HAL open science

Vers un système de taguage iconique basé sur Hypertopic

Xiaoyue Ma, Jean-Pierre Cahier

To cite this version:

Xiaoyue Ma, Jean-Pierre Cahier. Vers un système de taguage iconique basé sur Hypertopic. IC 2012 - 23e journées francophones d'ingénierie des connaissances, Jun 2012, Paris, France. pp.150. hal-00726846

HAL Id: hal-00726846

https://hal.science/hal-00726846v1

Submitted on 31 Aug 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Actes des 23e journées francophones d'ingénierie des connaissances (IC), 12p. , 26-29 Juin 2012, Paris, France. . Paris: Hermes, 2012, 16 p.

Vers un système de taguage iconique basé sur Hypertopic

Xiaoyue Ma[1], Jean-Pierre Cahier[1],

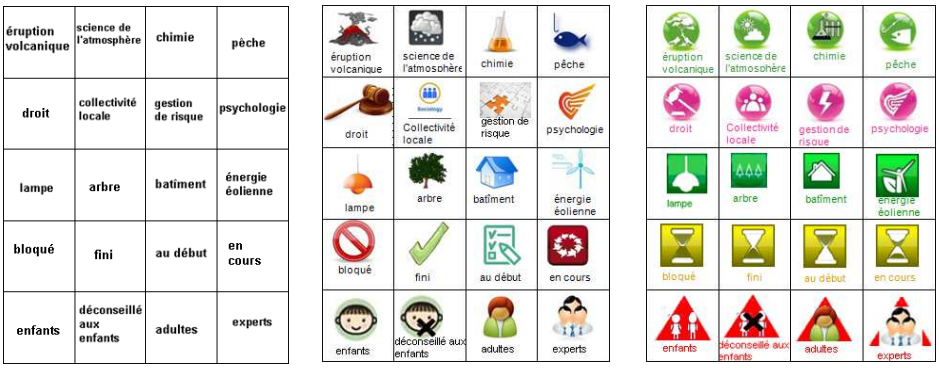
[1] Laboratoire ICD/Tech-CICO
UTT (Université de Technologie de Troyes)
10010 Troyes,
{xiaoyue.ma,cahier}@utt.fr

Résumé : Les systèmes de « taguage recommandé », où les tags sont fixés à l'avance de façon descendante (« top-down ») par des experts, proposent des tags utilisables pour l'organisation des ressources personnelles ou pour la classification, le partage et la recherche de connaissances. En cas de thèmes très nombreux croisant des sémantiques variées, les relations entre les tags et les documents indexés deviennent plus confuses, ce qui rend l'utilisation et la réutilisation des systèmes de taguage recommandé plus difficiles. Dans cet article, nous présentons une solution de Langage Visuel Distinctif pour construire un système de taguage iconique qui répondra à ce problème, en améliorant la représentation des tags ainsi que leur structure. Nous présentons l'évaluation de cette approche par une expérience comparative de taguage recommandé dont les résultats valident notre hypothèse.

Mots-clés : système de taguage, représentation des connaissances, visualisation, icône, Web2.0, Web socio-sémantique.

1. Introduction

Dans les systèmes de taguage recommandé, des documentalistes ou des experts prédéfinissent des tags utilisables pour la classification, le partage et la recherche de connaissances dans des ressources personnelles ou organisationnelles. Dans les utilisations collectives, le nombre et le volume croissants des vocabulaires de métiers, se référant à des sémantiques variées, peuvent impliquer que les relations entre les tags et les documents indexés deviennent trop ambigües ou confuses, qui rend l'utilisation de systèmes de taguage plus difficile lorsque ceux-ci sont purement textuels (Furnas *et al.*, 1987) (cf. (a) de la Figure.1).



(a) (b) (c)

FIGURE 1 – *Trois systèmes de taguage que nous avons comparés : (a) système de taguage textuel, (b) système de taguage iconique sans structure visuelle, (c) système de taguage iconique bien structuré*

Face aux limites des tags textuels dans un système collectif de taguage, l'amélioration de la représentation des tags est nécessaire. Des études en psychologie cognitive de l'image (Denis, 1989) et l'acceptation progressive de la Théorie du double codage (Paivio, 1986) ont imposé peu à peu l'idée que dans la représentation des informations, le code visuel et le code verbal sont utilisables de façon à la fois autonome et complémentaire. Ces recherches ont montré que la mémorisation de certaines informations textuelles est renforcée si un signe pertinent visuel est également présenté ou si l'apprenant peut imaginer une image visuelle reliée avec l'information verbale. En plus, des travaux antérieurs ont théoriquement prouvé que les icônes peuvent être une représentation visuelle efficace des connaissances en considérant les caractères graphiques et les caractères symboliques (Lohse *et al.*, 1990).

Comme l'illustre la Figure 1, un système de taguage s'intéresse non seulement à la représentation de chaque tag, mais aussi à la *structure* des tags. Malheureusement les travaux en sémiotique graphique et en psychologie cognitive, tels que la Théorie du double codage ont expliqué surtout l'amélioration de la compréhensibilité de chaque icone, en tant qu'image porteuse de connaissance, mais ne traitent pas de la structure. Par ailleurs les catégories pertinentes pour les connaissances concernées augmentent avec le nombre de points de vue existant chez les acteurs impliqués. Si ces catégories de connaissances sont représentées par des icônes non-structurées, les utilisateurs pourront ressentir une désorientation accrue face à trop de symboles isolés (cf. (b) de la Figure.1). Pour cela, la représentation de la structure de tags dans le système de taguage iconique a besoin d'être améliorée (cf. (c) de la Figure.1). La sémiotique graphique de (Bertin, 1967) a proposé six variables de base: la taille, la couleur, la forme, la texture, l'orientation et la valeur. Toutes ces variables ne possèdent pas la même aptitude à exprimer les mêmes informations. La couleur ou la forme,

par exemple, sont incapables de traduire des rapports quantitatifs alors qu'elles expriment parfaitement les différences. Définir le type d'information que nous désirons communiquer nous permet de choisir la variable visuelle la plus adéquate. Les variables graphiques sélectionnées pourraient être également prises en compte pour la conception de systèmes iconiques pour cataloguer des connaissances et des informations.

