task, I have to stop suddenly and do something else” [B-1].

The responses show that opinions are divided on the use of folders or facets for classifying documents, in both study groups. Two respondents indicated a preference for organizing documents according to folders rather than facets: “My hierarchical classification has proven its effectiveness and I find everything very well. The classification structure of the documents for a course is indeed self explanatory” [A-2]; and “Honestly, I wonder if it improves my performance at work because I’m already well organized in my folders” [B-1]. In another comment, one respondent stated that using facets would be a complementary solution to organizing documents into “more intuitive” folders, while overcoming the “one dimensional” view imposed by a traditional, one-dimensional folder hierarchy. And one respondent emphasized that the faceted classification could prove its full utility in the context of shared documents, stored for instance on a common server by all team members. (For the pilot study, the documents were saved on individual workstations.)

In addition, one respondent established an important link between the choice of documents to be classified according to the faceted classification and the archival value of these documents: “... I do not always know what document I need to keep, what is important to me in my daily work is not necessarily of great interest to my organization. There is no defined policy on the way in which I have to use the tool, what I have to classify or not” [B-1]. This comment demonstrates the need to establish both the business and archival value of administrative documents, while attempting to reconcile these perspectives in the choice of facets and values available to users.

## Discussion

The results represent the experience of a small group of users in the daily use of a faceted classification scheme to organize their electronic documents. These results, which involve a very small sample, have documented this specific experience, opening new avenues for future research.

Several respondents mentioned that it should be possible to categorize a document according to more than one value pertaining to the same facet, in

order to improve both classification and future document retrieval. This raises the following questions: Should faceted classifications be conceived on the same basis as hierarchical classifications in which a document is usually classified into one single folder? In the case of faceted classification, and considering that facets will also support the search function, is it relevant to tag a document using multiple values within one facet?

Other comments were related to the possibility provided by the faceted classification scheme to add values, under each facet, according to the work conducted by each respondent. Associated with this flexibility is the risk of producing a classification system that is too complex and difficult to understand and use, particularly for organizations with high staff turnover. It is therefore important to establish validation rules when proposing to add new facet values to enrich the classification system. As a tool, ISIS embeds governance mechanisms in order to maintain the balance between user preferences and customization needs, and the willingness of the archivists in reaching a more generic model that would be more representative of the organization.<sup>47</sup>

Moreover, the collaborative principle inherent in the use of ISIS for faceted classification tagging allows users to develop new information competencies (such as describing, indexing, and classifying documents) even if these competencies are not always sought in a busy work context, as emphasized earlier in this paper. Allowing users to suggest the addition of new facets or new terms supports the process of co-construction and constant updating of the classification model. The employee is thus actively contributing to the development of information systems, placing the user at the centre of the electronic architecture; the user is not only the consumer of a system, but also a contributor and initiator of new functionalities.<sup>48</sup>

Mas et al. recognized the problems associated with a unilateral (designed by one expert), unique, and formal approach to describing and organizing complex documents.<sup>49</sup> The discourse universe of the “archivist-classificationist” does not always coincide with that of the authors of the document or the users. This type of approach is mainly intended for categories of employees with similar work objectives but understanding a personal semantic related to their domain of expertise. It is not, however, suited to situations in which there is a range of different information needs, because knowledge and references

47 Alberts, Schellinck, Eby, and Marleau, “Bridging Functions and Processes for Records Management.”

48 Louise Merzeau, “L’intelligence de l’usager,” in *L’usager numérique* (Paris, 2010), pp. 9–37.

49 Sabine Mas, Aurélien Bénel, Jean-Pierre Cahier, and Manuel Zacklad, “Classification à facettes et modèles à base de points de vue: différences et complémentarité,” *Proceedings of the 36th Conference of the Canadian Association for Information Science, June 5–7, 2008, University of British Columbia, Vancouver, Canada* (2008), [http://www.caais-acsi.ca/proceedings/2008/mas\\_2008.pdf](http://www.caais-acsi.ca/proceedings/2008/mas_2008.pdf) (accessed on 28 June 2011).

can be very specific and scalable.

Within the document information conveyed explicitly or implicitly, structured or unstructured, we cannot fully predict which will present relevant search or referencing criteria for different employees. In a more decentralized context, designers of faceted classification schemes could benefit from the user's point of view; user participation in defining classification categories as well as which themes to classify is not new. The innovative aspect of the ISIS approach is the inclusion of collaborative and participative modalities, by which individuals can play an active role in the referencing of documents they create, receive, and use as part of their processes; this would greatly improve user adoption of the faceted classification.

