



**HAL**  
open science

# L'ICONOTHEQUE D'ART AUJOURD'HUI ET DEMAIN.

Henri Hudrisier

► **To cite this version:**

Henri Hudrisier. L'ICONOTHEQUE D'ART AUJOURD'HUI ET DEMAIN.. 3ème 2000 Workshop on Multimedia Art, Mediateca Fundatio La Caxia, Jan 2000, Barcelone, Espagne. hal-02594556

**HAL Id: hal-02594556**

**<https://hal.science/hal-02594556>**

Submitted on 15 May 2020

**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.

**L'ICONOTHEQUE D'ART AUJOURD'HUI ET DEMAIN.**  
**Par Henri Hudrisier, Maître de Conférences à l'Université de Paris 8, Laboratoire**  
**Observatoire International des Multimédias**  
**III 2000 Workshop on Multimedia Art**  
**21<sup>st</sup>, 22<sup>nd</sup> January 2000**  
**Mediateca, Fundatio « la Caxia », BARCELONA**

Car chaque image, à chaque coup, vous force à réviser tout l'univers.  
Aragon<sup>1</sup>

...l'art pur suppose la coïncidence visible de l'esprit du contenu  
avec l'expression des éléments de forme et celle de l'organisme formel.  
Paul Klee<sup>2</sup>

Les documents constitués d'images animées sonorisées (cinéma, vidéo), les images fixes (non seulement les photographies, mais les images d'art : tableaux, gravures, images numériques) et les sons (musique, paroles ou bruits) posent depuis longtemps des problèmes théoriques et pratiques complexes. Les historiens d'art, les sémiologues, les chercheurs de très nombreuses disciplines qui travaillent sur des sources audiovisuelles mais aussi les artistes eux mêmes ainsi bien sûr que les professionnels de la recherche audiovisuelle et de la conservation (médiathécaires, conservateurs de musées), tous, selon des axes d'approche nécessairement distincts sont soucieux de conserver, organiser, donner à voir et entendre ce foisonnement de documents tellement complexe dans son hétérogénéité tant de formes, que d'origine et de significations.

Face à une telle disparité de la matière audiovisuelle et de ses modes d'approches, les technologies de l'information et de la communication, qui peuvent répondre aux problèmes posés, reproduisent nécessairement en écho cette même complexité. Si elles ne le faisaient pas nous pourrions penser, bien à raison, qu'elles n'auraient pas d'intérêt.

Il est de ce fait fondamental que nous accompagnions la sophistication des techniques numériques de production du document et celle de la médiation patrimoniale. Cela nous demandera des efforts importants d'appropriation culturelle, mais cela augmentera considérablement nos possibilités de compréhension, de conservation, de modes d'accès et bien sûr de diffusion des fonds audiovisuels. C'est d'ailleurs la position qu'ont toujours pris les conservateurs de musées, les cinémathécaires, les phothotécaires et autres médiathécaires d'images qui ont toujours été obligés de s'inscrire dans une toute autre techno-culture que leurs collègues des documents textuels. Tout les y portait naturellement tant du côté des références catalographiques audiovisuelles qui par la nature beaucoup plus collective des auteurs et éditeurs que celle du texte édité posent d'inextricables problèmes spécifiques, que du côté de l'analyse de par l'évidente spécificité sémiotique de l'audiovisuel.

Le monde de la médiation patrimoniale de l'art s'est non seulement interrogé mais il s'est aussi équipé dans les 15 dernières années avec la première génération d'outils techniques qu'ont produit les industriels de l'audiovisuel, de l'informatique et des réseaux. Les acteurs du patrimoine ont pu ainsi déployer des systèmes de visualisations interactives évidemment liés à leur description et à leur catalogue informatique.

Mais la convergence des technologies (je ne dis pas des techniques<sup>3</sup>), pénètre maintenant en profondeur les codes et langages, ce qui génère une deuxième génération d'outils qui concernent bien

---

<sup>1</sup> Louis Aragon, Le paysan de Paris, Gallimard

<sup>2</sup> Le crédo du créateur, textes rassemblés in Paul Klee, Théorie de l'art moderne, éditions Gonthier, collection Bibliothèque Médiations, Genève, 1964. p34

sûr au premier chef les professionnels de l'audiovisuel mais aussi les médiateurs du patrimoine audiovisuel ainsi que les artistes<sup>4</sup>. Pour les médiateurs patrimoniaux cela va transformer en profondeur leurs métiers et leurs méthodes et les obliger à un renouvellement sans doute beaucoup plus fondamental que la première génération d'outils techniques de l'information et de la communication dont ils se sont déjà largement équipés.

La première génération de NTIC (Nouvelles Technologies de l'Information et de la Communication) qui a pénétré le monde des musées et médiathèques était essentiellement constituée d'instruments de saisie et de numérisation, de régie de stockage, de réseaux plus ou moins adaptés à la transmission d'images (ou de sons), par contre la deuxième génération est (ou plutôt sera parce que la plupart de ces technologies sont actuellement en cours de recherche-développement) constituée de langages et de codes qui vont pénétrer en profondeur la structure et la sémantique de l'audiovisuel lui-même. Le maître mot en sera MPEG<sup>5</sup> (surtout ses couches MPEG4, 7 et 21. Nous y reviendrons.)