De précédentes recherches empiriques ont mis en évidence la notion de « système d'icône » comme celui des signes de la route, les symboles de la sécurité incendie (Collins & Lerner, 1982) et les applications de svstème médical qui intéressent l'Ingénierie de Connaissances (Lamy, 2009, cf. Figure. 2). Les chercheurs qui ont travaillé sur ce système d'icônes médicales ont commencé à illustrer explicitement des corpus complexes avec une représentation graphique, où les icônes combinent plusieurs parties graphiques. Mais dans ces études, peu d'attention a été portée à expliciter la structure de connaissances du système de taguage par icônes. Ces approches demanderaient à être généralisées de façon plus systématique.

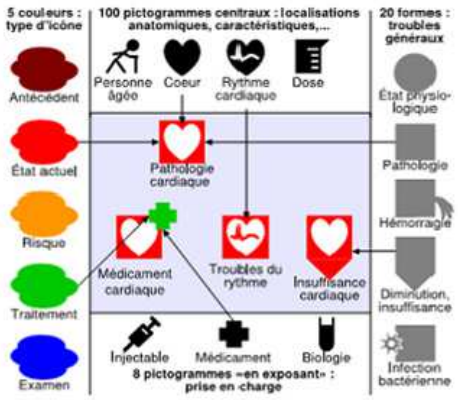


FIGURE 2 – *Application médicale d'un système d'icônes (Lamy, 2009)*

Dans cet article, nous proposons une approche de système de taguage iconique basé sur un modèle d'Ingénierie de Connaissances : Hypertopic[1] (Zhou *et al.*, 2006). Nous faisons l'hypothèse que les tags iconiques structurés avec les régularités graphiques d'un Langage Visuel Distinctif basé sur ce modèle améliorent l'efficacité des systèmes de taguage, en particulier pour les applications orientées vers l'annotation participative, le partage des connaissances et d'autres types d'activités concernant les tags. Cette approche conduit à un lien possible entre la conception graphique d'icônes, l'ingénierie des connaissances (IC) et le travail coopératif assisté par ordinateur (TCAO).

[1] Hypertopic est un modèle multi-point de vue pour le Web socio-sémantique. (cf. http://hypertopic.org)

2. Système de taguage iconique basé sur Hypertopic

Dans les systèmes de taguage iconique tels que nous les définissons, on doit s'intéresser à la visualisation à deux niveaux : celui de la structure de tags d'une part, et celui de la représentation de chaque tag d'autre part. Aussi notre approche propose-t-elle d'abord une manière de structurer les tags, puis d'iconiser ces tags ainsi que leur structure. Ce travail ouvre un nouveau champ de recherche dans le domaine de l'Ingénierie des Connaissances, car une fois la structure de tags confirmée, l'ensemble de tags pourra être considéré comme une catégorisation des connaissances aidant à organiser plus efficacement les informations. Lors de la qualification des items par des tags *via* le système de taguage, notre hypothèse est que l'usager bénéficie ainsi de catégories mieux identifiables.

Par ailleurs, le but à moyen terme de nos travaux sur le taguage iconique est le taguage iconique participatif : en raison de cette dimension coopérative de nos recherches, nous avons choisi d'utiliser un modèle de connaissances respectant les principes du Web Socio-Sémantique (Zacklad *et al.*, 2003) : le modèle « Hypertopic » (Zhou *et al.*, 2006). Il propose de classer des items par des thèmes (topics), des attributs, et des ressources (cf. Figure.3), avec la particularité que tous les thèmes appartiennent à des points de vue différents. Les tags seront pour nous des thèmes correspondant à des buts d'information et d'action différents des acteurs, donc à des points de vue différents.

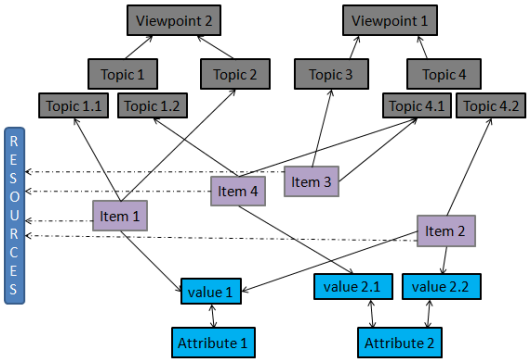


FIGURE 3 – *Système de taguage textuel proposé par Hypertopic*

Hypertopic propose une catégorisation mettant l'accent sur le point de vue qui est importante en classification collaborative des connaissances. Les thèmes sont catalogués sous une structure d'arbre en considérant le point de vue comme le nœud «parent» d'une catégorie. Les systèmes de connaissances qui sont construits selon Hypertopic permettent aux participants de partager et de rechercher des connaissances par les thèmes liés aux points de vue qu'ils préfèrent, mais aussi de découvrir s'ils veulent les points de vue et le langage des autres. Les attributs et leurs valeurs fournissent également des caractéristiques qui enrichissent les items avec

des informations supplémentaires. Ils sont organisés en paire attribut-valeur comme une facette (Mas *et al.*, 2008)..

Hypertopic fournit classiquement une structure significative pour gérer les tags textuels découlant de thèmes et d'attributs (cf. Figure.4, partie gauche). Mais lorsque les tags textuels ne portent pas les signes de la famille à laquelle ils appartiennent, ils seront moins distinctifs en particulier aux yeux des utilisateurs connaissant peu le domaine. Ce problème sera encore plus prononcé si le nombre de tags augmente.