One reason respondents continued to favour folders over faceted classification may be related to a natural resistance to change when encountering any new system, in conjunction with their habits of classifying their documents according to a folder hierarchy available from their desktop.<sup>50</sup> As mentioned earlier in the literature review, employees generally develop personal hierarchical structures to organize their documents and folders; this way of proceeding is perceived as stable and self-explanatory, and allowing efficient and effective document retrieval. Factors which are both cognitive and affective will influence the choice of the personal structures with which employees are familiar. In a work environment, personal classificatory structures have cognitive benefits for employees because they allow them to maintain control of their documents (e.g., by ensuring they are all in "one place") and establish links between these documents.<sup>51</sup> Consequently, for the respondents of this study, it would appear that the faceted classification scheme has to move *beyond* traditional hierarchies for users to be satisfied.<sup>52</sup> Another reason that may explain the preference for the use of folders is most likely related to the fact that the search module was not available during the pilot, which certainly had an impact on user perception of the classification system utility.

The study focused on the *use* of the faceted classification scheme for classifying documents and not on their *retrieval*. Nevertheless, the perceived ease of use and speed in learning faceted classification for multiple-concept document tagging was also an interesting outcome of the pilot study and highlighted the value of a faceted classification system for organizing administrative

50 M. Lynne Markus, "Power, Politics, and MIS Implementation," *Communications of the ACM*, vol. 26, no. 6 (1983), pp. 430–44; James A. O'Brien, *Introduction to Information Systems: Essentials for the E-Business Enterprise* (Montreal, 2003).

51 William Jones, Ammy Jiranida Phuwanartnurak, Rajdeep Gill, and Harry Bruce, "Don't Take My Folders Away!: Organizing Personal Information to Get Things Done," *CHI'05 Extended Abstracts on Human Factors in Computing Systems, April 2–7, 2005, Portland, Oregon* (2005), pp. 1505–08.

52 Mas, "Schémas de classification et repérage."

documents.

Furthermore, pilot study respondents did not have to work in a collaborative environment such as a common server, where issues related to the classification and identification of shared documents would be much more significant. Similar challenges can be related to change management when positions characterized by a high level of staff turnover require that a classification structure developed by an employee be efficiently transferred to his successor. Usability factors related to the movements of personnel and the sharing of documents among team members to achieve a common business process, would reinforce the need for a faceted classification structure as a complement (or possibly as a substitute) to a conventional hierarchical structure. In that sense, this research paper provides an interesting research avenue into the pertinence of supplementing traditional hierarchical classification schemes with a faceted classification approach for records organization in a decentralized digital environment.

## **Conclusion**

This research study investigated the feasibility of using faceted classification for records organization in electronic work environments. The design of two faceted classification models for the purpose of this study was based on an analysis combining a top-down and a bottom-up approach. Facets were identified and selected from a set of conventional dimensions (document author, recipient of the document, document type, dates, places, names of people, etc.). The aim of this approach was to propose a classification scheme that was logical, referential, and stable over time while presenting a limited number of predefined facets, and providing a classification that was both multidimensional and standardized for actors performing the same job within the same organization. The comments received on the conceptual design were generally very positive and encouraging. Meanwhile, the respondents made relevant comments and suggestions to improve the proposed faceted classification system.

In this sense, our research reveals that users can feel involved in the development of a classification structure, which is perhaps one of the ways to overcome the lack of motivation that normally characterizes the classification of documents in organizational contexts. In fact, the traditional records management discourse associated with one organization's need to comply with legislative requirements should be further adapted to the need for users to perform their daily activities. One of the major hurdles to overcome is the perception that employees' work is personal or belongs to their work unit, rather than pertaining to a broader organizational agenda. The individualism that characterizes the classification of organizational documents should motivate records managers and archivists to rethink their methods in order to provide solu-

tions that would not only standardize the personal organization of electronic records, but also prevent problems related to the retrieval of this information.

For that reason, we encourage a paradigm shift with the adoption of a more egalitarian approach whereby the user becomes a partner with the records manager in the construction and updating of classification schemes. As would be the case in the Web environment, the results of this study point out, however, that all users do not necessarily want to be contributors; the user of a faceted classification scheme may be reluctant to contribute to the classification in a context where he has to work under pressure. Hence the importance of demonstrating the usefulness of such a model to the user while offering a model that is as simple, specific, and comprehensive as possible in order to reduce the cognitive and emotional burden of any classification practice. In this perspective, the approach of organizing and describing electronic records that reflects a top-down strategy presents great value in a technological world that generally promotes bottom-up approaches.

The faceted structures developed for this pilot study are still evolving. The addition of several content facets along with semantic relationships is planned so as to improve the classification process. Furthermore, we are expecting to test the research function to assess if using more specific facets related to the document content could provide richer semantic descriptions going beyond records management; this would allow users to look for more complex information such as: Who were all the invited speakers within the past five years? What decisions have been taken on cases of plagiarism? What are the different formulas used in the appreciation letters to thank the guests?

A longitudinal research study involving more volunteers is required to provide reliable and consistent data on the classification structure's conceptual coverage, selection, ease of navigation, user friendliness, and flexibility of using faceted classification in a work setting. A study based on document and information research using a faceted classification scheme is also planned to empirically evaluate respondents' perception of the usefulness of faceted classification when used to search for documents.