Il est à craindre que le milieu de la médiation patrimoniale de l'audiovisuel qui, il y a vingt ans, attendait impatiemment les techniques qui allaient permettre de lier interactivement l'image à la référence et la glose à l'image, risque par contre de se montrer beaucoup plus rétif à des techniques qui pénétreraient plus en profondeur la forme même de l'analyse ou de la glose, qui renouvellerait radicalement la structure, le repérage référentiel et sémantique des grands corpus de documents qui constituent nos collections.

La deuxième génération des NTIC qui s'amorce devrait pouvoir répondre à un grand nombre de questions, que la complexité des oeuvres dont nous avons la charge laisse encore sans réponses. Il est à craindre aussi que ceux-là même qui refuseront de chercher à repenser leurs méthodes à la lumière des NTIC, refuseront plus vigoureusement encore de chercher à s'associer internationalement pour créer des ressources communes partagées et mettre en commun des méthodologies.

Quelle est la nature de cette deuxième vague de NTIC ? La vague de modernisation qui va suivre va-t-elle exiger encore plus d'équipements techniques matériels ou au contraire induire un redéploiement intelligent des capacités internes (et coopératives) de nos établissements ?

C'est pourtant tout le sens de la deuxième génération des NTIC. Celle-ci exigera proportionnellement une bien plus faible part d'équipements techniques matériels que la première génération, mais cela induira par contre un redéploiement important des professionnels de la médiation patrimoniale de l'audiovisuel. Cela devra se faire à la fois en interne dans chaque institution mais aussi collectivement pour construire de façon internationale les centres de ressources communs propres à rendre compatibles nos différentes approches, tout en les laissant plurielles.

Cela ne se fera pas sans d'importants efforts de dissémination et de divulgation : c'est la démarche qu'a déjà depuis longtemps entreprise la Fondation de la Caxia à travers notamment les Rencontres internationales qu'organise la médiathèque.

---

<sup>3</sup> Sur ce point on pourra lire Henri Hudrisier, Distinguer technique et technologie, un préalable à la réflexion sur les NTIC, in « Passerelles » (Revue de l'Université de Paris 8) n°24, Dossier spécial NTIC, Mars 1999.

Même auteur dans la même livraison :

- La « lecture assistée par ordinateur » et ses applications savantes ou pédagogiques dans un contexte interactif normalisé : la TEI (Text Encoding Initiative)

- Des abacistes aux algoristes : les risques potentiels d'un hypermédia mal compris.

« Passerelles » est disponible sur le serveur de l'Université de Paris 8 : [www.univ-paris8.fr](http://www.univ-paris8.fr)

<sup>4</sup> Pour une certaine avant-garde de l'art, cela peut avoir des effets non négligeables sur la définition de l'art lui-même. La production d'images et de sons directement en numérique, la multiplication d'œuvres d'art qui peuvent maintenant se dématérialiser au point de devenir des documents virtuels, des « programmations d'événements ou de fabrication en temps réel », autant de modernités inédites qui non seulement interrogent l'art lui-même mais posent de difficiles et quelquefois insolubles problèmes aux médiateurs du patrimoine notamment pour ce qui concerne la conservation de ces nouveaux objets d'art, on pourra consulter la thèse de Corinne Welger-Barboza: Le devenir documentaire du patrimoine artistique, perméabilité du musée aux technologies numériques, 1998, Université de Dijon (France)

<sup>5</sup> Moving Pictures Experts Groups

J'hésite beaucoup devant un public comme le votre à développer un raisonnement qui donnera l'impression de sortir du sujet et de s'éloigner de l'image. Ceux qui me connaissent savent pourtant que mon parcours professionnel ou universitaire n'a jamais abandonné l'image. Depuis ces dernières années, hormis mon enseignement à l'Université de Paris 8 qui est largement orienté sur la documentation audiovisuelle, j'ai participé à de nombreuses recherches et groupes de travail ayant l'audiovisuel pour sujet. J'en cite trois qui représentent une part importante du temps que j'ai pu consacrer à la recherche en matière de documentation audiovisuelle :

- la Bibliothèque nationale de France (BnF), où j'ai participé à deux études d'organisation de l'audiovisuel
- l'Institut national de l'audiovisuel (INA) où j'ai animé pendant trois ans un des quatre ateliers de préfiguration de l'Inathèque (atelier « les fonds »). Toujours à l'INA, je participe actuellement au Collège iconique
- à l'AFNOR (Association Française de Normalisation) pour ma participation à des réunions du groupe français MPEG 7.

Cependant pour comprendre le devenir d'une bibliothéconomie de l'image, il me semble de plus en plus indispensable de s'intéresser aussi au devenir du document général et à celui du texte notamment, d'où ce nécessaire détour.

