

# L'avenir du traitement plan par plan des images animées

Commençons par une distinction terminologique. Pour les francophones des deux côtés de l'Atlantique l'expression *image fixe* désigne un même concept, mais ce n'est plus le cas lorsqu'il s'agit d'*images animées*. En France, le terme *image animée* désigne plutôt l'image de type cinéma ou vidéo, alors qu'au Québec on préfère *image en mouvement* pour désigner ce concept. Au Québec, *image animée* suggère l'image en mouvement créée par technique d'animation – *Babar* ou *Tintin*, par exemple –, alors qu'en France, *image en mouvement* suggère l'image fixe qui se trouve être en mouvement parce qu'on se promène en projetant l'image.

**James M. Turner**

École de bibliothéconomie  
et des sciences de l'information  
Université de Montréal  
james.turner@umontreal.ca

Dans cet article, on s'en tiendra à l'usage français, mais il me semblait important de préciser ce détail pour celles et ceux qui consulteront d'autres textes que j'ai déjà publiés.

Du point de vue des professionnels des sciences de l'information, il importe de bien distinguer entre images fixes et images animées, parce que leurs pratiques de catalogage et d'indexation sont très différentes. Voici, en résumé, les différences les plus importantes.

## Image fixe

Dans un premier temps, il s'agit de faire la distinction entre les principaux types d'images fixes. On adoptera utilement une classification en trois grands types : l'image d'art, l'image documentaire et l'image « ordinaire » (la photographie de tous les jours, l'imagerie commerciale, etc.). Cette distinction entre ces trois types se traduit par des pratiques différentes d'organisation et de traitement

documentaire, qui méritent que l'on s'y attarde.

*L'image d'art* se caractérise par l'existence, dans la notice catalographique, de l'équivalent d'un auteur et d'un titre, ne serait-ce qu'un titre fourni par le catalogueur, tel *Nature morte*. Par ailleurs, l'art fait l'objet d'études, et il se crée autour d'une œuvre une masse documentaire que les systèmes d'information voudraient bien signaler. Les images d'art sont donc celles qui requièrent le traitement le plus complet et le plus élaboré.

*L'image documentaire*, c'est le type d'image, surtout photographique aujourd'hui, que l'on trouve dans les archives historiques, dans les agences de presse et dans les archives des médias. On peut aussi y inclure les photos familiales. Selon Barthes, une photo de presse n'est jamais sans commentaire écrit. Ainsi, sa structure n'est pas une structure isolée ; elle est reliée à au moins une autre structure, celle du texte. Le traitement documentaire de ce type d'image est

Docteur en sciences de l'information de l'université de Toronto, **James M. Turner** a enseigné à l'université de Montréal avant de devenir professeur à l'École de bibliothéconomie et des sciences de l'information de cette même université. Il est l'auteur de plusieurs articles sur l'indexation des images, et de l'ouvrage *Images en mouvement : stockage, repérage, indexation* (Presses de l'université du Québec).

caractérisé par l'association d'une légende ou d'informations supplémentaires (métadonnées) destinées à identifier l'image, son contenu, son contexte. Soulignons que la valeur de l'image disparaît souvent lorsque manque le texte accompagnateur.

L'image « ordinaire », c'est l'image de tous les jours, le *clip art* commercial, la photo d'illustration, etc., dont le traitement documentaire est très sommaire. Sauf exception, le photographe est sans importance, et un titre ou une légende sont inutiles. Ces images sont surtout recherchées en tant qu'illustration par le « chercheur », l'enseignant et le graphiste, et l'objet précis de leur recherche est rarement connu. Ainsi, une organisation très simple par grandes catégories (géographie, jeux, animaux, etc.) suffit normalement, et le mode de consultation privilégié consiste à balayer de nombreuses images pour trouver ce dont on a besoin. Pour ce type d'image, il est inutile de cataloguer et d'indexer, sauf par catégories très générales.

D'un point de vue technologique, des équipements et des logiciels à prix abordable sont disponibles pour la gestion de l'image fixe. L'image en ligne peut aussi être incluse dans une base de données. Par contre, lorsqu'il s'agit de l'image animée, certains obstacles technologiques restent à franchir, comme on le verra plus loin. Pour celle-ci, les bases de données peuvent contenir des métadonnées, mais pas des données, comme dans le cas de l'image numérique animée en ligne.

