

# Gérer son information personnelle au moyen de la navigation par facettes

*Florian Evéquo, Julien Thomet et Denis Lalanne*

DIUF, Université de Fribourg  
Bd de Pérolles 90  
CH-1700 Fribourg, Suisse  
{ florian.evequo, julien.thomet, denis.lalanne } @ unifr.ch

## RESUME

Cet article présente *Weena*, un système de navigation par facettes dans une collection personnelle, qui permet notamment d'accéder aux informations en passant par les personnes (facette sociale) et la période (facettes temporelles) en lien avec ces informations. L'évaluation du système montre de la part des participants une utilisation effective de ces facettes et un intérêt pour l'approche, ce qui désigne la navigation par facette comme une alternative viable et prometteuse à la navigation hiérarchique et la recherche textuelle, traditionnelles en gestion de l'information personnelle (GIP).

**MOTS CLES :** Gestion de l'information personnelle, GIP, navigation par facettes, social, temporel.

## ABSTRACT

This article introduces *Weena*, a personal information management (PIM) system enabling faceted navigation in a personal collection. Re-finding information items can be achieved in particular through the related people (social facet) and temporal period (temporal facets) in addition to traditional hierarchical browsing and text search. Participants in the evaluation effectively used those facets and expressed an interest for the approach. Faceted navigation is therefore a viable and promising alternative to hierarchical browsing and text search, the two more traditional re-finding means in PIM.

**CATEGORIES AND SUBJECT DESCRIPTORS:** H5.m. Information interfaces and presentation: Miscellaneous.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*IHM'2010*, September 20-23, 2010, Luxembourg, LU

Copyright © 2010 ACM ISBN 978-1-4503-0410-8/09/2010 ... \$10.00

**GENERAL TERMS:** Management, design, experimentation, human factors.

**KEYWORDS:** Faceted navigation, personal information management (PIM), social, temporal.

## INTRODUCTION

Le terme d'information personnelle désigne toutes les informations numériques qu'un individu possède et conserve pour son usage personnel et qui sont placées sous son contrôle direct [13]. La croissance régulière des capacités de stockage d'informations numériques entraîne de la part des utilisateurs une certaine prodigalité quant à l'archivage de leur collection d'informations personnelles. Rapidement, ces collections représentent une quantité telle d'information que leur ordonnancement et leur accès mettent au défi ses capacités cognitives. De fait, la métaphore du bureau, calquée sur le modèle physique d'un espace de travail disposant d'un bureau proprement dit où sont empilés des documents, et de dossiers d'archives, impose le classement de toute information numérique. Bien que cette métaphore soit quadragénaire, puisqu'Alan Kay en faisait la première présentation à Xerox PARC en 1970, aucune alternative sérieuse n'est venue jusqu'à ce jour contester son hégémonie. Pourtant, elle n'est pas sans défaut. Sans doute le plus saillant d'entre eux a-t-il trait à l'accès aux documents. Par analogie à un bureau traditionnel, en effet, un document n'est qu'à une seule place dans une hiérarchie unique. Retrouver un document classé au « mauvais endroit » peut ainsi se révéler une expérience coûteuse en temps et génératrice de frustrations. Bien sûr, les moteurs de recherche intégrés aux systèmes d'exploitation peuvent considérablement réduire l'effort requis. Néanmoins, ils exigent de la part de l'utilisateur de formuler des mots-clés précis, dont celui-ci pourrait ne pas se souvenir. Par ailleurs, des travaux récents en psychologie cognitive<sup>1</sup> mettent en lumière le fonctionnement de la mémoire. Pour un événement donné, cette dernière peine à se souvenir des expressions exactes le caractérisant. Par contre, elle conserve naturellement

<sup>1</sup>[6] en donne un résumé exhaustif.

un grand nombre d'informations plus floues, liées au contexte de cet événement. Quelles personnes sont concernées ? Quand a-t-il eu lieu ? Quelle en était la thématique générale ?

Forts de ce constat, nous proposons dans cet article un navigateur nommé *Weena* mettant à profit certaines des propriétés contextuelles, ou *facettes*, de l'information personnelle pour faciliter l'accès aux documents dans une collection personnelle. Une évaluation est ensuite présentée, qui permet de juger de l'adéquation de chacune des facettes à certains types de tâches d'accès, et de l'intérêt de la navigation par facettes pour la gestion de l'information personnelle.

Dans la première partie de cet article, la littérature scientifique en rapport avec la gestion de l'information personnelle et la navigation par facettes sont examinées. La partie principale de l'article précise les motivations sous-tendant la conception de notre système et décrit son interface. L'évaluation de notre approche de navigation par facettes est enfin exposée dans une dernière partie.

## RECHERCHE APPARENTÉE

**Gestion de l'information personnelle.** La gestion de l'information personnelle (GIP)<sup>2</sup> moderne commence avec la vision pionnière de Vannevar Bush [1]. Il décrit un appareil hypothétique, le *Memex*, grâce auquel un individu peut consulter « tous ses livres, archives et communications [...] avec une rapidité et une flexibilité exceptionnelles ». Cette première description informelle de la GIP met en évidence l'une des tâches cruciales qui la constituent : l'accès (par navigation ou recherche) aux éléments précédemment archivés. Nous présentons ci-dessous quelques travaux récents et innovants ayant trait aux interfaces d'accès à l'information personnelle.