C'est là un effet de la convergence des médias. Mais que les passionnés d'audiovisuel se rassurent, si l'information structurée, balisée, normalisée, mise en bibliothèque virtuelle qui se construit aujourd'hui, s'applique encore beaucoup au texte, par contre la télévision interactive qui s'élabore expérimentalement (MPEG4) en même temps que se dessine sa documentation structurée (MPEG 7) et, au delà son échange commercial en réseaux (MPEG 21), aura d'ici 2 ou 3 ans un pouvoir fédérateur et intégrateur sur la totalité des langages de l'information à l'orée des années 2005, 2010. Cette dynamique d'intégration autour de l'information audiovisuelle numérisée (normalisée sur la norme MPEG) touchera bien sûr l'image filmique, télévisuelle, le multimédia donc le son et l'image fixe qui en sont évidemment des sous-composantes. Le texte qui en est aussi une sous-composante a ainsi de fortes raisons d'être normalisé comme document répondant aux mêmes critères de description. En fait, dès les premières années de ce millénaire, nous ne devons plus raisonner, pour ce qui est de la gestion globale de l'information, en terme d'images et de sons distincts des textes, mais définitivement en terme de documents.

Les vidéodisques et aujourd'hui les CDrom, les DVD et autres mémoires numériques d'image associées aux moyens de diffusion planétaire sur l'Internet ont représenté une première révolution de méthodes : celle de l'hypermédia. Elle fut aussi institutrice de nouveaux modes de médiation patrimoniale ou d'histoire de l'art que l'avènement de la gravure permettant au discours sur l'art d'exister, de se théoriser et de prendre son envol, avait pu l'être en son temps.

Ainsi avait-on vu César Ripa proposer un dictionnaire iconologique en 1600 au moment<sup>6</sup> où la gravure sur bois associant facilement avec le livre imprimé et David Teniers entreprendre un demi siècle plus tard une démarche iconographique sur une large échelle : le *Theatrum Picturum*<sup>7</sup> s'appuyant sur la gravure sur cuivre<sup>8</sup>.

Aujourd'hui on ne compte plus les sites Internet des musées. La plupart d'entre eux éditent des CDrom. L'édition d'art multimédia est de plus en plus répandue et les catalogues des disquaires voient leur part patrimoniale augmenter considérablement : il existe un véritable engouement populaire pour le patrimoine filmographique, maintenant repris sur DVD, ou le patrimoine sonore. Les banques de données audiovisuelles aussi, qu'elles soient ou non accessibles sur Internet, sur un réseau spécialisé ou sur CDrom ont connu une extension sans précédent, mais dans la plupart des cas

---

<sup>6</sup> Cesare Ripa, *Iconologia*, première édition non illustrée Rome en 1593, la première édition illustrée date de 1603 à Milan

<sup>7</sup> Un ensemble de plus de 240 images organisées selon la thématique des oeuvres, publié à Anvers en 1658

<sup>8</sup> J'ai développé il y a presque 20 ans ce parallèle dans : Henri Hudrisier, *De l'estampe au vidéodisque : les nouvelles technologies de l'iconothèque*, in « Image et signification », Rencontres de l'Ecole du Louvre, ed. La Documentation française, Paris, 1983, p 87 à 99

on se contente encore d'associer des techniques de descriptions traditionnelles d'image à des systèmes plus ou moins sophistiqués et rapides de présentation d'images.

Ce foisonnement de visualisations et d'écoutes interactives, disponible sur toute la planète, en temps réel, potentiellement lié à la complexité de leurs références et susceptible d'être interrogé sur une infinité de systèmes d'interrogation, est ce dont nous disposons aujourd'hui. D'évidence les services rendus sont immenses et ils étaient inespérés (et même hors d'atteinte des rêves les plus fous) il y a une vingtaine d'années.

## **L'INFORMATION STRUCTUREE, BALISEE ET NORMALISEE : LA DEUXIEME GENERATION DES NTIC**

La deuxième génération des NTIC devrait introduire une révolution de méthode alors que d'évidence nous ne nous sommes pas encore appropriés les conséquences de la première. Cependant selon un processus classique de l'évolution des techniques, les générations technologiques successives ont pour vertu d'intégrer et de simplifier les dispositifs techniques des générations précédentes. En terme d'appropriation prospective, il sera donc opportun d'accorder effectivement un peu moins d'importance aux dispositifs actuels de l'iconothèque hypermédia d'aujourd'hui et d'essayer de les imaginer intégrés dans la 2<sup>ème</sup> génération.

Cette deuxième génération répond à quelques qualités clefs :

- langages à balises de l'information structurée
- intercompatibilité, prédominance des normes ouvertes sur les standards
- globalisation planétaire de l'information structurée (mise en bibliothèque virtuelle)
- convergence des média, des modes, des métiers, des disciplines, des langues...

Le multimédia et l'hypertexte restent des notions fortes mais on peut les considérer comme une vision encore partielle de la convergence.

### **Les langages à balises et l'information structurée**

Aujourd'hui le monde entier constate le fonctionnement d'Internet et les formidables performances des moteurs de recherche automatique comme Yahoo, Altavista<sup>9</sup>.