### Image animée

Comme pour les images fixes, on retrouve dans les images animées les mêmes types (images d'art, images documentaires, images « ordinaires » ou *stockshots*), mais c'est la précision du catalogage et de l'indexation qui importe pour les gestionnaires de collections. Dans le cas des images animées, on distingue les produits finis des éléments de production.

On peut dire que le traitement des produits finis – tel un film ou une émission de télévision – peut se faire à peu près comme celui d'un livre.

L'utilisateur peut visionner l'image fixe en toute liberté pour une durée très courte ou très longue, alors que le visionnement de l'image animée prend un temps défini pour le spectateur

Bien sûr, le livre n'a souvent qu'un seul auteur, alors qu'une production audiovisuelle nécessite toute une équipe pour sa création. En outre, d'autres éléments interviennent que l'on ne trouve pas dans les livres comme le temps de projection, la nature du produit en noir et blanc ou en couleurs, etc. Mais, globalement, on peut dire que le traitement des produits finis ressemble beaucoup à celui des livres.

### Le traitement plan par plan

En revanche, pour certaines collections d'images animées, il est nécessaire d'effectuer un traitement plan par plan ou séquence par séquence. Cela se passe notamment

ainsi dans les archives télévisuelles pour les informations d'actualités et pour les cinémathèques *stockshots*, ou autres collections de chutes.

Le coût élevé du traitement de ces collections plan par plan se justifie en arguant, dans le premier cas, qu'il est nécessaire au bon fonctionnement de la salle des nouvelles<sup>1</sup>, et, dans le deuxième cas, que l'utilisation des collections génère des revenus. Cependant, dans beaucoup d'autres cas, un catalogage et une indexation à un tel niveau de précision seraient souhaitables. À titre d'exemples, imaginons le critique de cinéma qui voudrait citer un moment précis dans un film ou dans une émission de télévision, ou l'étudiant qui compare certains événements survenus dans une série à la télévision, ou encore le chercheur en quête de toutes les occurrences d'une locution, d'un événement, etc., à travers une gamme de productions.

Les différences les plus importantes entre le traitement de l'image fixe et celui de l'image animée proviennent de ce que l'utilisateur peut visionner l'image fixe en toute liberté pour une durée très courte ou très longue, alors que le visionnement de l'image animée prend un temps défini pour le spectateur. En pratique, l'utilisateur à la recherche d'images fixes peut consulter le contenu d'une base de données au rythme d'environ 10 images/seconde (Ding *et al.*, 1997), ou 600 images/minute.

La consultation de l'image animée est plus complexe. D'une part, dans les systèmes actuels, l'image animée est rarement proposée en ligne. La consultation de chaque plan ou de chaque séquence nécessite par conséquent le repérage de ce plan sur une bobine ou une cassette dans le magasin de stockage, le transport de ce support vers un appareil de visionnement, et enfin la localisation du plan précis sur la bobine ou la cassette. On

1. Il s'agit de la salle, dans les chaînes de radio ou de télévision, où les informations sont reçues, rédigées et montées en vue de leur diffusion...

voit donc facilement combien il est utile d'inclure dans la base de données une représentation photographique du plan, mais surtout un texte de quelques lignes qui décrit l'action. Ce texte s'appelle le *synopsis visuel*.

D'autre part, cette représentation textuelle ne sert pas uniquement à dépanner en attendant qu'on puisse consulter l'image animée en ligne avec ses métadonnées. Même dans les systèmes qui offrent l'image animée au sein de la base de données, il est fort utile d'avoir accès à une description textuelle de ce qui se passe dans l'image animée, parce qu'il est facile de lire rapidement une longue liste de descriptions textuelles courtes. L'utilisateur peut donc éliminer les plans repérés qui s'avèrent inutiles, et réaliser ainsi des économies considérables de temps en ne visionnant que les plans les plus pertinents.