Les premiers moyens d'accès à l'information personnelle sont les outils mis à disposition par les systèmes d'exploitation eux-mêmes, respectant la métaphore du bureau. Il s'agit de navigation dans une hiérarchie unique : le système de fichier. Cette méthode se complète souvent d'un moteur de recherche locale (*Spotlight*, *Windows Search*, *Google Desktop Search*, *Copernic Desktop Search*) qui, similairement à ses ascendants web, offre des outils de recherche textuelle. *Stuff I've Seen* [5] est le précurseur de tels moteurs de recherche, introduisant l'indexation de l'information personnelle, la recherche par mots-clés et des techniques de filtrage utilisant les attributs de fichier offerts par le système d'exploitation (date, type, auteur, etc.) au sein d'une interface de recherche textuelle. L'évaluation de ce système, à grande échelle (+200 personnes), procure des données intéressantes quant au nombre moyen de mots-clés et aux filtres utilisés.

<sup>2</sup>En anglais : Personal Information Management (PIM).

*Phlat* [4] constitue un développement de l'interface du précédent pour faciliter la construction dynamique de requêtes de recherche. Son évaluation indique un intérêt pour le filtrage par personnes et types de données au moyen de la nouvelle interface.

Plusieurs systèmes exploitent l'aspect temporel de l'information en tant que moyen d'accès privilégié [10, 16, 17]. Aucune évaluation ne valide malheureusement suffisamment cette approche.

*Forget-Me-Not* [14] se fonde quant à lui sur la psychologie cognitive, tirant parti du fait que la mémoire associe naturellement les éléments à des épisodes. Un épisode est caractérisé dans la mémoire par un certain nombre de propriétés, dont les personnes et les lieux qui lui sont liés, ce qu'il s'est passé, ainsi que la période approximative durant laquelle il a eu lieu. Le prototype développé vise à mettre à profit les capacités naturelles de notre mémoire épisodique en conservant un riche contexte pour chaque élément archivé. Malheureusement, il ne permet pas de gérer ainsi toute l'information personnelle, puisque l'accent est mis sur les communications uniquement (courriels, échanges de fichiers, téléphones, etc.), et n'a pas été évalué formellement. *ContactMap* [20] est une interface d'accès à l'information personnelle utilisant le réseau social de l'utilisateur. A travers une représentation graphique de ce réseau, l'utilisateur peut accéder aux courriels et historiques de clavardages (*chats*) échangés avec un individu ou un groupe de contacts. Les auteurs présentent ce système comme un moyen d'accès à l'information personnelle centré sur les personnes plutôt que les messages. L'évaluation montre un fort intérêt des participants pour cette approche, bien qu'elle soit à l'instar de *Forget-Me-Not* limitée à un sous-ensemble de l'information personnelle. En outre, le réseau social formant l'interface de *ContactMap* est construit essentiellement par annotation manuelle. De nombreux travaux ont enfin été consacrés à la gestion des courriels uniquement, mais leur énumération serait hors de propos. [13, 11] y dédient plusieurs chapitres.

**Navigation par facettes.** La classification par facettes est une invention du bibliothécaire et mathématicien indien Ranganathan [15]. Elle consiste à organiser les éléments à classer selon différentes dimensions indépendantes, nommées *facettes*, plutôt que dans un ordre taxonomique déterminé une fois pour toute. Chaque facette contient des aspects ou propriétés clairement définis, collectivement exhaustifs et mutuellement exclusifs de l'ensemble d'éléments à classer, nommés *catégories* de la facette. Une collection de livres peut ainsi être classée selon une facette « auteur », une facette « type », une facette « sujet », etc. Les facettes peuvent ainsi être comparées à des hiérarchies indépendantes. Cette classification permet en-

suite naturellement de parcourir et d'effectuer des recherches dans la collection en utilisant l'une ou l'autre des facettes en alternance, ce qui permet de parvenir aux mêmes éléments en suivant des chemins différents. Ce paradigme se nomme navigation par facettes ou recherche par facettes. Depuis la popularisation de cette technique par Hearst et son équipe [21], les sites internet qui la mettent en application sont légion (*Yahoo!*, *Flamenco*, d'innombrables sites de commerce électronique, etc.), mais les applications locales restent rares, à l'exception notable des systèmes de gestion de bibliothèques musicales (*iTunes* et ses alternatives).

Quelques prototypes utilisent ce paradigme en GIP. C'est le cas de *FacetMap* [18], outil généraliste, ou *PhotoMemory* [8], destiné à la gestion de photographies personnelles. Dans le premier, les facettes utilisées sont mises à disposition par le système d'exploitation, dans le second elle sont partiellement issues d'annotations manuelles. *Feldspar*[3] utilise lui aussi un principe proche *in fine* de la navigation par facettes, bien que son but premier soit de faciliter la création de requêtes textuelles de recherche dans une collection personnelle.