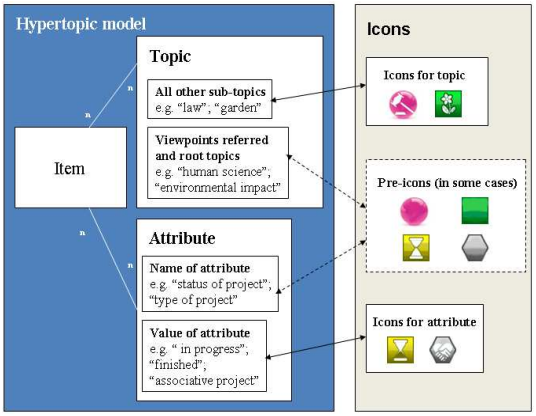
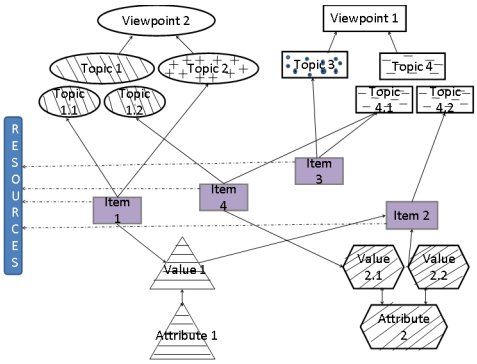


FIGURE 4 – *Modèle pour construire un système de taguage iconique basé sur Hypertopic*

Notre idée est de profiter de la structure de tags catégorisés par Hypertopic et ensuite d'iconiser les tags pour *mieux visualiser les tags et leur structure* dans un système de taguage. Les caractères symboliques interpréteront la connaissance représentée par les tags tandis que les caractères graphiques aideront à visualiser la structure de tags dans le système de taguage. Parmi les trois éléments principaux de Hypertopic, les thèmes (au sein de points de vue) et les valeurs d'attributs constituent les éléments, au départ textuels, que nous voulons iconiser (cf. Figure.4, partie droite). En particulier, nous proposons d'utiliser un groupe d'icônes spéciales appelées « pré-icônes » qui sont conçues pour représenter le point de vue, un thème de niveau supérieur, ou un nom d'attribut. Elles sont considérées comme les bases graphiques pour signifier la structure des tags iconiques. D'autres tags iconiques de chaque catégorie (les thèmes de niveau fin ou les valeurs d'attributs) seront précisés par la combinaison des signes concrets sur les pré-icônes correspondantes. Toutefois, la pré-icône de l'attribut n'est pas toujours nécessaire. Par exemple, lorsque nous représentons les valeurs de l'attribut « langue », il est assez clair de les représenter de manière indépendante avec des codes empruntés aux drapeaux nationaux, sans qu'il y ait de pré-icône ni de signe graphique communs.

Notre approche peut être considérée plus précisément comme un Langage Visuel Distinctif (LVD) (cf. Figure.5) qui vise à représenter

visuellement des objets classés selon Hypertopic. Un LVD s'intéresse plus à la structure de tags qu'au symbole d'icône. Parmi les six variables graphiques proposées par Bertin (Bertin, 1967), trois sont peu utiles pour cet objectif, aussi avons-nous choisi de les éliminer en première approche : la taille, l'orientation et la valeur. Pour des raisons esthétiques, les icônes sont censées être conçues dans une taille unifiée. Surtout quand on veut les afficher sur le web, il n'est pas pratique de concevoir la taille appropriée pour chaque tag iconique et pour la nouvelle catégorie de tags. L'orientation peut causer le même problème. Un objet 3D qui tourne à 45 ° fait peu de différence avec le même objet qui tournerait à 60 ° sur un écran Web. Enfin des icônes qui utiliseraient des valeurs variées se heurteraient aux paramétrages et aux conditions différentes d'un écran d'ordinateur à un autre.



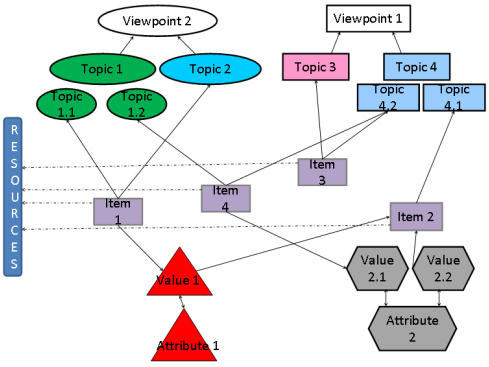


FIGURE 5 – *Langage Visuel Distinctif - Régularité graphique de la création d'un système de taguage iconique basé sur Hypertopic (à gauche : version colorée, à droite, version monochromatique équivalente)*

En revanche, la forme, la couleur et la texture sont très utiles, c'est pourquoi en première approche nous les avons privilégiées pour la création de LVD satisfaisant la structure du système de taguage iconique (cf. Figure.5 - gauche). Dans cet exemple, les pré-icônes pour les points de vue sont essentiellement distinguées les unes des autres par la forme. Des thèmes au premier niveau, sont également distingués par des pré-icônes avec une signification supplémentaire apportée par la couleur. Puisque les tags iconiques de thèmes sont encore catalogués dans des catégories différentes, il est logique d'ajouter la nouvelle variable graphique pour les thèmes dans le deuxième niveau. Cependant, d'une part le nombre de variables graphiques est limité ; d'autre part les variables graphiques excessives réduisent la lisibilité de la catégorisation de tags. Visant à simplifier le système de taguage iconique, les thèmes à partir du deuxième niveau gardent toujours le même code visuel (forme, couleur) sans être distingués par des nouvelles variables graphiques.