Les collections d'images animées dont le traitement se fait plan par plan exigent un traitement plus élaboré. On a parlé du synopsis visuel, mais l'indexation est un autre élément important pour le repérage. En pratique, c'est d'abord par une recherche dans les termes d'indexation que l'utilisateur va pouvoir construire une liste préliminaire de plans à consulter. Nous discuterons plus loin des divers outils employés dans ces collections. Pour le moment, attachons-nous sur les enjeux de l'indexation plan par plan.

### Les enjeux de l'indexation plan par plan

Les collections d'images animées sont rarement indexées plan par plan en raison des coûts que cela entraîne. Comme toute indexation, l'indexation plan par plan coûte cher, mais nous savons que si on ne la fait pas, la tâche incombera à l'utilisateur qui en subira les conséquences. Elle n'est effectuée actuellement que dans des situations particulières, bien qu'on ait vu qu'elle serait souvent très utile. La solution réside dans l'automatisation

des processus, la clé du succès étant l'emploi de textes pré-existants pour générer automatiquement l'indexation pertinente.

L'indexation automatique de textes remonte à environ quarante ans, à Cranfield en Angleterre, où l'on faisait des expériences de ce genre dans les années 1960 (Cleverdon, 1991). Ces

La solution réside  
dans l'automatisation  
des processus,  
la clé du succès  
étant l'emploi de textes  
pré-existants pour générer  
automatiquement  
l'indexation pertinente

tests étaient basés sur des notions cognitives plutôt simplistes ; on a affaire aujourd'hui à des logiciels de langues, de linguistique informatique, etc., et les progrès ont été considérables. En revanche, on ne peut toujours pas créer de façon automatique un index conceptuel convenable pour un livre, et l'on doit se contenter de logiciels informatiques pour assister l'indexeur humain.

Ce qui nous intéresse, c'est que la théorie générale qui régit le comportement du texte dans l'analyse automatique ne semble pas s'appliquer aux textes les plus utiles dans l'automatisation de l'indexation des images animées. C'est ce que nous démontrèrent apparemment des études réalisées depuis quelques années.

*Première observation* : ceux qui recherchent des images autres que des images d'art nomment les personnes, les objets ou les événements qu'ils souhaitent voir représentés sur l'image (Turner, 1994). Dans la description des images, on mentionne

de temps en temps d'autres aspects comme une forme, une émotion ou des couleurs présentes dans l'image ; toutefois, les données recueillies auprès de ceux qui participent à ces expériences se rapportent presque toujours à des personnes, des objets ou des événements.

Une autre caractéristique de ces objets informationnels (chutes de film dans le cas des études dont il est ici question) est qu'il y a une forte correspondance entre les termes qu'utilisent des indexeurs professionnels pour ces plans, et ceux qui sont le plus souvent nommés par des participants à qui on demande de décrire l'image par des mots-clés qu'on peut employer pour le repérage du plan. Par ailleurs, les termes sélectionnés pour l'indexation des plans (en nommant des personnes, des objets et des événements dans le plan) se trouvent aussi presque toujours dans le synopsis visuel (Turner, 1995). Il apparaît donc qu'on peut automatiser le processus d'indexation, du moins en grande partie, en générant automatiquement, à partir de textes déjà existants, des mots-clés qui serviront de termes d'indexation à l'image correspondante.

On constate également qu'il est possible, à partir de cet index, d'en générer d'autres concernant le même plan dans d'autres langues. Pourquoi ? Parce qu'on n'a pas affaire ici à du texte continu, mais plutôt à des listes de mots ou d'expressions. Ainsi, on élimine tous les problèmes grammaticaux qui désorientent les logiciels de traduction. Il s'agit de trouver le terme correspondant dans d'autres langues : en tenant compte des synonymes les plus probables, on arrive à une correspondance très élevée.

Les résultats très encourageants que nous avons obtenus dans nos recherches depuis quelques années, et notamment les taux de correspondance entre les termes donnés par des participants et ceux donnés par des indexeurs professionnels, diffèrent des résultats obtenus en sciences de l'information lorsqu'on analyse

des textes. Nous avons reproduit certaines expériences afin de confirmer les résultats, résultats qui s'expliquent probablement par la nature des objets informationnels. Il s'agit d'objets et d'événements de tous les jours (par exemple les autos, la circulation), qui font partie de l'expérience de tout un chacun. Il est normal que tout le monde nomme chaque objet ou événement par son nom habituel. Soulignons que cet état de fait est vrai pour l'image « ordinaire » dans toutes ses manifestations, mais que, lorsqu'il s'agit de métrage documentaire, d'un événement précis, ou encore d'un film d'art, il faudrait avoir aussi une connaissance spécialisée pour pouvoir identifier le contenu des images. Dans ce cas, on fera appel à de la documentation spécialisée ou à l'indexeur humain pour bien décrire et indexer le document.