Il faut savoir que ces moteurs de recherche automatique d'informations fonctionnent parce que les documents mis sur les différents serveurs Web de la planète ne sont pas seulement numérisés et accessibles en réseaux, mais aussi parce qu'ils sont normalisés et organisés sous la forme d'un unique langage balisé (HTML = HyperText Mark up Language, en français langage balisé hypertextuel).

En informatique, comme sur mer, sur route ou sur un aéroport, une balise signale et/ou borne une zone sensible ou un obstacle. Les balises d'un document permettent d'interagir (par exemple chercher des mots clefs, opérer des renvois de notes ou d'index, réserver une zone (icône d'appel) d'où il sera possible par choix interactif de visualiser, une image ou faire écouter un commentaire sonore, etc....

Comme spécialistes de l'audiovisuel nous connaissons bien les documents électroniques balisés et le vidéodisque en fut un premier exemple d'autant plus intéressant qu'il démontre qu'il n'est pas obligatoire qu'un document balisé soit intégralement numérique. Ainsi les premiers vidéodisques laser étaient analogiques pour ce qui est du son et de la vidéo et seul le système de balisage était numérique. La totalité des disques de la famille CD sont évidemment construits selon la logique des langages balisés.

Les balises que l'on retrouve en SGML, en VRML<sup>10</sup> et maintenant en XML ont pour effet de catégoriser les différentes sections ou points critiques d'un document selon une structure logique explicite.

---

<sup>9</sup> Pour ce qui est de l'audiovisuel on voit apparaître des « moteurs images » de façon non obligatoirement convaincante dans un certain nombre de sites Internet expérimentaux (notamment des sites proposant des fonds muséaux). Ce sont des banques d'images dans lesquelles les moteurs de recherche n'utilisent plus la description textuelle d'images mais activent des requêtes en exploitant uniquement les caractéristiques graphiques ou visio-sémantiques de l'audiovisuel. Il est finalement dommage que ces outils expérimentaux connaissent, du fait de la pénétration systématiquement planétaire d'Internet, une telle publicité, alors qu'ils ne se justifieraient véritablement qu'intégrés dans un environnement beaucoup plus large d'informations systématiquement structurées, nous y reviendrons.

<sup>10</sup> SGML (Standardised General Mark up Language) et VRML(Virtual Reality Mark up Language)

Pour un éditeur, par exemple le balisage SGML d'un livre consistera, non plus comme autrefois à penser le livre comme une succession de pages fabriquées selon la technologie de la composition au plomb, mais à le découper selon une structure de type conceptuel, celle du « front, body et back », dont il ne reste plus qu'à déclarer le "style de mise en page" à travers une feuille de style : DTD (Definition Type Document) en SGML.

L'idée que les textes doivent être balisés est aussi vieille que l'écriture, mais la codification de normes de balisage dépend à la fois des besoins culturels et de ce qui est technologiquement perçu et réalisable.

L'espacement entre les mots<sup>11</sup> est un des premiers besoins de balisage qui soit apparu dans l'histoire de l'écriture.

La mise en évidence de la structure logique et grammaticale de l'énoncé et des phrases qui le composent, vient plus tard. Elle entraîne, au sein de l'écriture, la codification d'autres balises : les signes de ponctuation.

Un autre niveau de balise plus macro-textuelle peut se déployer lorsqu'un livre (c'est à dire un texte dont la caractéristique principale est la grande longueur) n'est plus transcrit sur un rouleau (volumen) mais sur un cahier organisé en pages (codex).

En effet, celui qui ouvre une Bible ou un Coran pour la première fois est en droit de s'étonner de la fréquence des marques de repérage, livres, versets, sourates, qui reviennent une dizaine de fois par page. Ils sont la trace héraldique de la difficulté de se repérer dans un volumen dont le texte déroulé en séquences (les latins disaient "explicare", littéralement défaire les plis) ne bénéficiait pas des commodités d'ergonomie de lecture<sup>12</sup> que représentent les pages numérotées et résumées dans une table des matières qui seront rendues possibles dans le codex.

Cette évolution du livre<sup>13</sup>, l'émergence du codex se substituant aux anciens "volumen" débute avec notre ère (début du premier siècle) et elle est ressentie par le monde traditionnel des lettres (poètes et écrivains latins) avec la même méfiance face à la nouveauté technologique que le monde littéraire d'aujourd'hui perçoit le CDrom et Internet. Les intellectuels du siècle d'Auguste pensaient que le codex était bon pour les livres pratiques, scientifiques et encyclopédiques mais que les oeuvres littéraires devaient continuer à être produites sur rouleaux<sup>14</sup>.

La similitude entre le volumen et l'audiovisuel en bobine et les conséquences sur l'ergonomie archivistique du cinéma sont frappantes. D'après mes propres évaluations à la Cinémathèque Gaumont, dans la fin des années 70, j'évaluais à une semaine de travail pour une équipe de 10 personnes la préparation du matériel filmique<sup>15</sup> et son visionnement pour aboutir in fine à ce qui permettait de monter effectivement 30 mn d'émission à base d'archives filmiques. Grâce aux recopies vidéo d'une partie des fonds, qui se généralisent dans la plupart des cinémathèques d'archives<sup>16</sup>, ce

---

<sup>11</sup> On peut remarquer que les écritures idéographiques n'éprouvent pas (et pour cause) ce besoin de balisage entre les mots.

