

Méthodes et défis du repérage d'images sur le Web : Jean et John cherchent-ils de la même manière ?

Elaine Ménard, Ph. D.
Professeure adjointe
School of Information Studies, McGill University
elaine.menard@mcgill.ca

Résumé

Nous faisons ici le point sur les principales méthodes employées pour le repérage d'images numériques que l'on trouve dans la majorité des sites et pages Web. Les méthodes d'indexation manuelles et automatiques facilitant le repérage d'images sont également décrites. Nous nous attardons plus particulièrement sur les problèmes rencontrés par les chercheurs d'images lors du repérage en contexte multilingue, c'est-à-dire lorsque la langue de la requête diffère de la langue d'indexation. Le principal objectif de cette étude était d'identifier les similarités et les différences dans la manière de formuler les requêtes de quatre groupes de participants de langue maternelle différente au moment du repérage d'images sur le Web. Nous avons demandé à quatre groupes de dix participants de quatre communautés linguistiques différentes (française, anglaise, russe et chinoise) de préciser quelles catégories de termes ils utilisent en général dans leurs requêtes lors de la recherche d'images sur le Web. Les résultats de notre étude sur les comportements de recherche des chercheurs d'images de quatre communautés linguistiques différentes ont révélé des différences significatives dans la manière de chercher (nombre de requêtes utilisées, longueur des requêtes, attributs conceptuels et terminologiques contenus dans les requêtes). Ces différences doivent être prises en considération au moment de l'élaboration des fonctionnalités de recherche proposées par les moteurs permettant la recherche d'images. En outre, la manière de chercher des images varie en fonction de la langue maternelle du chercheur d'images. Le défi de l'accès aux images, peu importe leur mode d'indexation, reste entier.

Abstract

This article examines the main methods used for the retrieval of digital images we find in most websites and Web pages. Techniques for manual and automatic indexing simplifying the image retrieval process are also described. We specifically focus on the problems encountered by searchers during the image retrieval process in a multilingual context: that is to say, when the language of the query differs from the indexing language. The main objective of this study is to identify similarities and differences in the way four groups of participants with different native languages formulate their queries when searching for images on the Web. We asked four groups of ten participants from four different linguistic communities (French, English, Russian and Chinese) to indicate the categories of terms they commonly use in their queries when searching for images on the Web. The results of our study on image searchers' behaviours of these four different linguistic communities revealed significant differences in the way of searching (number of queries used, query length, conceptual and terminological attributes contained in the queries). These differences

must be taken into consideration when developing the search functionalities offered by search engines dedicated to image retrieval. In addition, the image search process varies according to the image searcher's mother tongue. The challenge of accessing images regardless of the indexing method remains puzzling and complex.

Mots-clés

images numériques; recherche d'information multilingue; indexation; recherche d'images; multilinguisme

Keywords

digital images; cross-language information retrieval; indexing; image retrieval; multilingualism

Introduction

L'expression « recherche d'information » (*Information Retrieval*), souvent mentionnée dans les écrits en bibliothéconomie et sciences de l'information, réfère principalement aux procédés facilitant l'accès à l'information, peu importe sa source. Le principal objectif de la recherche d'information est de combler un besoin informationnel et de résoudre le problème à la base de ce besoin.

De nos jours, la question de l'accès à l'information multilingue est un sujet qui demeure d'actualité. L'information multilingue se définit de multiples manières dans la littérature. Lorsqu'un document (textuel ou autre) comporte plus d'une langue, on le considère comme bilingue, voire même multilingue. En 2011, « le » parfait système capable de traiter et de repérer l'information en contexte multilingue se fait toujours attendre. Certains moteurs de recherche s'en approchent grâce à certains mécanismes et fonctionnalités de recherche de plus en plus sophistiqués. Toutefois, il reste beaucoup de chemin à faire pour assurer un accès universel à de nombreux types de documents textuels ou multimédias.

Comme toute forme de documents que l'on retrouve sur le Web, l'image numérique pose sa part de problème au moment du repérage. Cet article propose de faire le point sur les principales méthodes employées pour le repérage d'images numériques que l'on trouve dans la majorité des sites et pages Web et s'attarde plus particulièrement sur les difficultés rencontrées par les chercheurs d'images lors du repérage d'images en contexte multilingue, c'est-à-dire lorsque la langue de la requête diffère de la langue d'indexation.

Comment cherche-t-on les images ?

La recherche par caractéristiques physiques

Les systèmes de recherche d'images par contenu (*Content-Based Image Retrieval* - CBIR) exploitent les caractéristiques physiques de l'image. Ces caractéristiques, généralement extraites automatiquement des fichiers d'images, sont considérées de bas niveau, par opposition aux éléments dits de haut niveau (termes d'indexation) attribués suite à l'analyse de l'image par un indexeur humain ou par traitement

automatique. Avec les systèmes de CBIR, les images sont décrites non par des éléments textuels, mais plutôt par des valeurs associées à certains paramètres tels que la couleur, la texture et la forme de l'objet représenté sur l'image. L'individu ne se voit pas dans l'obligation de conceptualiser sa requête avec des mots. Les contraintes linguistiques sont ainsi abolies (Boudry & Agostini, 2004).