La règle graphique est pareillement appliquée aux attributs. Le nom d'attribut est iconisé par la pré-icône avec les formes colorées, puis on détaille la valeur en ajoutant le symbole à la pré-icône (sauf les cas spéciaux

tels que les langues, évoquées dans le paragraphe précédent). Dans le cas monochromatique, la couleur sera remplacée par la texture comme le deuxième variable graphique (cf. Figure.5 - droite) ne changeant pas les autres règles dans la version colorée.

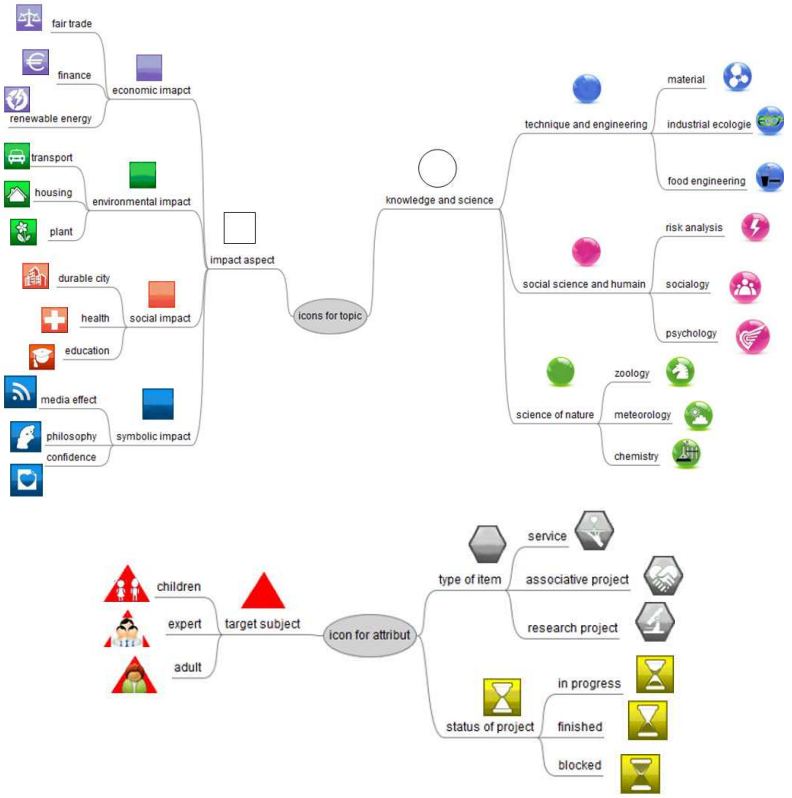


FIGURE 6 – *Une partie du système de taguage iconique basé sur Hypertopic et respectant un LVD, utilisé dans l'expérience dans la section 3*

Un LVD basé sur Hypertopic est donc guidé par les éléments d'Hypertopic – points de vue, thèmes, attributs, résultant d'une analyse d'IC - pour la construction du système de taguage iconique. Bien que la structure des tags iconiques et le code visuel (pré-icône) soient pour l'instant proposés pour un système de « taguage recommandé » (top-down, sans possibilité par exemple de taguage folksonomique), l'étape suivante à laquelle le LVD basé sur Hypertopic sera encore plus utile, sera celle du « taguage coopératif » où les acteurs seront invités à coopérer pour apporter de façon ascendante ou hybride leur contribution en termes de tags et de relations

La Figure.6 présente un extrait d'un système de taguage iconique respectant le LVD, où les tags, recommandés par des experts, portent sur le domaine du développement durable. Ce système de taguage a servi à l'expérience que nous allons maintenant décrire.

3. Expérimentation

Pour évaluer l'efficacité du système de taguage iconique basé sur Hypertopic nous avons réalisé une expérience de taguage « sur papier » dans notre Université. Pour concevoir cette expérience, nous considérons que *l'efficacité d'un système de taguage* est un sujet complexe associé à de très nombreux facteurs cognitifs chez l'utilisateur. Aussi nous nous sommes concentrés principalement sur l'amélioration concrète apportée par notre approche pour l'interprétation de tags et la visualisation de la structure de tags, par rapport aux autres deux types de systèmes de taguage : le système de taguage textuel pur et le système de taguage iconique sans structure visuelle. Le résultat de cette expérience nous amènera à une discussion plus approfondie sur l'argument psychologique.

3.1. Participants et matériel

Vingt-sept étudiants en dernière année à l'Université de Technologie de Troyes ont participé dans cette expérience, séparés au hasard en trois groupes, en tirant au sort un code de participant noté A, B, ou C: groupe A pour les tags textuels, groupe B pour les tags iconiques sans structure visuelle, groupe C pour les tags iconiques basés sur Hypertopic.

Le matériel expérimental se compose de trois « présentoirs » de système de taguage au format A3 et de 24 fiches d'items à taguer. Ils sont tous imprimés sur papier en couleur.

Les 87 tags ont été affichés au hasard mais gardent la même disposition dans tous les trois présentoirs des groupes A, B et C en marquant le numéro correspondant en haut à gauche (voir figure 3) : les tags textuels (groupe A), les tags iconiques sans structure visuelle (groupe B) et les tags iconiques basé sur Hypertopic (groupe C). Les tags textuels sont sélectionnés depuis les tags structurés venant de sept catégories de thèmes et trois catégories d'attributs de CartoDD[2], un système de connaissances dans le domaine du développement durable expérimenté antérieurement à l'UTT (Cahier et al., 2010). Dans CartoDD, basé sur le protocole Hypertopic, les items sont tagués par des thèmes depuis les points de vue et complétés par les attributs. Il a donc été facile de construire le LVD à partir de cette structure Hypertopic. Sept catégories de thèmes ont été choisies parmi les aspects d'impact (aspect économique, aspect environnemental, aspect social et aspect symbolique) et les aspects techno-scientifiques (technologie et l'ingénierie, sciences sociales, sciences naturelles). Ont été aussi retenus trois noms d'attributs, en forme de facette : le type d'item (projet, article, activité…), le statut du projet (en cours, terminé…) et le public cible (enfant, expert…).