Il reste maintenant à construire un système pour tester ces résultats intéressants qui constituent le cœur des

**Les résultats de cette étude ont démontré que l'emploi d'un thésaurus aux fins d'indexation des images animées est loin d'être universel**

activités de recherche qui visent à automatiser le processus. C'est ce que nous espérons faire dans le cadre des travaux des prochaines années. Comme on cherche à construire un système en réseau, il faut faire appel aux ressources actuelles et futures du

Outils utilisés par les 14 collections pour l'indexation des plans	
Mots-clés simples	7 collections
Classification	3 collections
Thésaurus	6 collections
Vedettes-matière	6 collections
« Autres » (non spécifiés)	2 collections

web et employer des protocoles lisibles par les outils du web. Un tel système devra aussi tenir compte des travaux des informaticiens qui font des manipulations statistiques sur les primitives<sup>2</sup> des images. Les informaticiens et les professionnels en sciences de l'information peuvent ainsi travailler ensemble pour concevoir les systèmes les plus performants (Turner, 2000).

**Approches de l'indexation**

Nos recherches récentes ont permis de dresser un tableau des outils d'indexation des collections nord-américaines d'images animées, dont l'unité documentaire est le plan (Hudon *et al.*, 2001). On cherchait à étudier spécifiquement le thésaurus comme outil d'indexation, dans le but d'évaluer la faisabilité d'un thésaurus universel pour l'image animée. Nous avons pu analyser les outils de quatorze collections dans onze institutions.

Les résultats de cette étude ont démontré que l'emploi d'un thésaurus aux fins d'indexation des images animées est loin d'être universel. Six de ces onze institutions utilisent un outil qu'on peut qualifier de thésaurus, mais les pratiques varient et ces thésaurus ne sont pas nécessairement construits selon des normes précises. À titre d'exemple, rares sont les thésaurus étudiés qui séparent les noms communs des noms propres, repèrent les angles de caméra ou les descripteurs géographiques. Signalons que le nombre moyen de termes servant à indexer chaque plan est de douze, selon les huit institutions qui ont fourni des données sur ce point.

Pour ce qui est de la précision de l'indexation, onze des quatorze collections le font au niveau du titre, cinq au niveau de la séquence, huit au niveau du plan, et cinq à ces trois niveaux. Cinq ont dit indexer à d'autres niveaux (par exemple une bobine ou une cassette).

Nos recherches actuelles et à venir auprès de systèmes européens permettront de comparer les méthodes en vigueur sur les deux continents.

**L'harmonisation des pratiques**

Quels sont les enjeux pour l'avenir de l'organisation des collections en réseau ? Actuellement, c'est l'anarchie qui règne dans la gestion des collections. Si certaines solutions théoriques commencent déjà à voir le jour, en pratique elles ne seront en place que dans plusieurs années.

Ce qu'il faut pour permettre la communication entre systèmes, la découverte de l'existence de ressources et l'échange de fichiers, ce sont des métadonnées communes. Certaines approches envisagent des structures communes de bases de données, mais ce n'est pas une solution pratique. Même d'un point de vue théorique, on constate que, selon la nature archivistique plutôt que bibliothéconomique des collections d'images animées cataloguées et indexées plan par plan - sachant que le matériel de chaque collection est unique et que les besoins des usagers sont très variés -, il est logique que chaque collection possède son organisation particulière. C'est la situation que l'on retrouve le plus souvent actuellement. Chaque structure de base de données contient les métadonnées nécessaires pour gérer les collections, mais si on pouvait s'entendre sur un noyau commun de métadonnées, on pourrait créer un certain niveau de communication entre systèmes.