Par exemple, il est possible d'indexer, puis de repérer des images en utilisant la similarité de couleurs obtenue par un histogramme identifiant la proportion de pixels de chaque couleur à l'intérieur d'une image (McDonald & Tate, 2003). Toutefois, l'utilisation de cette caractéristique de l'image demeure problématique pour l'individu au moment du repérage (Boudry & Agostini, 2004). En effet, il semble que l'œil humain soit capable de distinguer jusqu'à 10 000 nuances de couleurs différentes, mais des études ont également démontré que la majorité des individus sont uniquement capables de nommer une douzaine de couleurs au total (Jørgensen, 2003).

Par ailleurs, la texture de l'objet illustré sur l'image peut aussi être utilisée pour l'indexation de l'image, grâce à la modélisation selon un effet à deux dimensions utilisant des variations de gris (Goodrum, 2000). Des modèles mathématiques de la texture des images ont été peu à peu intégrés dans les systèmes de CBIR. Ces modèles ont prouvé leur utilité lors de la recherche d'images, plus particulièrement lorsqu'ils sont associés à d'autres caractéristiques physiques. Cependant, il demeure assez complexe de décrire automatiquement les différentes textures contenues dans une image.

Finalement, l'indexation des images peut se faire à l'aide de la forme. Cette caractéristique de l'image dépend d'une combinaison de plusieurs paramètres comme la couleur, le contour, l'ombrage et la profondeur. En général, les requêtes utilisant la forme sont effectuées par la sélection ou l'esquisse d'une image similaire fournie par le chercheur d'images. Cependant, en raison de la distorsion due à certains facteurs (perspective, ombre, etc.), la reconnaissance de la forme n'est pas aussi aisée à réaliser par un ordinateur. En effet, aucune technique ne parvient encore à identifier de manière automatique et consistante les sections de l'image à contraste élevé (*high-contrast regions*) aussi bien que l'œil humain (Jørgensen, 2003).

Quelques exemples

Depuis quelques années, les moteurs permettant le repérage d'images offrent des fonctionnalités de plus en plus sophistiquées. Il est maintenant possible de chercher des images avec des requêtes autres que textuelles. Google Images (Google, 2012) permet, par exemple, de faire glisser une image dans la zone de recherche pour trouver une image identique à l'image glissée. De plus, ce moteur permet d'importer une image directement dans la barre de recherche pour repérer un ou des images similaires. Le contenu de l'adresse URL peut également être utilisé pour la recherche.

D'autres moteurs proposent aux utilisateurs de dessiner ce qu'ils espèrent repérer et le système tentera de trouver une ou plusieurs images similaires. Le moteur Retrievr

(Retriever, 2012) par exemple, permet d'explorer et de chercher parmi une sélection d'images offertes par Flickr en traçant une simple esquisse (Flickr, 2012). Il est également possible de chercher avec une couleur dominante. Cette fonctionnalité est offerte par plusieurs moteurs (Multicolour Search Lab, 2012). Ainsi, certains moteurs permettent de chercher parmi des millions d'images offertes par Flickr Creative Commons (Flickr Creative Commons, 2012) en utilisant une technologie mettant en correspondance une couleur sélectionnée avec la couleur dominante des images à repérer. Il suffit de sélectionner une ou plusieurs couleurs de la palette de couleurs et la recherche s'effectue automatiquement. Il est également possible d'ajouter plus de couleurs, ou d'en enlever de manière à raffiner les résultats. D'autres moteurs vous permettront de chercher l'image de votre sosie parmi les photographies de célébrités (Facedouble, 2012) ou encore de repérer une image dans une catégorie spécifique (Art, Biologie, Paysage, Animal) (Macroglossa 2012) ou encore, d'identifier une personne aperçue dans la rue et croquée sur votre portable (Alipr, 2012).

Malgré certaines faiblesses sur le plan de l'utilisabilité (pertinence des résultats obtenus, vitesse d'exécution des tâches de repérage, satisfaction du chercheur d'images), les systèmes de CIBR semblent tout de même prometteurs. Par exemple, on considère que l'utilisation de l'approche basée sur les caractéristiques physiques de bas niveau de l'image peut être intéressante dans le cas du furetage, c'est-à-dire lorsque l'individu n'a pas une idée très précise de ce qu'il cherche (Markkula & Sormunen, 2000). Toutefois, on constate que les caractéristiques physiques extraites du contenu de l'image n'ont pas nécessairement de sens pour l'individu et que ces systèmes restent complexes à mettre au point et à utiliser. Pour le moment, les méthodes d'indexation et de repérage utilisant les caractéristiques physiques de l'image demeurent au stade expérimental (Jørgensen, 2003).