Notre interface d'accès à l'information personnelle décrite ci-après applique le paradigme de la navigation par facettes. Notre approche poursuit et dépasse les travaux précédents en ce qu'elle propose à la fois : (1) une unification de toute l'information personnelle (fichiers, courriels, événements de calendrier), (2) un choix de facettes personnelles fondé sur des études ou résultats empiriques reconnus, (3) une génération des facettes sans recours à l'annotation manuelle, et (4) une évaluation du système.

## CONCEPTION DU SYSTÈME

Afin de déterminer quelles facettes de l'information personnelle doivent être privilégiées, nous nous sommes appuyés prioritairement sur les résultats d'études précédentes, qui nous ont encouragés à sélectionner les facettes suivantes :

- le type de document [5, 4]
- les personnes (réseau social) liées au document [17, 14, 20]
- la date du document [10, 16, 17]
- la localisation du document dans la hiérarchie de fichiers [12]

La catégorisation d'un document conformément à chacune de ces facettes présente également l'avantage de pouvoir être obtenue automatiquement, sans annotation manuelle de la part de l'utilisateur.

En outre, il a été décidé d'intégrer un champ de recherche, puisqu'il est avéré que la recherche textuelle reste un moyen apprécié et utilisé de retrouver des

informations [2]. Bien que la recherche textuelle ne puisse pas être considérée comme une facette au sens strict, puisque l'ensemble des requêtes de recherche possibles ne constitue pas un ensemble de catégories mutuellement exclusives, nous la nommons tout de même « facette textuelle » dans la suite de ce travail par souci de simplification.

Finalement, les résultats d'un questionnaire en ligne<sup>3</sup> impliquant 147 personnes (55 femmes et 92 hommes) ont également apporté deux conclusions intéressantes relativement à notre choix de facettes : (1) les utilisateurs considèrent prioritairement leurs collections de documents sous l'angle de la hiérarchie de fichiers, ce qui nous a conforté dans l'idée de l'intégrer comme facette à part entière ; (2) leur rapport à l'aspect temporel de l'information est complexe, avec d'un côté l'utilisation assidue d'agendas logiciels et de l'autre une utilisation sporadique de critères temporels pour la recherche de documents, ce qui plaide pour l'intérêt d'une évaluation de pertinence de la facette temporelle dans les tâches d'accès à l'information.

## DESCRIPTION DU SYSTÈME

Le système de navigation par facettes que nous avons développé se nomme *Weena*. Il permet de naviguer et d'accéder à des documents dans une collection d'information personnelle.

Les facettes utilisées par *Weena* sont obtenues principalement à partir de l'indexation réalisée par *Windows Search*. L'interface de *Weena* (cf. fig. 1) se compose de trois parties principales décrites dans les sections suivantes.

**Facettes.** Sept facettes peuvent être utilisées pour naviguer dans sa collection d'informations personnelles : *textual search*, *file explorer*, *type*, *social network*, *date accessed*, *date modified*, et *date created*. Chaque facette, exceptée *textual search*, affiche les différentes catégories qui la constituent. La sélection d'une (resp. de plusieurs) catégorie(s) dans une facette filtre l'ensemble des documents personnels pour ne garder que ceux qui correspondent à la (resp. aux) catégorie(s) choisie(s). Les documents résultants sont affichés dans le panneau de résultat. Conformément au paradigme de la navigation par facettes, toutes les facettes sont mises à jour pour ne refléter que les catégories correspondant aux documents affichés dans le panneau de résultat. Les catégories disponibles dans une facette sont alors affichées en noir, alors que les indisponibles sont affichées en gris et ne peuvent plus être sélectionnées, ce qui interdit de naviguer vers une « impasse » où il n'y aurait aucun document.

Les catégories qui constituent chacune des facettes sont déterminées automatiquement à partir notam-

<sup>3</sup>Les résultats détaillés peuvent être consultés dans [19].