---

[2] http://initiatives21.cartodd.org/

Dès que les 87 tags textuels ont été confirmés, nous avons cherché des icônes appropriées afin de composer les tags iconiques pour le groupe B. Ensuite il a fallu déterminer les pré-icônes du LVD pour le système de taguage du groupe C, et créer sur ces pré-icônes les tags iconiques basés sur Hypertopic, mais en prenant soin de s'appuyer sur mêmes symboles présents dans les tags iconiques du groupe B, comme montré dans la Figure.7 (si le tag « eau » du groupe B comporte un symbole de robinet, celui du groupe C utilise aussi ce symbole). Les deux types de tags iconiques sont *toujours sous-titrés*, avec des textes totalement identiques à ceux des tags textuels. La couleur des sous-titres des tags iconique pour le groupe C reste compatible avec les pré-icônes tandis que celle des sous-titres du groupe B reste en noir.

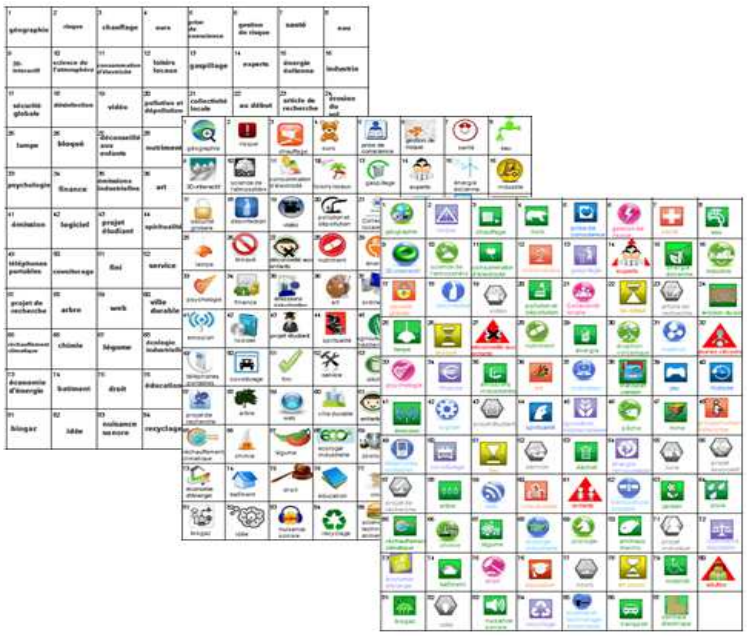


FIGURE 7 – *Présentoirs des trois systèmes de taguage (de g. à dr.): tags textuels, iconiques sans structure visuelle, iconiques basés sur Hypertopic*

Le matériel utilisé consiste aussi en 24 fiches (formatA4) des items à taguer. Tous ces items sont choisis depuis CartoDD, avec des sujets variés du développement durable, tel que l'eau, l'énergie ou le climat. Le type des items est aussi très diversifié ; certains d'entre eux présentent des résultats ou des recherches en cours, tandis que d'autres décrivent une idée à ses débuts. Le principal critère de choix de ces items est de couvrir la plupart des 87 thèmes et valeurs d'attributs des présentoirs. Chaque item est présenté avec une photo, une dizaine de lignes de description et 15 cases vides pour mettre les tags choisis. Les 24 items sont proposés aux participants des trois groupes dans le même ordre fixe, selon un numéro indiqué sur chaque fiche.

3.2. Protocole

L'ensemble des 27 participants étaient présents ensemble dans la salle pour tirer au sort le code de participant Ax, Bx ou Cx (9 codes pour chaque groupe, e.g. A3 = le 3ème participant du groupe A), à rappeler ensuite sur tous les matériels expérimentaux (par ailleurs anonymes). Puis ils ont été séparés en trois salles indépendantes, avec un animateur surveillant dans chaque salle. Puis, pour que les expérimentateurs puissent connaître les niveaux en français et en développement durable, chaque participant dut remplir un pré-questionnaire.

Une fois que tous les participants ont terminé le pré-questionnaire, ils se voient distribuer les 24 fiches d'items et le présentoir du système de taguage : groupe A avec les tags textuels, groupe B avec les tags iconiques sans structure visuelle et groupe C avec les tags iconiques basés sur Hypertopic. Après avoir compris la consigne, ils taguent les items l'un après l'autre en utilisant les tags du le présentoir et en notant le numéro correspondant dans les cases vides. Il suffit de mettre entre 1 et 15 tags, sans priorité. Les participants ne peuvent pas revenir aux items précédents, et doivent avoir fini de taguer l'item précédent avant de prendre la fiche suivante. Le processus de taguage est limité à 50 minutes, même s'ils n'ont pas terminé la tâche. Nous les avons encouragés à trouver des tags avec la meilleure qualité possible sans se soucier trop de quantité ni de vitesse.

L'expérience se poursuit par un post-questionnaire en deux parties. Dans la 1ére partie, conçue pour tous les participants, ils devaient répondre à trois questions en regardant le présentoir. La question n°1 concerne la compréhensibilité du présentoir utilisé en proposant des annotations et commentaires, en écrivant cinq tags les plus compréhensibles ainsi que cinq les moins compréhensibles. Dans la question n°2 il est demandé à tous de signaler une catégorie de tags et de la nommer. Nous avons ensuite comparé cette catégorie avec la catégorisation du LVD : sept catégories des thèmes et trois catégories d'attributs (marquées par les pré-icones accessibles au seul groupe C). Les participants étaient aussi encouragés à commenter l'utilisabilité du système de taguage et à donner des suggestions précieuses.