Compte tenu de ce qui précède, on peut conclure que le repérage à l'aide des différentes caractéristiques physiques des images pose toujours certains problèmes. Les systèmes permettant le repérage de l'image à l'aide de la couleur, la texture et la forme ne sont toujours pas utilisés à grande échelle par la majorité des utilisateurs. En outre, il demeure très difficile pour un système de CBIR de repérer toutes les images d'un coucher de soleil ou d'un enfant riant aux éclats par exemple, en se basant uniquement sur les caractéristiques physiques telles que la couleur, la texture ou la forme. Par conséquent, l'approche d'indexation à l'aide de la description textuelle de l'image demeure encore la voie la plus utilisée menant à son repérage.

La recherche textuelle ***Indexation manuelle d'images***

Le repérage de l'image peut se faire à l'aide du texte qui lui est associé ou encore avec l'information qui décrit l'image, c'est-à-dire les métadonnées créées manuellement ou automatiquement au moment de l'indexation. On distingue deux points de vue divergents au sujet du meilleur scénario à adopter quant au choix du schéma de métadonnées à utiliser pour la description de l'image : ceux qui estiment que les images sont si différentes du texte qu'il ne peut y avoir similitude dans les méthodes d'indexation, et ceux qui considèrent qu'un seul et même ensemble de règles peut être appliqué de manière satisfaisante à tous les types de documents, y

compris les images (Jørgensen, 2003). Dans son étude, Greenberg (2001) a comparé différents schémas de métadonnées utilisés pour la description des images. Elle révèle que le nombre d'éléments de métadonnées contenus dans un schéma varie en fonction de son domaine d'utilisation. Elle affirme également que les schémas de métadonnées employés pour la description des images devraient contenir principalement quatre classes de métadonnées : la découverte, l'utilisation, l'authenticité et la gestion (Greenberg, 2001). Par exemple, les métadonnées peuvent inclure de l'information sur le contenu de l'image et son origine (date, heure, luminosité, l'usage ou non d'un flash, etc.). Elles renferment parfois des données permettant d'identifier l'image à l'intérieur d'une collection (indexation sujet, catégorisation, usage d'un schéma de classification, etc.). D'autres métadonnées administratives et techniques sont essentielles à la gestion d'une collection d'images numériques. Par exemple, elles renseignent sur les technologies qui ont permis la production de l'objet informationnel et celles qui en permettront la reproduction, ainsi que sur les procédures qui permettent la gestion des données. L'image, en effet, offre plus de complexité que le texte pour un système d'information. Il faut également pouvoir rendre compte de la source originale (par exemple, une sculpture comme le David et la photo de la sculpture de David). Ce problème ne se pose évidemment pas avec les documents textuels.

Pour sa part, Baca (2003) avec son étude sur les principaux problèmes liés aux schémas de métadonnées révèle qu'il n'existe aucun schéma pouvant convenir à toutes les situations ou à tous les environnements. Cette étude suggère également qu'un choix inapproprié de schéma de métadonnées pour l'indexation des images peut nuire considérablement à leur repérage. En effet, les schémas de métadonnées prévoient les éléments servant à la description des images et qui seront éventuellement employés à leur repérage, tels que les termes d'indexation. Par conséquent, la sélection d'un schéma de métadonnées inapproprié ne sera d'aucune utilité pour ses éventuels utilisateurs. En outre, il est généralement peu recommandé d'appliquer aux ressources visuelles telles que l'image, des schémas de métadonnées développés formellement pour les documents textuels et qui démontrent souvent une nature trop générale pour bien décrire les images (Baca, 2003).

En règle générale, les schémas de métadonnées développés pour l'indexation des images suggèrent l'utilisation de vocabulaires contrôlés pour certains éléments, comme le sujet de l'image par exemple. L'utilisation des vocabulaires contrôlés pour l'indexation de l'image offre de nombreux avantages pour le repérage, le furetage et l'interopérabilité entre divers vocabulaires. Cependant, comme le soulignent Greisdorf et O'Connor (2008) ces vocabulaires ne sont pas, pour la plupart, assez riches pour fournir de l'information descriptive adéquate ou encore, des points d'accès utiles pour tous les types d'images. En outre, un autre désavantage de l'indexation à l'aide de vocabulaires contrôlés est que ceux-ci deviennent rapidement désuets. En effet, les néologismes mettront souvent beaucoup de temps à être intégrés aux différents vocabulaires contrôlés utilisés pour l'indexation.

Par ailleurs, l'indexation de l'image peut se faire également en vocabulaire libre, c'est-à-dire un type d'indexation qui n'est régi par aucune règle préétablie quant au choix des termes à employer pour décrire une image. Les termes d'indexation sont

tirés de la langue naturelle. À titre d'exemple de l'approche d'indexation en vocabulaire libre, soulignons l'apparition depuis quelques années d'un phénomène d'un genre nouveau. Il s'agit de l'indexation collaborative dont la principale caractéristique est de n'utiliser aucun langage documentaire ou vocabulaire contrôlé pour l'indexation (Macgregor & McCulloch, 2006). L'étiquetage collaboratif ou social désigne le processus avec lequel les utilisateurs attribuent eux-mêmes des mots-clés aux documents.