<sup>12</sup> L'Evangile de Luc nous relate la "posture" du Christ lisant à la synagogue "on lui remit le livre du Prophète Isaïe et déroulant le livre, il [le Christ] trouva le passage où il était écrit [...] Il replia le livre, le rendit au servent et s'assit. (Luc 4- 17 à 20) "

<sup>13</sup> La subtilité du vocabulaire prend ici son importance. Un livre, c'est à dire une oeuvre généralement textuelle, d'une longueur conséquente, peut être indifféremment un volumen ou un codex. Il serait évidemment fautif d'assimiler obligatoirement le livre au codex. Le volumen mais aussi le livre sur CDrom, sur cassette audio, ou sur réseaux sont indubitablement des livres.

<sup>14</sup> CHARTIER (Roger), Du codex à l'écran : les trajectoires de l'écrit in : Pour une nouvelle économie du savoir, sous la direction de Ghislaine Chartron, Sylvie Fayet-Scribe, Brigitte Guyot et Jean-Max Noyer, Presses Universitaires de Rennes, Rennes, 1994 pp 65 à 75

<sup>15</sup> Chercher des références sur le fichier papier, aller les chercher en cinémathèque de sécurité -films flammes - préparation du visionnement

<sup>16</sup> Dans les très grandes cinémathèques (plusieurs dizaines de milliers d'heures de films) pour des raisons économiques, mais aussi archivistiques évidentes, une partie seulement du fonds est généralement recopiée en vidéo. Par exemple à la Cinémathèque d'Actualité Gaumont, à Paris, seuls les « Journaux hebdomadaires d'actualité » sont exhaustivement disponibles en vidéo dans les locaux même du siège de la Cinémathèque. Nous ne disposons ainsi que d'un dixième environ du fonds. La quasi totalité que représentent les chutes et les documents « non utilisés » dans les Journaux hebdomadaires d'actualité n'est souvent accessible que sur les documents originaux demeurés le plus souvent encore à l'état de « film flamme ». Il va de soi que recopier ces documents originaux en vidéo sans disposer des crédits pour en faire des recopies de sécurité revient à vouloir volontairement détruire le fonds.

temps s'est énormément raccourci. Cependant 10 heures<sup>17</sup> de film d'archives pourraient être consultées (pour ce qui est de leur aspect visuel) à travers une extraction automatique d'une image numérique par plan (imassettes, telle que cela se pratique à la BnF, sur la SLAV de l'Inathèque ou dans de nombreuses autres médiathèques dans le monde. On peut alors évaluer que 10 heures d'archives à raison d'une imasette par plan représente environ 10000 imassettes. Si on a pour finalité de sélectionner ce que l'on voudra effectivement visionner au rythme cinématographique on peut consulter et sélectionner les 10000 imassettes confortablement en 1 ou 2 heures (et ce par une seule personne). De plus ces imassettes servent de descriptif et de memo et pourront ensuite être réutilisées pour structurer et organiser la sélection finale du matériel par le réalisateur, qui pourra même servir au monteur.

Aujourd'hui, l'information structurée et les langages à balises se sont généralisés et normalisés sous l'unique famille SGML (essentiellement sous l'influence du Web et de HTML). HTML est en effet un cas particulier de la norme SGML acceptant de surcroît des balisages hypertextuels et multimédia (i.e. des liens hypertextuels, des images, des sons). Si le Web est devenu si facilement populaire, ce fut grâce à la relative simplicité de HTML (qui n'est pas un langage rigoureusement structuré) ; c'est aussi cette simplicité qui constitue aujourd'hui les limites du Web.

Cette question concerne particulièrement la documentation parce que si le Web HTML permet de lier « hypertextuellement » des pages Internet à d'autres pages par contre, sauf similitude absolue de leur système documentaire, deux banques de données ne peuvent pas être cumulées dans une même consultation. C'est l'avantage absolument primordial qui va être rendu possible (et certainement assez vite se généraliser<sup>18</sup>) dans la deuxième génération NTIC actuellement en émergence (fondée elle sur le langage à balises XML) et dont on comprend qu'elle sera celle d'un fantastique essor de la documentation<sup>19</sup> en réseaux et non plus les interminables « promenades » quelquefois peu pertinentes à travers le Web.

### **Intercompatibilité, prédominance des normes ouvertes sur les standards :**

Nous avons déjà une culture de l'intercompatibilité puisque nous n'hésiterions pas à traverser l'Atlantique avec pour tout bagage professionnel, non plus un cartable rempli de textes, de dossiers de diapositives, mais une ou plusieurs disquettes, certains que nous sommes d'être capables de les ouvrir sur un logiciel compatible. Nous pouvons aussi avoir tout simplement expédié les fichiers en question par Internet dans l'université ou l'institution qui nous accueille. Plus désinvolte (mais nous courrons alors le risque d'une panne du réseau), nous pouvons penser qu'il nous sera possible d'aller, au dernier moment, chercher nos fichiers à distance sur notre ordinateur resté branché dans notre bureau. Cependant cette compatibilité logicielle est essentiellement le fruit du quasi-monopole Microsoft et de la large diffusion de ses logiciels standardisée et non normalisée. Cela en fait une réalité fragile puisqu'une faillite (peu probable dans le cas de Microsoft) ou une cessation d'activité pourrait réduire à néant des quantités importantes de données.