L'autre partie du post-questionnaire a consisté à tester la mémorisation la compréhension et la reconnaissance des tags iconiques comparée des deux types de tags iconiques (pour les seuls groupes B et C cf. Fig. 7), en enlevant cette fois le présentoir. Le test prenait la forme d'un quizz mettant en correspondance 12 tags iconiques et des textes.

3.3. Résultats

3.3.1. Pré-questionnaire

Le Test d'homocedasticité Levene n'a pas révélé une hétérogénéité significative entre les variances sur les scores aux tests de connaissances préalables ($F < 1$, $P = 0{,}902$). Les scores moyens sur le test de connaissances

sont 1,1 (SD = 1,05) pour le groupe A, 0,9 (SD = 1,05) pour le groupe B et 1,0 (SD = 1,00) pour le groupe C. Une analyse de variance réalisée sur les performances dans le pré-questionnaire n'a révélé aucune différence significative (F <1).

3.3.2. Processus de taguage

Nous proposons une méthode de la matrice d'experts pour évaluer la qualité de taguage. Cinq experts sur le développement durable ont été invités à taguer les 24 items avec 87 tags textuels donnés dans l'expérience sans limite de temps (en tenant compte tags dans trois types de présentoir représentent les mêmes objets, les tags textuels sont choisis pour vérifier leurs pertinences aux items). Il était demandé aux experts de donner à chaque tag une note de 0 à 5 pour représenter le degré de corrélation aux items. 5 correspond à la plus forte corrélation (0 à aucune relation)**.** La notation des cinq experts compose une matrice appelée la matrice d'experts montrant la relation entre les tags et les items (cf. tableau 1).

TABLEAU 1 – *Matrice de taguage d'expert*

	Item 1	Item 2	Item i	…	Item 24
Tag 1	0	0	1	…	0
Tag 2	3	4	2	…	1
Tag j	2	2	1	…	0
…	…	…	…	…	…
Tag 87	5	4	5	…	3

De même, le résultat de taguage de chaque participant pourra composer 27 matrices de taguage qui sont appelées la matrice de participant X. La différence unique est que les éléments de la matrice de participant ne seront que 1 ou 0: 1 se réfère aux tags utilisés alors que 0 aux tags non indiqués dans les boîtes.

Pour analyser la qualité de taguage de participant X, la formule ci-dessous est appliquée.

$$Rx = \sum_{i=1}^{24}\sum_{j=1}^{87} TE_{ij} \cdot TPx_{ij} \quad \Sigma TE_{ij} \bullet TPx_{ij} \qquad (1)$$

TE_{ij}: chiffre dans le rang i la colonne j de la matrice d'experts de taguage
TPx_{ij}: chiffre dans le rang i la colonne j de la matrice de participant de taguage (participant x)

Par conséquent, Rx est considérée comme le critère de comparaison sur la qualité du taguage entre les trois groupes. En plus de la qualité de marquage, la vitesse de taguage est considérée (dans l'expérience, personne ne remit les items tagués à l'avance). En effet la vitesse de taguage est affectée par Rx (1), par exemple, la matrice de participant A1 laissant vides

les quatre dernières colonnes signifie qu'il arrête à l'item n ° 20 et sa vitesse implicite sera Rx.

Après avoir vérifié que l'homoscédasticité des variances a été respectée (F <1), Rx (1) sont analysés par les analyses répétées des variances (ANOVA). Le groupe C utilisant le système de taguage iconique basé sur Hypertopic (M = 246.3, SD = 49,1) manifeste nettement mieux que le groupe A avec les tags textuels (M = 198,7, SD = 39,4), MD = 47,6, P = 0,037. De même, le groupe C (M = 246.3, SD = 49,1) ont montré un effet significatif dans Rx par rapport le groupe B utilisant le système de taguage iconique sans structure visuelle (M = 199,2, SD = 40,2), MD = 47.1, p = 0,041. Par contre, il n'y avait pas de différence significative entre Rx dans le groupe A (M = 198,7, SD = 39,4) et le groupe B (M = 199,2, SD = 40,2), MD = 0,5, P = 0,977.

TABLEAU 2 – *Moyenne des Rx (1) des trois groupes pour évaluer la performance de taguage*

	Moyenne	**SD**
Groupe A	198.7	39.4
Groupe B	199.2	40.2
Groupe C	246.3	49.1

	Différence de moyenne (MD)	**P**
Groupe A vs Groupe B	0.5	0.977(>0.05)
Groupe B vs Groupe C	47.1	0.041(<0.05)
Groupe A vs Groupe C	47.6	0.037(<0.05)

3.3.3. Post-questionnaire

La prédiction critique de l'identification de la structure des tags (la première partie du post-questionnaire) concerne la comparaison avec les catégories du LVD (sept catégories des thèmes et trois catégories des attributs). Les participants qui sont en correspondance complète par rapport au profil de taguage proposé par les experts obtiendront 2 points. Ceux dont la catégorie correspond partiellement à l'une des catégories proposées sont crédités de 1 point. Enfin aucun point n'est obtenu pour les participants en cas d'adéquation avec aucune catégorie proposée. En respectant l'homocedasticité des variances (F <1), la procédure d'analyse a été mise en jeu. Les résultats du groupe C manifestent une grande cohérence à la catégorisation de CartoDD (M = 1,4, SD = 0,7), qui a été sans surprise mieux réalisées que le groupe A (M = 0,7, SD = 0,7), MD = 0,7, P = 0,035 et le groupe C (M = 0,5, SD = 0,8), MD = 0,9, P = 0,019. Au contraire, aucun écart important a été fait entre A (M = 0,7, SD = 0,7) et B (M = 0,5, SD = 0,8), MD = 0,2, P = 0,645.