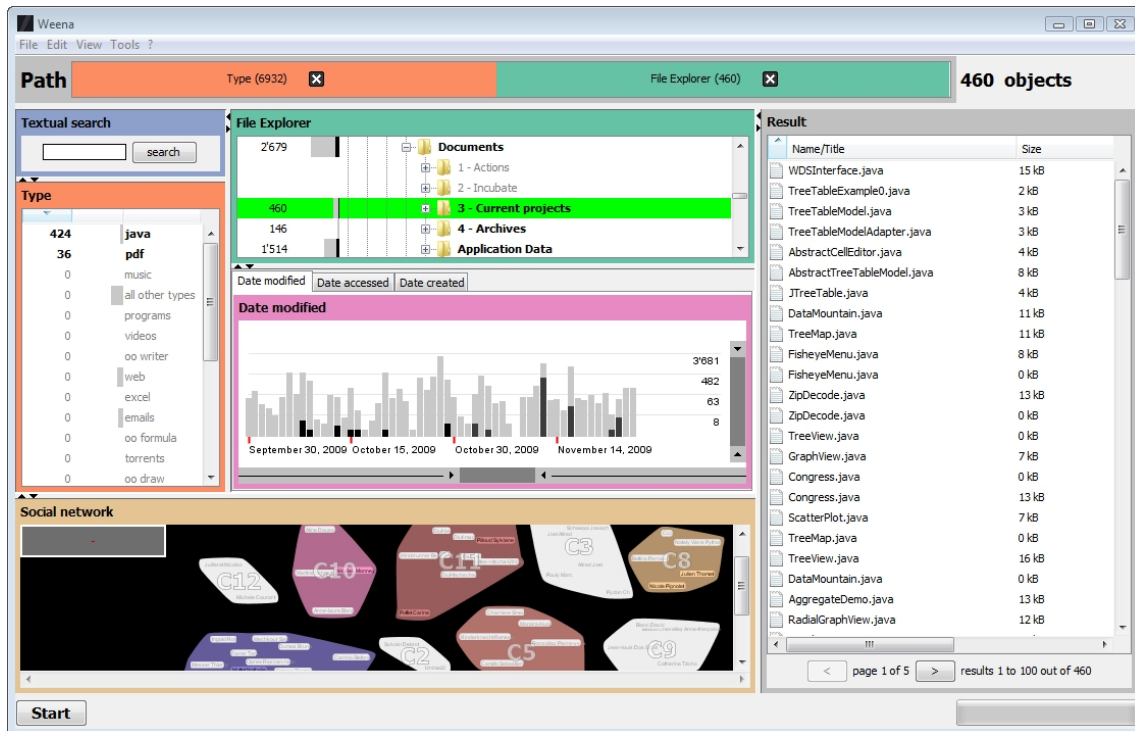


FIGURE 1 : L'interface de *Weena*. Chaque facette de l'information personnelle est représentée dans un panneau à bordure colorée. Dans cet exemple, l'utilisateur a tout d'abord retenu les documents *pdf* et *java* dans la facette *type*, puis a sélectionné le dossier *3 - Current projects* dans la facette *file explorer*. En haut, le fil d'Ariane (*Path*). A droite, le panneau de résultat (*Result*).

ment de l'index généré par *Windows Search* (par exemple les différentes dates pour les facettes temporelles, les différents types de fichier pour la facette *type*, ou encore le réseau social pour la facette *social network*). L'extraction du réseau social est réalisée à partir des archives de courriels de l'utilisateur. Les contacts sont ensuite regroupés automatiquement en communautés. Une description de la méthode utilisée dépasse le cadre de cet article [9]. Le tableau 1 documente les catégories constituant chaque facette ainsi que le type de représentation graphique de ces catégories.

Les facettes temporelles affichent des histogrammes qui indiquent le nombre de documents modifiés, créés ou accédés à chaque date. Les histogrammes des facettes *type* et *file explorer* informent l'utilisateur sur le nombre de documents correspondant à chaque type, resp. contenu dans chaque dossier, sous-dossier, etc. Tous les histogrammes sont formés d'une barre bicolore. La couleur noire représente le nombre d'éléments courants; la couleur grise, le nombre total de documents avant l'application de filtres. Le chiffre adjacent à l'histogramme indique le nombre d'éléments courants. La facette sociale représente le réseau social de l'utilisateur. Les noms de ses contacts y sont affichés et organisés spatialement sous forme d'un graphe, partagé en diverses communautés de per-

sonnes. Les liens entre les contacts (arêtes du graphe) sont masqués par défaut. Ils se révèlent lorsque la souris survole un contact.

**Fil d'Ariane.** À l'instar des fils d'Ariane présents sur la plupart des sites web, celui de *Weena* montre à tout moment à l'utilisateur le chemin parcouru, soit la séquence des sélections effectuées dans les diverses facettes. Il affiche, pour chaque filtrage appliqué, un panneau comportant le titre de la facette, sa couleur et mentionnant le nombre d'éléments issus du filtrage correspondant. Les filtres appliqués par une sélection dans une facette peuvent être annulés en utilisant l'icône en forme de croix affichée sur le panneau correspondant au filtre dans le fil d'Ariane.

**Panneau de résultat.** Le panneau de résultat présente les documents correspondant à l'application successive de tous les filtres actifs sous forme de tableau paginé doté des colonnes suivantes : type de document, titre, taille, dernière date d'accès et de modification et date de création. Ce tableau propose les interactions usuelles de tri par colonne et permet à l'utilisateur d'opérer sur les éléments de diverses manières à l'aide d'un menu contextuel (ouverture du document, suppression, accès au dossier parent, etc.).

**Cas d'utilisation.** Dupont désire lire la thèse que son ami Durand lui a envoyé par courriel il y a 6 mois.

Facette	Représentation graphique	Catégorie
<i>file explorer</i>	explorateur de fichier avec histogrammes	arborescence hiérarchique des documents
<i>textual search</i>	champs de texte	titre, contenu, auteur, etc.
<i>type</i>	tableau avec histogrammes	type de document
<i>date modified</i>	ligne temporelle avec histogrammes	date de la dernière modification
<i>date created</i>	ligne temporelle avec histogrammes	date de création
<i>date accessed</i>	ligne temporelle avec histogrammes	date du dernier accès
<i>social network</i>	graphe social avec communautés	auteur, participants

TABLE 1 : Catégories et représentation graphique des différentes facettes.