Il en va donc de notre survie culturelle que nous sachions (cela s'impose) distinguer les standards (qui sont des logiques volontairement fermées par les constructeurs), des normes (auxquelles peuvent correspondre d'ailleurs des standards). Le standard est la volonté d'un ou de plusieurs industriels, la norme, elle, est le fruit d'un consensus issu de la consultation animée par un organisme de normalisation (national ou international). Cette consultation, qui a pour but d'aboutir à un consensus constructif, est menée de façon conjointe avec les industriels, les états éclairés par les utilisateurs et des experts.

---

<sup>17</sup> Cette durée est calculée sur la base du ratio de 1 pour 20, ce qui signifie que pour finaliser 1 heure de film d'archives montée on peut compter qu'il faut visionner (après sélection des références sur catalogue) environ 20 fois le métrage final. (Emissions, Le Grand Album de la Gaumont, Production de Jacques Martin et Daisy de Galard. Réalisation Pierre Philippe, Antenne 2, Paris 1978, Série hebdomadaire de 25 émissions.

<sup>18</sup> RDF (Ressources Document Format) sera le langage spécialisé (une application XML) qui permettra cette mise en Internet cohérente des différents sites documentaires.

<sup>19</sup> Ce sera surtout l'essor sans précédent du commerce électronique, étant entendu que l'on pourra cumuler dans une même requête des centaines et des milliers de catalogues commerciaux et industriels pour y trouver au meilleur prix et selon les caractéristiques techniques ou d'usage que l'on souhaite, le produit ou le service que l'on recherche. Cet aspect commercial, pour ce qui est de l'audiovisuel sera l'objet de la couche MPEG 21 (nous y reviendrons).

En matière d'information structurée et de langages balisés, le paysage s'est considérablement éclairci et nous disposons pratiquement d'une seule norme SGML (et des normes filles qui en sont issues HTML, VRML, XML). MPEG 4 et 7 sont aussi conçues à partir de XML mais ce sont des applications

Le fait que nous disposions aujourd'hui de normes mondiales d'échange normalisé d'informations est une sorte d'accident bénéfique de l'histoire et un sous-produit des besoins militaro-industriels américains. L'armée américaine qui est aussi le plus gros client du monde et qui peut donc facilement imposer ses désirs et ses besoins avait (il y a une quinzaine d'années) de grandes difficultés pour organiser en un tout cohérent les informations diverses qui étaient fournies sous des formats certes numériques mais totalement disparates. Les militaires devaient pouvoir disposer notamment de documentation aéronautique, de documentation d'armement, logiquement structurées disponibles sous le même format. C'est pour répondre à ce besoin que l'ARPA (Advanced Research Projects Agency) proposa que le Pentagone impose à tous les fournisseurs le format SGML<sup>20</sup>. Il faut voir les retombées en cascade de cette décision sur la plupart des fournisseurs des armées dans de nombreux pays au monde qui sont liés par des accords de coopération, OTAN (NATO) Alliance Est-Pacifique, alliances des pays occidentaux avec leurs ex-colonies. SGML est devenu de ce fait la norme de base de toutes les informations structurées normalisées actuellement mises en oeuvre. Cette décision historique a depuis produit des effets culturels importants.

Ainsi est né une véritable technoculture du document structuré non seulement ésotérique et réservée à ceux qui ont besoin de structurer des données à très haute valeur ajoutée (voire des données à valeur stratégique), mais une techno-culture qui chez les professionnels de l'édition et de la technologie des logiciels de traitement de texte commençait à devenir une technologie assez largement répandue. Cependant, même en ayant été l'objet d'une normalisation mondiale par l'ISO, SGML serait sans doute restée une culture professionnelle connue seulement des éditeurs de catalogues industriels ainsi que des éditeurs de données complexes comme les annuaires, les dictionnaires etc. Le fait que cette culture professionnelle se soit répandue comme un savoir beaucoup plus largement diffusé dans la technoculture est dû à l'initiative des chercheurs physiciens du CERN à Genève lorsqu'ils ont créé l'Internet Web.

Le Web<sup>21</sup> est la première manifestation à l'échelle planétaire d'un système d'informations structurées, mais il est volontairement conçu sous une forme dégradée et simplifiée, ce qui a permis son succès<sup>22</sup>, mais ce qui en trace aussi les limites notamment lorsqu'il s'agit de documentation, de références, de bases de ressources.