Étant donné sa grande popularité, l'indexation collaborative s'est rapidement répandue et ne concerne plus uniquement les documents textuels. En effet, les services en ligne de partage d'images ont également recours à ce genre d'indexation (Mathes, 2006 ; Angus, Thelwall & Stuart, 2008 ; Rorissa, 2008). Ces systèmes permettent l'indexation et le partage de ses propres images. À partir de leurs étiquettes (*tags*) ou mots-clés, les individus peuvent partager des images sur un même thème et créer une communauté d'échange. Ce type d'indexation a pour principal avantage de fournir des points d'accès supplémentaires et souvent très différents de l'indexation traditionnelle utilisant le vocabulaire contrôlé (Ménard, 2008). Les étiquettes peuvent ensuite être utilisées pour des recherches ciblées.

La principale caractéristique de l'indexation collaborative que l'on retrouve dans les systèmes de partage d'images est l'utilisation du vocabulaire libre, qui est la langue que l'individu parle et écrit couramment. Ainsi, pour l'indexation des nouveaux sujets, ce type d'indexation est susceptible de gagner haut la main. En effet, la précision sera meilleure puisque l'indexation en vocabulaire libre offre, en principe, un maximum de spécificité des termes d'indexation. En outre, le rappel est également susceptible d'être supérieur puisque le chercheur d'images n'a pas besoin de deviner quel sont les meilleurs termes à utiliser dans ses requêtes.

Les étiquettes peuvent prendre toutes les formes possibles, selon le désir de l'internaute et surtout, selon sa culture et sa maîtrise de la langue. Parallèlement, les autres utilisateurs du système ont la possibilité de mettre quotidiennement à jour l'indexation des images en ajoutant des mots-clés ou en inscrivant des commentaires sur toute image à laquelle ils ont accès. En outre, le vocabulaire libre que l'on trouve dans les systèmes de partage d'images peut être tiré d'une seule langue ou combiner plusieurs langues. L'indexation multilingue qui en résulte est ainsi avantageuse pour une majorité d'individus qui peuvent ainsi formuler des requêtes en multiples langues pour repérer des images. Cependant, un bémol à ce type d'indexation en vocabulaire libre doit être mentionné. En effet, les termes d'indexation attribués demeurent tributaires de l'habileté personnelle des indexeurs profanes qui souvent ont peu ou pas conscience de leur importance.

L'indexation collaborative réussira-t-elle à détrôner l'indexation traditionnelle en vocabulaire contrôlé ? L'utilisation du vocabulaire contrôlé ou du vocabulaire libre pour l'indexation documentaire est une question longuement débattue dans la littérature. Ces méthodes d'indexation manuelle demeurent fort coûteuses en temps et en argent. En outre, étant donné la croissance vertigineuse du nombre d'images disponibles sur le Web, il devient indispensable de pouvoir indexer celles-ci plus rapidement et à un moindre coût.

Indexation automatique de l'image

Puisque les systèmes permettant la recherche d'images à l'aide des caractéristiques physiques telles que la couleur, la forme ou la texture ne sont toujours pas complètement au point, la recherche d'images à partir d'une requête textuelle demeure la manière de chercher la plus répandue parmi les internautes. Le principal avantage de ce type de repérage comparé à la recherche d'images par contenu (CBIR) est que les requêtes textuelles peuvent être formulées plus naturellement par les chercheurs d'images (Jørgensen, 2003). Comme nous l'avons mentionné précédemment, dans les systèmes de CBIR l'utilisateur doit rechercher les images avec des concepts tels que la couleur, la forme ou la texture ou encore, donner en exemple une image pour laquelle on souhaite trouver des images similaires.

Pour leur part, les moteurs de recherche permettant la recherche d'images par requêtes textuelles proposent de mettre en correspondance ces requêtes à une forme de texte associé. Celui-ci peut être constitué par les informations contenues dans l'image elle-même (par ex., les métadonnées de l'International Press Telecommunications Council - IPTC), par les informations contenues dans une base de données externe associée à l'image, ou encore par le contexte ou l'environnement textuel immédiat de l'image.

Étant donné le nombre élevé d'images à indexer se retrouvant sur les pages Web disponibles, la plupart des moteurs de recherche utilisent le texte en périphérie de l'image pour l'indexation. Par exemple, Google Images (Google, 2012) analyse le texte de la page qui entoure l'image, le titre de l'image et de nombreux autres paramètres afin d'identifier le contenu graphique de l'image. Cette forme d'indexation automatique donne des résultats intéressants bien que parfois fantaisistes. La principale critique vis-à-vis de ce type d'indexation est le manque de pertinence de certains résultats affichés. Puisque la pertinence est déterminée grâce aux mots-clés contenus dans la requête de l'utilisateur sur la base de paramètres tels que le nom du fichier ou le texte encadrant l'image dans la page d'origine, il ne faut pas s'étonner que les images ne correspondent pas toujours à l'image cherchée. Par exemple, on retrouvera une image de la candidate républicaine Sarah Palin parmi les résultats affichés avec la requête « car » (*voiture*). L'image ayant été indexée par le nom de fichier « palin-in-the-car.jpg », l'image d'une Sarah Palin souriante et possiblement assise dans une voiture s'affiche parmi les Mercedes, les BMW et autres voitures (on ne voit toutefois pas la voiture). De la même manière, le moteur repérera l'image d'une jolie jeune fille répondant au nom de « Annick Car ». Ces résultats souvent sans rapport avec la requête formulée engendrent inévitablement une grande frustration chez les chercheurs d'images.