Il sélectionne donc « Durand » dans la facette sociale puis « pdf » dans la facette *type* pour filtrer les nombreux documents relatifs à cet ami. Le panneau de résultat affiche désormais 20 éléments, mais la thèse n'y est pas. Se souvenant de la date approximative de la réception du message, il décide d'annuler sa dernière sélection à l'aide du fil d'Ariane, puis de sélectionner le mois de septembre dans la facette *date created*. Le fichier désiré est maintenant visible dans le panneau de résultat : il s'agissait d'un document *Word*.

## ÉVALUATION

Il n'est pas anodin que seul un petit nombre de systèmes de GIP ont été évalués. En effet, plusieurs difficultés se dressent face au chercheur désirant évaluer un tel système. Premièrement, l'information est personnelle. Chaque personne possède et gère une collection unique, liée à sa propre expérience. Forcer les gens à utiliser un corpus commun d'informations dans le cadre d'une évaluation rend cette dernière peu réaliste. Deuxièmement, les utilisateurs ne sont pas prêts à partager toute leur information personnelle avec un évaluateur. Certaines parties sont et doivent rester privées. Enfin, les chercheurs ne peuvent pas connaître à l'avance les raisons qui poussent les gens à accéder à telle ou telle information personnelle. Il est donc extrêmement périlleux de définir des tâches qui soient pertinentes et réalistes pour tous.

La méthode d'évaluation proposée par [7] permet en partie de contourner ces problèmes pour évaluer une interface d'accès à l'information personnelle de manière qualitative. En effet, l'évaluation a lieu avec les données personnelles du participant et les tâches sont réalistes, puisqu'elles sont communiquées à l'évaluateur par le participant lui-même dans une phase préalable. Le protocole de notre évaluation est conforme à [7]. Nous complétons néanmoins la taxonomie de tâches définie par [7] avec les dimensions d'indices et ressources, décrites ci-après.

**Objectifs et méthodologie.** L'évaluation présentée ci-dessous vise à juger de l'adéquation de chacune des facettes à certains types de tâches d'accès, et à déterminer l'intérêt de la navigation par facettes pour la

GIP. De par son protocole, elle a une valeur essentiellement qualitative.

Dix participants (P1-P10) dont cinq femmes ont été recrutés pour cette évaluation. Ils sont âgés de 18 à 46 ans (moyenne : 28) et proviennent de domaines divers, principalement l'informatique (3), la psychologie (2) et les arts (2). Six sont étudiants à des degrés divers (du gymnase au doctorat), deux indépendants, une secrétaire et un chef de projet. Le volume de leur information personnelle est très variable : de 2'500 à 100'000 éléments (fichiers, courriels, événements de calendrier), pour une moyenne de 20'000 éléments, seront accessibles au moyen de *Weena* lors de la deuxième phase de l'évaluation.

La première phase de l'évaluation est une *diary study*. Chacun des participants a dû prendre note d'au moins 10 tâches d'accès à son information personnelle pendant une période d'une semaine, dans un *diary*. Au terme de cette première phase, les tâches ont été rassemblées et catégorisées selon une taxonomie à 3 dimensions. La première dimension, proposée par [7], distingue les différentes **cibles** des tâches :

- *lookup* (L) : l'information recherchée est contenue dans un document (p. ex. « Recherche du titre exact d'un chant dans un fichier »-P7).
- *item* (I) : l'information recherchée est un document complet (p. ex. « Je cherche la présentation initiale pour le Projet XYZ »-P3).
- *multi-item* (M) : l'information recherchée consiste en plusieurs documents (p. ex. « Je cherche tous les documents concernant le projet ABC des années précédentes. »-P3).

Les tâches ont été catégorisées indépendamment par deux personnes avec une *intercoder reliability* (ICR) de 93% selon le *Stempel's Percentage Agreement Index*, soit davantage que ce qui est admis comme généralement acceptable (80%) pour les études en sciences sociales.

La deuxième dimension, intitulée **indices**, catégorise les tâches en fonction des indices qu'elles contiennent relativement à chacune des facettes décrites plus haut. Par exemple, une mention temporelle telle que

	Taux util. facettes (%)					Nbre de tâches
	FE	TS	T	SN	D	
<i>mail</i>	-	44	56	67	11	9
L	-	33	33	83	17	6
I	-	67	100	33	-	3
M	-	-	-	-	-	0
<i>doc</i>	61	61	54	04	07	28
L	67	33	67	-	-	6
I	61	72	44	06	11	18
M	50	50	75	-	-	4
<i>mail-doc</i>	-	67	67	67	33	3
L	-	-	-	-	-	0
I	-	100	100	-	-	1
M	-	50	50	100	50	2
<b>Total</b>	<b>43</b>	<b>58</b>	<b>55</b>	<b>23</b>	<b>10</b>	<b>40</b>

TABLE 2 : Taux d'utilisation des facettes par catégorie de tâches (selon ressource et cible). La colonne de droite indique le nombre de tâches réussies par catégorie.