TABLEAU3 – *Performances moyennes de l'identification de la structure des tags des trois groupes*

	Moyen	**SD**
Groupe A	0.7	0.7

Groupe B	0.5	0.8
Groupe C	1.4	0.7

	Différence de moyenne (MD)	**P**
Groupe A vs Groupe B	0.2	0.645(>0.05)
Groupe B vs Groupe C	0.9	0.019(<0.05)
Groupe A vs Groupe C	0.7	0.035(<0.05)

L'évaluation de la mémorisation des tags iconiques est mise en œuvre par des points : une icône correspondant permet de gagner 1 point et le score total est 12. Les résultats du groupe B et C respectent l'homocedasticité des variances (F <1). Une analyse plus approfondie d'ANOVA montre une performance équivalente entre les deux groupes: le groupe B (M = 13,1, SD = 0,8), le groupe C (M = 13,6, *S*D = 0,7), MD = 0,5, P = 0,273.

TABLEAU 4 – *Performances moyennes de la mémorisation des tags iconiques des trois groupes*

	Moyenne	**SD**
Groupe B	13.1	0.8
Groupe C	13.6	0.7

	Différence de moyenne(MD)	**P**
Groupe B vs Groupe C	0.5	0.273 (>0.05)

3.4. Discussion

Compte-tenu de la petite taille de l'échantillon de participants - à laquelle nous souhaitons remédier par une expérience ultérieure, cette fois informatisée permettant notamment d'augmenter ce nombre - les résultats actuels sont partiellement en accord avec nos prédictions. Comme prévu, le système de taguage iconique basé sur Hypertopic, dans le cas d'icones sous-titrées, apparait améliorer, d'une façon relativement significative, l'efficacité de taguage par rapport aux deux autres types de systèmes testés.

Cela peut s'expliquer par la Théorie du double codage (Paivio, 1986), parce que l'information verbale explicitement exprimée est associée à de l'information visuelle. Le symbole figurant au premier plan des tags iconiques agit comme une illustration de texte et il aide à améliorer la compréhensibilité et la mémorisation des tags textuels. Leur sous-titre étant là en cas de besoin pour éviter les équivoques, les icônes fournissent une traduction visuelle dans les deux types de taguages iconiques B et C utile pour les connaissances complexes, Cet avantage permet de mieux identifier le sens du tag. Le temps de réponse à l'image est beaucoup plus court que le temps de réponse au texte (Mayer & Gallini, 1990), ce qui conduit à une évocation plus rapide des tags iconiques que des tags textuels. En plus, grâce à la régularité graphique, l'utilisateur passe moins de temps à trouver et mémoriser le tag dans le cadre d'un LVD.

Bien que les deux systèmes de taguage iconique montrent une capacité équivalente à l'interprétation symbolique et à la mémorisation imagée, une explication particulière rend compte de la différence entre eux : *la représentation de structure*. L'identification de la structure de tags intervient comme étape dans le processus de taguage, de trois façons :

Premièrement, lorsqu'un utilisateur découvre le présentoir, sa première perception cognitive le pousse à tenter de comprendre la structure des tags. La plupart des participants, de n'importe quel groupe, ont commenté dans le post-questionnaire qu'ils obtenaient une mauvaise performance dans les cas où ils ressentaient une catégorisation peu claire des tags. Ce problème est explicitement confirmé par les résultats des groupes A et B. En particulier, les participants du groupe B disent avoir essayé de trouver une certaine structure existante au système de taguage iconique, alors que ces icônes « dépareillées » sont aussi diverses que les tags textuels pour trier les tags qui relèveraient de la même catégorie. En plus, disent-ils, le temps limité les força à terminer la tâche plutôt qu'à identifier la structure de tags. La meilleure performance sur l'identification (post-questionnaire) de la structure, obtenue par le groupe C a signifié que les codes graphiques du LVD, comme la couleur ou la forme, ont permis de visualiser la catégorisation. Certains des participants du groupe C ont décrit comment ils ont bénéficié de la structure visuelle pour le taguage : dès que le présentoir leur a été distribué, ils ont trouvé que les différentes couleurs de fond et les formes communes étaient un signe de la structure implicite des tags.

Deuxièmement, l'utilisateur redécouvre de temps en temps les tags ayant été utilisés pour les items précédents, associés à des thèmes ou des valeurs d'attribut proches, pour taguer l'item actuel. Il essaie de parcourir les tags dont les « membres de famille » ont été utilisés. Par exemple, si quelqu'un a tagué un item en utilisant un tag iconique vert, il tente de consulter d'autres tags verts pour taguer cet item ou pour taguer un autre item qui parle d'un sujet proche.

Troisièmement, certains des participants ont déclaré que, après avoir identifié la structure visuelle des tags iconiques, ils ont pris l'habitude de choisir régulièrement au moins un tag depuis chaque catégorie visuelle (tags verts carré, tags roses ronds ...), alors que dans les groupes A et B, les participants ignorent, conceptuellement, la catégorisation des tags. Aucune information explicite de la structure de tags ne passe dans ces deux types de système de taguage qui pourrait être capitalisée dans le processus de taguage. Dans le cas C, sélectionner un tag approprié parmi 87 autres devient un choix depuis chaque catégorie visuelle (7 pour les thèmes, 3 pour les attributs). Cette efficacité croîtra quand le nombre de tags augmente.

Cependant un résultat inattendu a été noté dans la comparaison entre les groupes A et B. Nous avions fait l'hypothèse que les tags iconiques sans structure amélioreraient le processus de taguage par rapport aux tags textuels. Mais cela n'a pas été significativement montré par le résultat. Plusieurs arguments pourraient être proposés pour l'expliquer. Tous les tags iconiques, rappelons-le, sont sous-titrés dans l'expérience, et le choix entre

l'information visuelle et l'information verbale va servir aux participants dans la compréhension ou la redécouverte de tags. Or de précédents travaux (Chandler & Sweller, 1991) (Jamet 2000) ont observé que la restauration de la coréférence entre texte et illustration ne produit aucune amélioration des performances de compréhension. Certains participants du groupe B ont commenté qu'ils ont fait davantage attention aux icônes lorsqu'ils éprouvaient des difficultés à comprendre des mots professionnels dans le sous-titre. Les tags iconiques sans structure auraient alors aussi bien fonctionné que les tags textuels. D'autre part, comme mentionné, Rx (1) est une intégration de la qualité et de la vitesse de taguage. Le problème pourra aussi se poser au groupe C, mais l'amélioration apportée à la structure rééquilibre ce problème de coréférence.