Par ailleurs, les moteurs de recherche d'images offrent des fonctionnalités de plus en plus raffinées. Ainsi, certains moteurs donnent la possibilité de chercher par couleur (incluant le noir et blanc), par taille, par type (visages, images clipart, photographies), ou encore par format (jpeg, gif). Depuis 2010, Google Images (2012) offre une nouvelle manière d'afficher les résultats. Les images s'affichent telle une mosaïque. Lorsque l'on amène la souris sur une image, celle-ci s'agrandit et le nom du site d'où provient l'image apparaît. De plus, en cliquant sur l'image l'utilisateur est dirigé sur le site éditeur. Depuis quelques mois, ce moteur de

recherche offre également la possibilité de raffiner les résultats obtenus par « sujet ». Cette nouvelle fonctionnalité de tri des images est par ailleurs disponible sur la plupart des versions linguistiques du moteur de recherche. Elle permet, une fois les résultats de la requête affichés, de trier ceux-ci selon une série de sous-catégories. Par exemple, les résultats obtenus avec la requête « chien » peuvent être triés par les catégories telles que chien drôle, chien dangereux, chien mignon, chien méchant, chien moche, petit chien, gros chien, bébé chien, beau chien, race de chien, chien berger, chien labrador, etc. Évidemment, le chercheur d'images doit alors se débrouiller pour faire la différence entre l'image d'un chien méchant et celle d'un chien dangereux, ou encore entre un chien mignon et un beau chien, puisque les tenants et les aboutissants de cette catégorisation demeurent bien mystérieux.

Cette nouvelle fonctionnalité pourrait bien s'avérer efficace pour affiner les requêtes des internautes, tout en leur faisant gagner du temps. Toutefois, le principe de fonctionnement reste le même que lors de la recherche textuelle initiale, c'est-à-dire que les résultats sont classés en fonction des mots les plus fréquemment associés à la requête. Cette catégorisation automatique est néanmoins intéressante puisqu'elle semble solutionner du même coup le problème de l'affichage de trop nombreux résultats proposant une même image provenant de différents sites.

La recherche d'images en contexte multilingue

Par ailleurs, un utilisateur peut désirer repérer des images indexées dans différentes langues qu'il connaît ou ne connaît pas. La recherche d'information multilingue (RIML) est intéressante et offre des possibilités fascinantes pour les institutions culturelles comme les bibliothèques et les musées. Ce type de recherche a pour objectif de satisfaire ce besoin en permettant aux utilisateurs de trouver les documents dans des langues différentes de celle de la requête. Selon Oard et Dorr (1996), la RIML est utile dans de nombreuses situations : (1) la collection regroupe des documents en plusieurs langues ; (2) les documents eux-mêmes sont écrits en plusieurs langues ; (3) l'individu ne connaît pas suffisamment la langue d'un document, mais veut quand même l'obtenir ; (4) la collection est indexée dans une langue non familière à l'individu ; (5) un chercheur veut connaître tout ce qui a été écrit sur un sujet précis, peu importe la langue ; et finalement, (6) l'individu possède les capacités nécessaires pour traduire un document dans une langue qu'il comprend.

Choix conceptuels et terminologiques

Dans une étude récente (Ménard, 2012), nous avons demandé à quatre groupes de dix participants de quatre communautés linguistiques différentes (française, anglaise, russe et chinoise) de préciser quelles catégories de termes ils utilisent en général dans leurs requêtes lors de la recherche d'images sur le Web. Le principal objectif de cette étude était d'identifier les similarités et les différences dans la manière de formuler les requêtes de quatre groupes de participants de langue maternelle différente au moment du repérage d'images sur le Web.

Ainsi, l'examen des termes utilisés dans les requêtes sur l'image a indiqué que certains termes sont utilisés plus fréquemment que d'autres. Par exemple, il n'est

pas surprenant que nous ayons observé que les termes spécifiques, noms composés, adjectifs de couleur et noms propres se retrouvaient fréquemment ou très fréquemment dans les requêtes de la majorité des participants. Cependant, d'autres types terminologiques tels que les superlatifs, abréviations, articles, adverbes n'étaient que rarement ou très rarement utilisés. Notre analyse a révélé également des similarités parmi les quatre groupes linguistiques au sujet de la fréquence d'utilisation des noms composés, adjectifs de couleur utilisés fréquemment par les participants, alors que les adjectifs de dimension n'étaient que rarement utilisés par les participants des quatre communautés linguistiques étudiées.