« mai 2009 » dans l'énoncé de la tâche est un *indice relatif à la facette temporelle*. Notons que les subdivisions de cette dimension ne sont pas exclusives, une tâche pouvant évidemment contenir des indices relatifs à plusieurs facettes. Les indices possibles sont : *file explorer* (FE), *textual search* (TS), *type* (T), *social network* (SN), et *date* (D) qui regroupe les indices relatifs aux trois facettes temporelles. Les indices ont été encodés indépendamment par deux personnes sur un échantillon de tâches (ICR 80.1%).

Enfin, la troisième dimension distingue les tâches selon les **ressources** impliquées. Deux évaluateurs ont réalisé cette dernière catégorisation ensemble en discutant des éventuels conflits. Les subdivisions de cette dimension sont exclusives. Il s'agit de :

- *mail* : tâches liées aux courriels.
- *doc* : tâches liées aux fichiers enregistrés localement ou sur un disque réseau.
- *mail-doc* : tâches liées à des fichiers joints aux courriels ou recherches d'information nécessitant les ressources *mail* et *doc*.

À la suite de cette catégorisation, entre 4 et 6 tâches par participant ont été sélectionnées pour un total de 45 tâches. La deuxième étape de l'évaluation a eu lieu en moyenne quatre semaines après la distribution des *diaries*. Le prototype a été installé et configuré chez chaque participant. Après un tutoriel, chaque participant a dû réaliser les tâches sélectionnées pour lui, issues de son propre *diary*. Au cours de la réalisation des tâches, des *logs* incluant la séquence de facettes utilisées, le nombre d'éléments sélectionnés dans chaque facette ou encore le temps de complétion de la tâche ont été récupérés automatiquement. De plus, l'évaluateur présent a également pris des notes. Après l'évaluation, les participants ont complété un questionnaire. Les résultats provenant des *logs*, des

notes de terrain et des questionnaires sont présentés et discutés ci-dessous.

**Résultats.** 40 tâches sur 45 ont été réussies ; les 5 autres ont échoué pour des raisons externes au système lui-même (p. ex. fichier présent lors du *diary study* mais supprimé avant l'évaluation). Seules ces 40 tâches réussies sont prises en compte pour l'analyse des résultats. Au total, 147 opérations (sélections de facettes, annulation ou restauration de sélection) ont été réalisées donc en moyenne 3.675 opérations par tâche (min : 1, max : 16) ; 4 tâches ont nécessité plus de 7 sélections et 22 annulations de sélections ont été effectuées.

Le tableau 2 présente le taux d'utilisation des facettes en fonction des cibles et ressources des tâches. Les facettes *textual search* (TS) et *type* (T) sont les plus sollicitées puisque resp. 58% (23 tâches) et 55% (22 tâches) des tâches réussies les utilisent. L'explorateur de fichiers (FE) est utilisé pour 43% des tâches. La facette sociale (SN) est utilisée pour environ un quart des tâches, tandis que les facettes temporelles (D) sont utilisées pour une tâche sur dix.

La facette *file explorer* (FE) est utilisée exclusivement dans le cas de la recherche de documents (fichiers). Son taux d'utilisation dans ce contexte est de 61% (17 tâches). L'habitude et l'aisance des utilisateurs à naviguer dans la hiérarchie de fichiers en est sans doute la raison principale. Les facettes *textual search* (TS) et *type* (T) ont quant à elles été utilisées de manière plutôt homogène quelles que soient les ressources impliquées pour environ une tâche sur deux. L'utilisation de la facette *social network* (SN) est prépondérante dans le cas de courriels. En particulier, les tâches de *lookup* dans des courriels ont mené à une utilisation de la facette sociale dans 83% des cas (5 tâches sur 6). Elle y supplante toutes les autres facettes. Dans les tâches *mail* également, l'utilisation supérieure à la moyenne de la facette *type* vient du fait qu'elle permet de séparer efficacement les courriels des autres documents. Notons enfin que si le nombre modeste de tâches *mail-doc* (impliquant aussi les courriels) ne permet pas de formuler des observations fiables, une importance notable de la facette sociale semble quand même se distinguer (67%, 2 tâches sur 3). L'utilisation réduite de la facette temporelle corrobore les déclarations des participants à notre premier questionnaire en ligne qui montraient une retenue dans l'utilisation de critères temporels pour la recherche. Néanmoins, les notes prises par l'évaluateur durant la réalisation des tâches révèlent que les facettes *date modified* et *date created* se sont montrées efficaces lorsque l'utilisateur avait le souvenir d'un indice temporel. De plus, les *logs* indiquent que les sélections dans les facettes temporelles n'ont jamais été annulées.

	<i>FE</i>	<i>TS</i>	<i>T</i>	<i>SN</i>	<i>D</i>
Util. facette avec indice (%)	24	30	<b>68</b>	<b>67</b>	50
Util. facette sans indice (%)	<b>76</b>	<b>70</b>	32	33	50
Nombre d'util. facette	17	23	22	9	4

TABLE 3 : Taux d'utilisation des facettes en fonction de la présence ou non d'indices correspondants dans la description de la tâche. La ligne finale indique le nombre de fois où la facette a été utilisée.