Nous connaissons tous l'Internet et sa fascinante aptitude d'interactivité hypertextuelle mondiale. On touche aussi très vite aux limites d'un tel système qui se contente de tirer des liens logiques entre toutes les informations rassemblées sur la planète, ce qui rend d'énormes services mais finit par lasser de par sa monotonie d'intelligence fonctionnant uniquement sur le rappel des références.

Au delà de cette première génération, le Web se développe maintenant, sur une aussi large échelle, une deuxième génération non plus fondée sur HTML mais sur XML, i.e. un nouveau langage à balises<sup>23</sup> lui aussi issu de SGML (qui le définit presque complètement) permettant la mise en place d'informations intégralement, systématiquement (canoniquement pourrait-on dire) structurées. Cette évolution technologique d'apparence anodine représente en soi une révolution de mon point de vue aussi importante que l'ensemble des innovations NTIC depuis l'apparition du vidéodisque et la généralisation de la culture de l'interactivité et de l'hypermédia.

Adopter des technologies d'information structurée et normalisée de la deuxième génération répond à un besoin évident. Maintenant que la plupart des musées, des médiathèques, des iconothèques de

---

<sup>20</sup> Ce qui à l'époque réduisit à néant une autre norme européenne ODA.

<sup>21</sup> ... et non pas les formes anciennes de l'Internet qui n'étaient qu'une généralisation à l'échelle de la planète de simple logique d'échanges de messagerie électronique

<sup>22</sup> Les chercheurs du Cern qui ont conçu le Web ont volontairement simplifié le dispositif d'information structurée pour qu'un large public d'utilisateurs puisse se l'approprier sans trop d'obstacles techniques.

<sup>23</sup> XML apporte comme innovation principale, l'organisation rationnelle des caractéristiques du langage SGML qui est organisée en sous ensembles (par exemple l'espace logique des mises en pages, l'espace du réseau...) et qui est associée judicieusement à des outils logiciels externes qui organisent ainsi des applications.



tout ordre ont informatisé presque systématiquement leurs données catalographiques et amorcé dans une large mesure (soit directement, soit indirectement<sup>24</sup>) la numérisation de leurs images et de leurs sons, on comprend aisément que l'étape qui doit suivre consiste à pouvoir lier ensemble la diversité de ces ressources, sous peine, ce qui est déjà en partie le cas de se décourager devant l'extrême disparité des différents types d'approches, des diversités des méthodes, de l'hétérogénéité des formats disponibles sur Internet ou dans des produits numériques édités.

Ce type de déclaration de principe consistant à vouloir disposer de systèmes intercompatibles, structurés, susceptibles de se lier entre eux, suscite toujours une crainte légitime : que deviendra la nécessaire disparité des approches (qu'ils s'agisse des types d'images : tableaux, gravures, photographies ; des modes de médiations : films, vidéos, images fixes, sons... ; des typologies d'usages ou d'institutions : agences de presses, musées, bibliothèques publiques, cinémathèques ou photothèques d'archives, collections de Dépôt Légal...), qu'en sera-t-il des disparités linguistiques, juridiques, des approches nécessairement disparates pour telle ou telle discipline ou tel ou tel usage professionnel.

Il est important de souligner avant tout la particularité de l'information structurée et des normes de langages informatiques qui permettent de mettre en œuvre cette mise en structure (SGML, XML, MPEG 4 et 7...). Ce sont des normes ouvertes et des langages susceptibles de s'articuler (donc de générer un nombre infini de codes, donc de significations).

Cette pétition de principe est nécessaire avant tout développement traitant de ces questions, sous peine, ce qui est malgré tout trop souvent le cas, de se heurter à des refus d'entendre. La normalisation fait peur, à raison, dans un monde où trop de choses se normalisent brutalement : la nourriture, les hôtels, les modes de vie, les médias et leurs contenus...

C'est pourquoi il est important d'insister sur le fait que les langues (fait culturel par excellence) sont fondées sur une norme, un arbitraire du code qui présente la particularité d'être doublement articulée : ce qui lui permet d'être une norme ouverte. De nombreux autres codes de langage non nécessairement liés à la langue, comme la musique, la chorégraphie, la langue des signes qu'utilisent les sourds, mais aussi le cinéma, la peinture ou la photographie sont des codes que nous pouvons qualifier de codes articulés de façon plurielle<sup>25</sup> : tous sont donc des normes ouvertes.

### **La globalisation planétaire de l'information structurée et sa mise en bibliothèque virtuelle : la « machine grammatologique »**

La globalisation planétaire de l'information structurée est évidemment un enjeu très complexe parce qu'il exige que notre « réalité sociale<sup>26</sup> » fasse sienne des notions non encore éclaircies, même dans l'esprit des spécialistes. C'est donc une avancée qui doit s'appuyer sur le développement de plusieurs logiques d'usage parallèle, sur plusieurs utopies techno-culturelles.

Je citerai ainsi :

- Les grandes bibliothèques virtuelles en réseaux et la lecture savante
- Le catalogue industriel et commercial global (que je cite pour mémoire)
- Les grandes vidéothèques et cinémathèques en ligne

En d'autre texte j'ai proposé que ce nouvel environnement technoculturel puisse s'appeler la « machine grammatologique »,

### **Les grandes bibliothèques virtuelles en réseaux et la lecture savante**

La notion de bibliothèque virtuelle est une dimension importante de l'avancée des technologies de l'information structurée.