Les résultats de cette étude ont mis aussi en lumière des différences significatives entre les quatre groupes au sujet de l'utilisation des verbes et adjectifs qualificatifs dans la formulation des requêtes. En outre, les répondants de langue anglaise incluaient davantage d'adjectifs qualificatifs dans leurs requêtes, de même que de noms propres, proportionnellement aux participants des trois autres groupes linguistiques. La fréquence d'utilisation d'autres types de termes différait selon les communautés linguistiques. Par ailleurs, il est intéressant de noter que les participants anglophones ont confirmé qu'ils formulaient rarement des requêtes avec des termes tirés d'une autre langue, tandis que les participants des trois autres groupes linguistiques affirmaient eux inclure souvent des termes tirés d'une langue différente de leur langue maternelle lors de la recherche d'images numériques sur le Web.

Traduction des requêtes

D'autres différences significatives entre les quatre groupes linguistiques de chercheurs d'images ont également été observées lors de notre étude (Ménard, 2012). Ainsi, les résultats obtenus indiquaient également une différence significative de la longueur moyenne des requêtes formulées parmi les quatre groupes linguistiques à l'étude (2,8 termes/requête pour les participants de la communauté russe ; 2,7 termes/requête pour les participants francophones ; 2,4 termes/requête pour les répondants anglophones et 2,2 termes/requête pour les participants de langue chinoise).

La recherche d'images semble ainsi générer des requêtes plutôt courtes. En contexte de recherche d'information multilingue, cette constatation est importante. En général, deux approches sont utilisées par les systèmes de RIML. La première suppose la traduction de tous les documents dans toutes les langues. Le principal avantage de cette méthode est d'offrir une grande précision de recherche apparente puisqu'un texte plus long posera moins de problèmes de polysémie lors de la traduction. Cette manière de procéder s'avère cependant longue et coûteuse. De plus, traduire tous les documents dans toutes les langues semble une tâche difficilement réalisable et très encombrante en ce qui concerne le stockage.

La deuxième approche consiste à traduire la requête de l'individu dans la langue du ou des documents à repérer. Dans le cas du repérage d'images, cette deuxième approche s'avère plus logique puisque les images ne sont pas nécessairement associées à un texte pouvant être traduit. Ainsi, la solution adoptée par les systèmes

de RIML dédiés au repérage d'images est de traduire les requêtes qui pourront être en suite mises en correspondance avec le texte associé à l'image, le cas échéant, ou aux termes d'indexation. Le principal avantage de la traduction des requêtes est que celle-ci est moins coûteuse à mettre en œuvre que la traduction complète des documents. En effet, comme nous l'avons mentionné précédemment, le nombre de termes d'une requête est incontestablement beaucoup moins élevé que le nombre de mots contenus dans un document entier.

En général trois techniques sont utilisées dans les systèmes de recherche d'information multilingue pour la traduction des requêtes : (1) l'utilisation d'un dictionnaire bilingue ; (2) l'utilisation d'un système de traduction automatique ; (3) l'utilisation de textes parallèles. De plus en plus disponible dans les moteurs de recherche, la traduction automatique, ou logiciel de traduction automatique, est un processus utilisant un logiciel informatique pour transcrire un texte d'une langue source vers une autre. Ce processus fort populaire parmi les internautes et apparemment simple à réaliser est en réalité passablement complexe. La traduction ne se limite pas à une simple substitution mot à mot. Un traducteur humain doit analyser et interpréter le texte et comprendre les relations entre les mots qui peuvent influencer son sens. Ceci requiert une connaissance de la grammaire, de la syntaxe (structure de la phrase) et de la sémantique (sens des mots) à la fois dans la langue source et dans la langue cible. Malgré des avancées importantes, les systèmes de traduction automatique ne parviennent pas toujours à fournir une traduction de qualité.

Il serait légitime de croire que les requêtes utilisées pour le repérage d'images peuvent être correctement traduites par un système de traduction automatique puisqu'en général celles-ci ne contiennent que peu de mots, c'est-à-dire en moyenne 3,7 termes par requête, tel que constaté par plusieurs chercheurs (Goodrum & Spink, 2001 ; Chen, 2001 ; Spink & Jansen, 2004 ; Pu, 2005). Cependant, une traduction mot à mot, ou phrase par phrase, n'aboutira pas nécessairement au texte original. Le sens porté par les mots doit être pris en considération également. La traduction a besoin de contexte et les requêtes formulées pour repérer des images sont souvent trop courtes, c'est-à-dire qu'elles ne fournissent pas suffisamment de contexte pour être bien interprétées et par conséquent, sont souvent incorrectement traduites. Par conséquent, la traduction automatique fait souvent face à de nombreux problèmes d'ambiguïtés sémantiques et syntaxiques. Cette ambiguïté principalement due à un manque de contexte peut avoir comme conséquence la production de contresens. Ceci constitue d'ailleurs l'un des principaux problèmes des systèmes de RIML offrant un mécanisme de traduction automatique (Kishida, 2005). Comme résultat, les images repérées et affichées ne correspondent pas toujours à l'image désirée ou encore, les requêtes dont la traduction correspondante est erronée ne génèrent aucun résultat.