C'est exactement ce que tentent de faire certaines initiatives récentes. Parmi celles qui sont pertinentes pour

2. En informatique, éléments de base, pixel ou groupe de pixels ayant quelque chose en commun. Les primitives sont aussi des petits programmes qui ont pour but d'effectuer une fonction répétitive.

notre travail, mentionnons le Dublin Core Metadata Initiative (Dublin Core 2001), dont le noyau est déjà très utilisé dans beaucoup de situations. Un document en français concernant l'utilisation du Dublin Core est d'ailleurs disponible (Teasdale, 2001). Le VRA Core Categories (VRA Core, 2001), noyau de métadonnées développé par la Visual Resources Association pour le traitement de l'image fixe, en est maintenant à sa troisième version. Enfin, citons le Digital Audiovisual Council (DAVIC, 2001), le Preservation Metadata for Digital Collections (pmeta, 2001), et la norme MPEG7 (MPEG7, 2001).

L'implantation de l'International Standard Audiovisual Number ou Numéro international normalisé des œuvres audiovisuelles (ISAN 2001), équivalent de l'ISBN pour les produits audiovisuels, contribuera à l'identification unique de documents audiovisuels. On espère que cette norme sera prête pour la fin de 2001.

L'Universal Preservation Format (UPF, 2001) se voulait un contenant destiné à assurer la longévité des données numériques. Le concept était prometteur, mais les travaux semblent arrêtés. La dernière mise à jour du site web date de décembre 1999.

Parmi les nombreuses ressources offertes par le J. Paul Getty Trust de Los Angeles pour la gestion de collections culturelles, le Getty Research Institute offre des passerelles pour les métadonnées (*metadata crosswalks*). Celles-ci servent de traducteur entre les divers systèmes d'ensembles de métadonnées en transférant automatiquement les étiquettes d'un ensemble à un autre (Getty Research Institute, 2001).

Plus encourageants encore sont les travaux du World Wide Web Consortium (W3C, 2001) sur le Resource Description Framework (RDF, 2001), un métacontenant<sup>3</sup> pour des métadonnées de toutes sortes. Ce conte-

3. Contenant renfermant d'autres contenants, chacun comportant un ensemble de métadonnées.

nant est très flexible et servira de bassin de structures de données pour permettre aux logiciels de traduction de métadonnées de construire les ponts souhaités entre les divers systèmes. Pour que tout cela marche convenablement, il faudra compter

Il est évident que,  
dans le nouvel  
environnement en réseau  
où nous vivons,  
il faut uniformiser  
dans une certaine mesure  
les pratiques d'organisation  
des collections

plusieurs années de travaux avant de s'entendre sur les noyaux de métadonnées à utiliser. Ces travaux vont déjà bon train, mais il faudra compter quelques années pour créer les nouvelles métadonnées nécessaires, ou encore pour traduire l'encodage de métadonnées existantes vers les formats qu'exigent les nouvelles normes.

Cette liste des initiatives ne prétend pas être exhaustive. Il y en a beaucoup d'autres, mais celles qui sont mentionnées suffisent pour donner une idée des tâches à conduire pour harmoniser les pratiques. Plus les usagers adopteront rapidement ces normes en devenant, plus on pourra assurer la pérennité des normes et des documents dont les institutions ont la garde.

## Conclusion

Contrairement aux méthodes bibliothéconomiques développées au cours d'une période beaucoup plus longue, les méthodes de gestion des collections d'images animées ne présentent plus actuellement aucune forme d'uniformisation. Pourtant, il

est évident que, dans le nouvel environnement en réseau où nous vivons, il faut uniformiser dans une certaine mesure les pratiques d'organisation des collections. Lorsqu'on considère que ce nouvel environnement n'existe que depuis sept ans à peine, on constate cependant que les initiatives de normalisation de métadonnées témoignent d'une reconnaissance assez rapide de cette nouvelle problématique.

Si on était encore dans les années 1970, on dirait qu'il faut trouver les moyens pour que tout le monde s'organise de la même façon, mais ce n'est plus le cas. Pour toutes sortes de raisons, il importe de conserver des méthodes locales, lesquelles reflètent les particularités des collections, les besoins variés des usagers et la nature unique des documents qui se trouvent dans les collections. De ce point de vue, la situation anarchique actuelle n'est pas tragique. Il reste qu'il y a beaucoup à faire.