Le tableau 3 compare la présence d'indices relatifs aux facettes dans la description des tâches issue du *diary* et l'utilisation effective de la facette concernée lors de la réalisation de la tâche durant l'évaluation. Les facettes *type* (T) et *social network* (SN) sont majoritairement utilisées lorsqu'un indice est présent dans la tâche demandée. A l'inverse, les facettes *file explorer* et *textual search* sont utilisées plus souvent en l'absence d'indice. Notre hypothèse est que les utilisateurs sont habitués à utiliser quotidiennement ces dernières, qui sont présents nativement sur les systèmes d'exploitation. En revanche, leur inexpérience dans l'utilisation des facettes sociales et de type fait qu'ils songent moins à les utiliser sans un indice explicite dans la description de la tâche.

Par ailleurs, les résultats du questionnaire indiquent les facettes préférées des utilisateurs après qu'ils ont utilisé le système. Il y était demandé de classer les facettes par ordre de préférence (de 1 à 7, 1 étant la préférée). La figure 2 montre la distribution de ce classement. Les facettes les plus habituelles sont également celles que les participants préfèrent. La facette sociale est généralement appréciée, même si elle fait l'objet de notes irrégulières (écart interquartile = 5), les différents participants ayant eu un succès inégal avec celle-ci. Les facettes temporelles sont les moins appréciées. Il est toutefois intéressant de comparer ces résultats avec l'utilisation réelle des facettes sur la base des *logs* (fig. 3) : la comparaison révèle en effet que l'utilisation effective des facettes est étroitement liée à la préférence perçue.

Enfin, dans le questionnaire, les participants jugent l'approche de *Weena* « utile », avec une note moyenne de 4.05 sur 5 (min : 3, max : 5). Leurs commentaires indiquent qu'ils trouvent le système utile pour la recherche de documents en général et la recherche dans les courriels en particulier. Deux choses ont principalement plu : (1) l'unification (courriels, documents et pièces jointes sont visibles dans la même interface) et (2) la possibilité de combiner aisément plusieurs critères de recherche (facettes).

**Discussion.** L'évaluation tend à confirmer l'intérêt de l'unification de l'information personnelle et de la navigation par facettes. Si les participants utilisent et préfèrent les facettes les plus habituelles (hiérar-

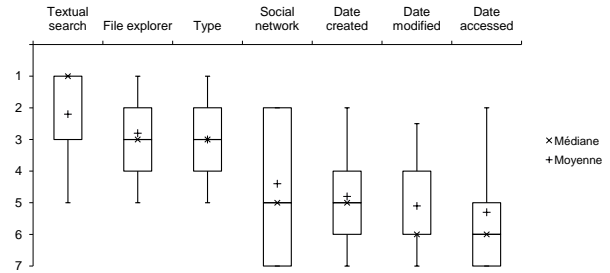


FIGURE 2 : Classement subjectif des facettes par ordre de préférence, issues des réponses au questionnaire.

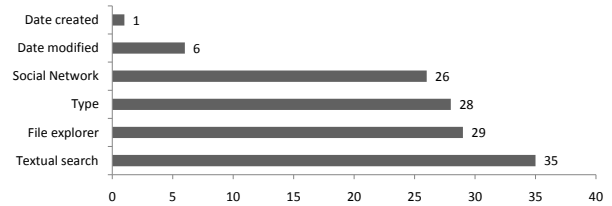


FIGURE 3 : Nombre de sélections effectuées dans chaque facette, issues des *logs* d'utilisation. La facette *Date accessed* manque car elle n'a jamais été sélectionnée.

chie de fichiers, recherche textuelle, type), l'utilisation et l'intérêt pour la facette sociale est remarquable en particulier dans les tâches utilisant des ressources de communication (courriel, pièces jointes). Ainsi, comme il est naturel d'associer des personnes aux documents qui les concernent, il est naturel d'accéder à ces documents en passant par ces personnes. Remarquons en outre que l'association automatique entre personnes et documents n'était pas parfaite dans notre prototype, ce qui a lésé certains participants et amené une plus grande variété d'appréciation de la facette sociale. Les résultats de l'évaluation encouragent à améliorer encore l'association entre personnes et documents. Les facettes temporelles ont été moins utilisées que les autres, mais lorsqu'elles l'ont été, elles se sont montrées efficaces. Il faut noter également qu'aucune tâche liée au calendrier, a priori plus dépendant du temps, n'a pu être proposée, puisqu'aucune n'était mentionnée dans les *diaries*. L'ajout de repères temporels personnels (p. ex. rendez-vous issus du calendrier) dans les facettes temporelles rendraient sans doute celles-ci plus faciles à utiliser. En outre, dans une tâche annexe d'auto-analyse de leurs pratiques de GIP, il était demandé aux participants de décrire sur quels projets ils avaient travaillé six mois plus tôt. Tous ont alors mis à profit la facette temporelle, et apprécié l'aisance avec laquelle ils pouvaient bénéficier d'une vue d'ensemble de la période concernée.