Cette déficience des systèmes de traduction automatique confirme les propos de Fluhr (2004) qui soulignait que « la qualité des systèmes de traduction [automatique] actuels ne peut assurer que cette traduction est bonne. En cas de mauvaise traduction des concepts de la requête, la recherche risque de ne pas aboutir ». Par exemple, les requêtes « bicyclette » ou « vélo » devraient en principe être facilement traduites par l'équivalent anglais *bicycle*. Cependant, il est possible que selon la

forme d'indexation choisie, un terme anglais différent (*bike*) soit retenu pour indexer une image contenant cet objet. Or, le mot *bike* traduit en français par le même système de traduction automatique, correspond au mot « vélo » et non pas au mot « bicyclette ». Autrement dit, si un individu utilise le mot « vélo » dans sa requête il peut repérer l'image représentant cet objet, mais pas s'il emploie le mot bicyclette, ce qui n'est pas juste puisque les mots « vélo » et « bicyclette » sont considérés comme des synonymes.

Conclusion

La recherche d'images en contexte multilingue (lorsque la langue de la requête diffère de la langue d'indexation (manuelle ou automatique) reste problématique. Il s'agit souvent d'un problème de traduction. Dans de nombreux contextes, la traduction française est souvent incompréhensible et ne génère que peu de résultats pertinents. La faute en revient en partie au recours à des logiciels de traduction automatique. Économiques et pratiques ? Malgré tous les progrès de la technologie depuis l'apparition du premier système de traduction automatique vers le début des années 1950, la machine ne peut vraiment remplacer le traducteur humain. Les exemples cocasses qui continuent à se multiplier le démontrent bien. D'ailleurs, les concepteurs de logiciels de traduction automatique le reconnaissent eux-mêmes. Google Traduction (2012) indique : « Même les logiciels actuels les plus perfectionnés ne peuvent maîtriser une langue aussi bien qu'une personne de langue maternelle ou posséder les compétences d'un traducteur professionnel. La traduction automatique est un domaine extrêmement complexe, car la signification des mots dépend du contexte dans lequel ils sont utilisés. La mise en place d'un service de traduction automatique rapide et efficace risque de prendre encore un certain temps ».

Pour sa part, Bing Translator (2012) avertit ses utilisateurs en toute bonne foi : « La qualité d'une traduction par une machine des plus perfectionnées est encore aujourd'hui bien en-dessous de la précision et de la fluidité d'une traduction humaine, et de nombreuses phrases sont encore incompréhensibles. Les chercheurs travaillent sans relâche à des améliorations, mais cela peut prendre encore de nombreuses années avant qu'une machine puisse offrir une traduction de grande qualité. C'est pourquoi le texte original et sa traduction sont proposés, afin que vous puissiez plus facilement comprendre la traduction en vous référant au contenu original. ». Il est certain que les systèmes de traduction automatique existants peuvent parfois être utiles pour permettre à un individu, qui ne connaît pas la langue dans laquelle un document est rédigé, de se faire une idée générale de ce sur quoi porte le document. Toutefois, même dans le meilleur des cas, la traduction automatique n'est pas de qualité suffisante pour communiquer correctement un message dans une autre langue. Dans le cas de la recherche d'images, compter sur un système de traduction automatique équivaut sans nul doute à s'acheter un billet de loterie et espérer que le bon numéro sortira.

Par ailleurs, il ne faudrait pas penser que les difficultés rencontrées par les chercheurs d'images en contexte de repérage multilingue n'est qu'un simple problème de traduction. En effet, les résultats de notre étude sur les comportements de recherche des chercheurs d'images de quatre communautés linguistiques

différentes ont révélé des différences significatives dans la manière de chercher (nombre de requêtes utilisées, longueur des requêtes, attributs conceptuels et terminologiques contenus dans les requêtes) (Ménard, 2012). Ces différences doivent être prises en compte au moment de l'élaboration des fonctionnalités de recherche d'images.

En outre, il faut penser que la manière de chercher des images varie en fonction de la langue maternelle du chercheur d'images. Puisque le repérage d'images à l'aide des caractéristiques physiques est toujours à l'état embryonnaire, les chercheurs d'images doivent effectuer leurs recherches à l'aide du vocabulaire qu'ils connaissent, souvent sans trop savoir où cela va les mener. Trop fréquemment, les chercheurs d'images utilisent des expressions obscures, alambiquées, voire même métaphoriques dans leurs requêtes pour trouver des images. Le processus d'indexation des images s'avère crucial pour en assurer leur repérage. Indexation par caractéristiques physiques, indexation automatique à l'aide d'algorithmes sophistiqués, indexation manuelle traditionnelle à l'aide de vocabulaire contrôlé ou par étiquetage social, combinaison de l'une ou l'autre de ces méthodes ? Toutes ces méthodes semblent offrir des possibilités passionnantes pour l'accès aux ressources visuelles. Réussiront-elles à combler la distance entre les langues de manière à connecter des gens qui pensent au delà des frontières linguistiques ? Le défi de l'accès aux images, peu importe leur mode d'indexation, reste entier.