L'approche générale qui pourrait nous guider dans la gestion d'images animées serait de privilégier surtout des méthodes d'organisation variées en fonction des besoins locaux. Toutefois, il est essentiel que la communauté professionnelle des sciences de l'information continue à investir temps et énergie, afin de développer des noyaux de métadonnées appropriées et des couches logicielles de traduction entre les ensembles de métadonnées, dans le but de faciliter le repérage et les recherches que mènent les usagers des banques d'images, qu'elles soient fixes ou animées.

Mai 2001

## RÉFÉRENCES BIBLIOGRAPHIQUES

CLEVERDON, C.W., « The significance of the Cranfield tests on index languages », In Bookstein, A. ; Chiamarella, Y. ; Salton, G. ; Raghavan, V.V., editors, *Proceedings of the 14th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 3-12, Chicago, ACM Press, 1991.

**DAVIC**, URL : <http://www.davic.org>  
[consulté le 11 mai 2001].

**DING, W. ; MARCHIONINI, G. ; TSE, T.**, « Previewing video data: browsing key frames at high rates using a video slide show interface », *Proceedings of the International Symposium on Research, Development, and Practice in Digital Libraries*, Tsukuba (Japon), Nov 18-21, 1997, 151-158.

URL : <http://www.DL.ulis.ac.jp/ISDL97/proceedings/weid/weid.html> [consulté le 12 mai 2001].

**Dublin Core**, URL : <http://dublincore.org/>  
[consulté le 11 mai 2001].

**Getty Research Institute**,

URL : [http://www.getty.edu/research/institute/standards/intrometadata/3\\_crosswalks/index.html](http://www.getty.edu/research/institute/standards/intrometadata/3_crosswalks/index.html)  
[consulté le 11 mai 2001].

**HUDON, Michèle ; TURNER, James M. ; DEVIN, Yves**, « Description et indexation des collections d'images en mouvement : résultats d'une enquête, 2001 ». Document soumis pour publication mais pas encore accepté.

**ISAN** (International Standard Audiovisual Description),

URL : <http://www.nlc-bnc.ca/iso/tc46sc9/isan-f.htm>  
[consulté le 11 mai 2001].

**MPEG7**,

URL : <http://www.darmstadt.gmd.de/mobile/MPEG7/>  
[consulté le 11 mai 2001].

**pmeta**,

URL : <http://www.nla.gov.au/preserve/pmeta.html>  
[consulté le 11 mai 2001].

**RDF** (Resource Description Framework),

URL : <http://www.w3.org/RDF/>  
[consulté le 11 mai 2001].

**TEASDALE, Guy**, traducteur, « Guide d'utilisation du Dublin Core ». URL : <http://www.bibl.ulaval.ca/DublinCore/usageguide-20000716fr.htm>  
[consulté le 11 mai 2001].

**TURNER, James M.**, Determining the Subject Content of Still and Moving Image Documents for Storage and Retrieval: an Experimental Investigation, PhD thesis, University of Toronto, 1994.

**TURNER, James M.**, « Comparing user-assigned terms with indexer-assigned terms for storage and retrieval of moving images: research results », *Proceedings of the 58th ASIS Annual Meeting*. Chicago, Illinois, October 9-12, 1995, vol. 32, 9-12.

**TURNER, James M. ; HUDON, Michèle ; DEVIN, Yves**, « Text as a tool for organizing moving image collections », *ACSI 2000 : Les dimensions d'une science de l'information globale. Association canadienne des sciences de l'information, Travaux du 28<sup>e</sup> congrès annuel, Edmonton, 2000*.

URL : <http://www.slis.ualberta.ca/cais2000/turner.htm>  
[consulté le 15 mai 2001].

**UPF** (Universal Preservation Format),

URL : <http://info.wgbh.org/upf/>  
[consulté le 11 mai 2001].

**VRA Core**, URL : <http://www.gsd.harvard.edu/~staffaw3/vra/vracore3.htm> [consulté le 11 mai 2001].

**W3C** (World Wide Web Consortium),

URL : <http://www.w3.org/> [consulté le 11 mai 2001].