## CONCLUSION

Cet article propose d'aborder la gestion de l'information personnelle par l'utilisation de facettes de navigation. L'implémentation de cette approche et l'évalua-

tion qui s'ensuit montrent sa faisabilité et un intérêt des utilisateurs pour les facettes. En particulier l'utilisation du réseau social extrait de leurs courriels, ainsi que l'axe temporel qui peut être relié à leurs activités, trouvent une place non négligeable pour soutenir des tâches de recherche d'information particulières : par exemple pour retrouver des documents reçus par des personnes, resp. pour analyser leurs activités passées. Une intégration en amont de ces facettes, au moment de l'archivage de l'information, ainsi que la possibilité pour l'utilisateur de les personnaliser en entrant des méta-données, devrait faciliter encore davantage leur adoption. Dans les travaux à venir, nous désirons nous concentrer sur l'amélioration de la facette sociale, aussi bien au niveau de la représentation que de l'alignement automatique avec les autres facettes, afin de raviver le lien implicite qui existe entre le réseau social d'un individu et son information personnelle au sens large.

## BIBLIOGRAPHIE

1. Bush, V. As We May Think. *The Atlantic Monthly*, 176(1) :101–108, 1945.
2. Capra, R., and Pérez-Quinones, M. Using web search engines to find and refine information. *Computer*, 38(10) :36–42, 2005.
3. Chau, D. H., Myers, B., and Faulring, A. Feldspar : A system for finding information by association. In *PIM 2008 - Proceedings of Personal Information Management, Workshop at CHI 2008*, Florence, Italy, 5-6 April 2008 2008.
4. Cutrell, E., Robbins, D., Dumais, S., and Sarin, R. Fast, flexible filtering with Phlat. In *CHI '06 : Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 261–270, New York, NY, USA, 2006. ACM.
5. Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., and Robbins, D. C. Stuff I've seen : a system for personal information retrieval and re-use. In *SIGIR '03 : Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 72–79, New York, NY, USA, 2003. ACM Press.
6. Elswailer, D. *Supporting Human Memory in Personal Information Management*. PhD thesis, Department of Computer and Information Sciences, University of Strathclyde, 2007.
7. Elswailer, D., and Ruthven, I. Towards task-based personal information management evaluations. In *SIGIR '07 : Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 23–30, New York, NY, USA, 2007. ACM.
8. Elswailer, D., Ruthven, I., and Jones, C. Towards memory supporting personal information management tools. *J. Am. Soc. Inf. Sci. Technol.*, 58(7) :924–946, 2007.
9. Evéquo, F., and Lalanne, D. "I thought you would show me how to do it" - a study on self-analysis of PIM practices. In *ASIS&T PIM Workshop, Vancouver, BC, Canada.*, 2009.
10. Fertig, S., Freeman, E., and Gelernter, D. Lifestreams : an alternative to the desktop metaphor. In *CHI '96 : Conference companion on Human factors in computing systems*, pages 410–411, New York, NY, USA, 1996. ACM.
11. Jones, W. *Keeping Found Things Found : The Study and Practice of Personal Information Management*. Morgan Kaufmann, November 2007.
12. Jones, W., Phuwanartnurak, A. J., Gill, R., and Bruce, H. Don't take my folders away! : organizing personal information to get things done. In *CHI '05 : CHI '05 extended abstracts on Human factors in computing systems*, pages 1505–1508, New York, NY, USA, 2005. ACM.
13. Jones, W., and Teevan, J., editors. *Personal Information Management*. University of Washington Press, Seattle, 2007.
14. Lamming, M., and Flynn, M. Forget-me-not : intimate computing in support of human memory. In *Proceedings FRIEND21 Symposium on Next Generation Human Interfaces*, 1994.
15. Ranganathan, S. R. *Colon classification*. The Madras Library Association, 1933.
16. Rekimoto, J. Time-machine computing : a time-centric approach for the information environment. In *UIST '99 : Proceedings of the 12th annual ACM symposium on User interface software and technology*, pages 45–54, New York, NY, USA, 1999. ACM.
17. Ringel, M., Cutrell, E., Dumais, S. T., and Horvitz, E. Milestones in Time : The value of landmarks in retrieving information from personal stores. In Rauterberg, M., Menozzi, M., and Wesson, J., editors, *INTERACT*, pages 184 – 191, Zurich (Switzerland), 1 -5 September 2003. IOS Press.
18. Smith, G., Czerwinski, M., Meyers, B., Robbins, D., Robertson, G., and Tan, D. S. FacetMap : A scalable search and browse visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5) :797–804, 2006.
19. Thomet, J. Weena, un système de navigation par facettes de documents personnels. Master's thesis, University of Fribourg, 2010.
20. Whittaker, S., Jones, Q., Nardi, B., Creech, M., Terveen, L., Isaacs, E., and Hainsworth, J. Contactmap : Organizing communication in a social desktop. *ACM Trans. Comput.-Hum. Interact.*, 11(4) :445–471, 2004.
21. Yee, K.-P., Swearingen, K., Li, K., and Hearst, M. Faceted metadata for image search and browsing. In *CHI '03 : Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 401–408, New York, NY, USA, 2003. ACM.