Biographie

Elaine Ménard est professeure adjointe à la School of Information Studies, Université McGill (Montréal, Canada). Son domaine d'enseignement inclut le catalogage, l'indexation, la classification et le repérage d'information. Ses recherches portent principalement sur la recherche d'information multilingue, l'indexation d'images, et les métadonnées. Elle a publié dans un certain nombre de revues savantes, y compris *The Canadian Journal of Library and Information Practice and Research*, *The Indexer*, *Journal of Information Ethics*, *Knowledge Organization*, *Library Resources & Technical Services*, *Library Hi Tech*, et *Documentation et bibliothèques*.

Bibliographie

Angus, Emma, Mike Thelwall et David Stuart. "General Patterns of Tag Usage Among University Groups in Flickr." *Online Information Review*. 32.1 (2008): 89-101. Print.

Automatic Photo Tagging and Visual Image Search - ALIPR. 21 mars 2012. Web.

Baca, Murtha. "Practical Issues in Applying Metadata Schemas and Controlled Vocabularies to Cultural Heritage Information." *Cataloging & Classification Quarterly*. 36.3/4 (2003): 47-55. Print.

Bing Translator. 21 mars 2012. Web.

Boudry, Christophe and Clemence Agostini. "Étude comparative des fonctionnalités des moteurs de recherche d'images sur internet." *Documentaliste*. 41.2 (2004): 96-105. Print.

Chen, Hsin-liang. "An Analysis of Image Queries in the Field of Art History." *Journal of the American Society for Information Science and Technology*. 52.3 (2001): 260-73. Print.

Face Double. 21 mars 2012. Web.

Flickr Creative Commons. 21 mars 2012. Web.

Fluhr, Christian. "Les SRI multilingues." Dans *Méthodes avancées pour les systèmes de recherche d'informations*, éd. par M. Ihadjadene. Paris: Hermès Science, 2004: 117-36 Print.

Goodrum, Abby A. "Image Information Retrieval: an Overview of Current Research." *Informing Science*. 3.2 (2000): 63-67. Print.

Goodrum, Abby et Amanda Spink. "Image Searching on the Excite Web Search Engine." *Information Processing and Management*. 37.2 (2001): 295-311. Print.

Google Images. 21 mars 2012. Web.

Google Traduction. 21 mars 2012. Web.

Greenberg, Jane. "A Quantitative Categorical Analysis of Metadata Elements in Image-Applicable Metadata Schemas ." *Journal of the American Society for Information Science and Technology*. 52.11 (2001): 917-24. Print.

Howard F. Greisdorf et Brian C. O'Connor. *Structures of images collections: from Chauvet-Pont d'Arc to Flickr*. Westport, Conn: Unlimited Libraries, 2008. Print.

Jørgensen, Corinne. *Image Retrieval : Theory and Research*. Lanham MD: Scarecrow Press, 2003. Print.

Kishida, Kazuaki. "Technical Issues of Cross-Language Information Retrieval: a Review." *Information Processing and Management*. 41.3 (2005): 433-55. Print.

Macgregor, George et Emma McCulloch. "Collaborative tagging as a knowledge organisation and resource discovery tool." *Library Review*. 55.5 (2006): 291-300. Print.

Macroglossa. 21 mars 2012. Web.

- Markkula, Marjo et Eero Sormunen. "End-User Searching Challenges Indexing Practices in the Digital Newspaper Photo Archive." *Information Retrieval*. 1.4 (2000): 59-85. Print.
- Mathes, Adam. *Folksonomies - Cooperative Classification and Communication Through Shared Metadata*. 21 mars 2012. Web.
- McDonald, Sharon et John Tait. "Search Strategies in Content-Based Image Retrieval." Dans *SIGIR 03*, Toronto Canada, 28 juillet-1er août 2003. 21 mars 2012. Web.
- Ménard, Elaine. "Formulation de requêtes pour le repérage d'images : une étude comparative de quatre groupes linguistiques." Dans *L'organisation des connaissances, dynamisme et stabilité*, Traité des sciences et techniques de l'information, Mustafa Widad El Hadi (Éd.). Paris, Hermès-Lavoisier, Collection : Structuration de l'information, 2012. Print.
- Ménard, Elaine. *Étude sur l'influence du vocabulaire utilisé pour l'indexation des images en contexte de repérage multilingue*. Doctoral Dissertation. Montréal: Université de Montréal, École de bibliothéconomie et des sciences de l'information. 21 mars 2012. Web.
- Multicolour Search Lab. 21 mars 2012. Web.
- Oard, Douglas W. et Bonnie J. Dorr. "A Survey of Multilingual Text Retrieval." Technical Report. University of Maryland, Institute for Advanced Computer Studies, UMIACS-TR-96-19: 1996. 21 mars 2012. Web.
- Pu, Hsiao-Tieh. "A Comparative Analysis of Web Image and Textual Queries." *Online Information Review*. 29.5 (2005): 457-67. Print.
- Retrievr. 21 mars 2012. Web.
- Rorissa, Abebe. "User-Generated Descriptions of Individual Images Versus Labels of Groups of Images: a Comparison Using Basic Level Theory." *Information Processing & Management*. 44.1 (2008): 1741-53. Print.
- Spink, Amanda et Bernard J. Jansen. *Web Search : Public Searching of the Web*. Boston: Kluwer Academic Publishers, 2004. Print.