En considérant l'ensemble des résultats de cette première expérience, il apparaît que le système de taguage basé sur Hypertopic améliore l'efficacité du processus de taguage par rapport aux autres deux types. Cette amélioration profite essentiellement de la représentation visuelle de la structure des tags qui aide à la compréhension ainsi qu'à la mémorisation. Bien que l'information complémentaire des sous-titres ait prolongé la durée de compréhension et de relocalisation des tags, l'avantage apporté par le LVD est sensible. En revanche, les tags iconiques sans structure visuelle n'ont pas montré d'amélioration remarquable par rapport au texte seul.

4. Perspectives

Dans un proche avenir, nous souhaitons prolonger l'expérience décrite par une autre expérience, cette fois informatisée et réalisable sur le Web, en gardant pour l'essentiel le même protocole, enrichie par certaines leçons intéressantes de cette expérience « sur papier ». Cela nous apportera aussi des éléments plus précis sur le comportement de taguage utiles pour explorer ensuite les aspects collaboratifs (catégorisation iconique collaborative, partage d'informations…) sur le Web. Comparer les performances de taguage avec des dispositions différentes dans le présentoir (au hasard, par des catégories…) nous intéresse aussi pour la suite.

Par ailleurs, dans le cadre du projet MiipaDoc, des comparaisons sont effectuées avec l'approche « Hypertagging » (Zacklad *et al.,* 2011) développée dans ce projet (autour de l'outil Semiotag basé sur la classification à facettes et permettant d'utiliser des icones).

Le système étudié est pour le moment dans un mode « recommandé par des experts » (top-down), mais nous voulons également associer le cas « bottom-up » de taguage folksonomique et de partage du système d'icones toujours en utilisant un système de taguage iconique basé sur Hypertopic, et étudier la co-construction participative d'un tel système, en supposant qu'elle serait favorisée si elle se conforme à cette approche de LVD basé sur l'Ingénierie de Connaissances

Remerciements

Ces travaux ont été en partie financés par l'Agence Nationale de la Recherche (ANR) dans le cadre du projet Miipa-Doc n°2008 CORD 014 03

Références

BENEL A., ZHOU C., et CAHIER J.P. (2009). Beyond Web 2.0... And Beyond the Semantic Web. Chapter 1. *Design of Cooperative Systems*. Springer.

BERTIN J. (1967), *Sémiologie Graphique*, (2nd ed.), Paris, Gauthier-Villars

CAHIER J.-P., ZAHER L., ISOARD G. (2010).Document et Modèle pour l'Action, une méthode pour le Web socio sémantique. Document Numérique, 978-2-7462-3233-4, 2010, vol. 13, n° 2/2010, p. 75-96.

CHANDLER P. & SWELLER J. (1996). Cognitive load zhile learning to use a computer program. Applied Cognitive Psychology 10, 151-170.

COLLINS B. L. & LERNER N. (1982). Assessment of fire safety symbols. *Human Factors* 24(1), 75-84.

DENIS M. (1989). Image et Cognition. Collection Psychologie d'aujourd'hui. ISBN 2 13 042250 0.

FURNAS G.W., FAKE C., VON AHN L., SCHACHTER J., GOLDER S., FOX K., DAVIS M., MARLOW C., ET NAAMAN, M. (2006). Why do tagging systems works? Extended Abstracts of the SIGCHI Conference (CHI' 06).

JAMET E. (2000).L'intégration spatiale d'éléments textuels et illustratifs améliore-t-elle la performance? Revue d'intelligence artificielle,14,167-188.

LAMY J., DUCLOS C., ET VENOT A. (2009) De l'analyse de texte à la conception d'une interface graphique d'accès aux connaissances sur le médicament, IC2009, Hammamet, Tunisia.

LOHSE J., RUETER H., BIOLSI K., ET WALKER N. (1990). Classifying visual knowledge representations: a foundation for visualization research, IEEE. Siggraph: ACM Special Interest Group on Computer Graphics and Interactive Techniques. ISBN: 0-8186-2083-8, pp. 131 – 138.

MAS S., BENEL A., CAHIER J.-P., ET ZACKLAD M. (2008). Classification à facettes et modèles à base de points de vue. CAIS 2008, Vancouver, 10 p.

MAYER R.E. & GALLINI J. (1990). When is an illustration worth ten thousand words? Journal of Educational Psychology 82, 715-726.

PAIVIO A. (1986). Mental representations: a dual coding approach. Oxford. England: Oxford University Press.

ZACKLAD M., CAHIER J.P. ET PÉTARD X (2003). Du web cognitivement sémantique au web socio sémantique : Exigences représentationnelles de la coopération. Web sémantique et Sciences humaines et sociales, 2003.

ZACKLAD M.,DESFRICHES O.,BERTIN G.,MAHE S.,RICARD B.,MUSNIK N., CAHIER J.-P.,BENEL A.,LEWKOWICZ E. (2011). Miipa-Doc: une gestion de l'hétérogénéité des classifications documentaires en entreprise, Hypermédias &pratiques numériques (H2PTM'11),Metz, France,p.79.

ZHOU C., LEJEUNE, C.H. ET BENEL A. (2006). Towards a standard protocol for community driven organizations of knowledge. Proc. ISPE CE'06, IOS Press, Amsterdam, pp 338 - 349.