Aujourd'hui le savoir littéraire, l'histoire ou la critique d'art, l'analyse du cinéma, la sociologie, la philosophie... peut être l'objet d'une accumulation collective normalisée.

---

<sup>24</sup> J'entends par là que la numérisation d'image a pu être faite pour des besoins internes et documentaires ou au contraire que cette saisie numérique s'est inscrite dans une politique éditoriale externe.

<sup>25</sup> Une telle affirmation peut être développée, évidemment l'image photographique ne s'articule pas selon les mêmes principes que la langue, mais cela sort du sujet directement traité. Il est par contre fondamental de savoir qu'une norme comme MPEG dans ses couches 4 et 7, qui est à la base des systèmes d'information structurée susceptibles d'organiser des grands corpus d'audiovisuel, est conçue selon une logique répondant à une approche en terme d'articulation.

<sup>26</sup> John R. Searle, La construction de la réalité sociale, Gallimard, Paris, 1998

De grandes bibliothèques virtuelles et numériques (grands catalogues collectifs, projet Gutenberg, Bibles, patristiques et concordances, Thesaurus Lingua Graeque, base Frantex...) disponibles à la fois en réseaux et sur CDrom, constituent à la fois de nouvelles bases patrimoniales et de nouveaux enjeux.

A l'évidence ces différents patrimoines exigent des solutions que la technique hypertextuelle, et hypermédia, et la facilité participative des réseaux et des supports rendent possibles mais que l'information structurée rend encore plus pertinente.

Au Moyen-âge celui qui possédait personnellement des livres les couvraient de notes personnelles, et de gloses. Le livre médiéval annoté, est ainsi à lui seul comparable à un petit secteur de bibliothèque et de références bibliographiques liées.

Les NTIC autorisent précisément les utilisateurs d'une bibliothèque virtuelle (en site comme en réseaux) à réactiver les comportements de l'intellectuel médiéval couvrant son livre de notes, de gloses, de références et de citations. À dire vrai ce comportement n'a jamais cessé d'être pour tous ceux qui ne se sentaient pas distants de la production de textes imprimés, les écrivains et les savants, qui comme beaucoup d'intellectuels n'ont jamais cessé d'annoter leurs livres personnels.

### **Les stations de lecture assistée par ordinateur**

Pour ceux qui ont utilisé ou vu des démonstrations du poste LAO (à la BNF) ou de la SLAV (à l'Inathèque) ou d'un certain nombre de stations de lecture savante sur ordinateur dans le monde, la lecture assistée par ordinateur est un concept opératoire. Le chercheur qui utilise de tels postes de travail ne dispose pas seulement du foisonnement en réseaux d'informations présentes sur le "net" ou dans la bibliothèque qu'il fréquente, mais d'un outil à même de l'aider à sélectionner et à structurer toutes ces données et références, à les mettre en parallèle avec les siennes propres, à disposer d'outils d'analyse linguistique, de statistiques lexicales, de pose de signets, balises et autres annotations, tout en même temps que d'un environnement hypermédia d'écriture. Dans le cas de la SLAV, les mêmes dispositifs sont élargis pour répondre aux spécificités audiovisuelles (structuration automatique ou à la demande d'une bande vidéo en plan élémentaire, saisie d'images, pose de balises et de zones de notes, mise en correspondance d'un document et de ses références documentaires etc.).

On comprend que de tels postes de travail séduisent les chercheurs qui les ont essayé.

L'inconvénient de la LAO, en l'état où elle a été étudiée à la BnF, puis abandonnée tient précisément en son avantage fondateur d'être adaptable à la personnalité et aux hypothèses du chercheur.

Dans l'un et l'autre cas, (LAO et SLAV et CD-ROM édités), la difficulté est la même, la richesse et la créativité des liens interactifs enrichissant les documents constituent l'obstacle majeur à la prise en compte globale de gros corpus hypermédia.

Paradoxalement si le lien multimédia, permet de se relier facilement à l'information présente dans le "Cybermonde entier", sur les stations de lecture les plus sophistiquées et sur tous les CD-ROM possibles, le chemin inverse est beaucoup plus difficile.

Pour pouvoir proposer des données hypermédia en réseaux, il est nécessaire de s'entendre sur une norme de fonctionnement syntaxique non seulement formelle (comme c'est le cas avec HTML) mais aussi capable d'être porteuse de sémantique mais surtout il est nécessaire de s'entendre sur des consensus de méthode entre chercheurs. Il est ainsi fondamental que les avantages de l'information structurée puissent s'appliquer rigoureusement au traitement des références et au traitement d'un balisage consensuel (évidemment à l'intérieur de communautés d'analyse des contenus qui ont préalablement défini des règles de fonctionnement intradisciplinaire et transdisciplinaire) c'est le cas de la TEI.

### **La TEI : "Text Encodind Initiative"**

La TEI que l'on pourrait traduire par groupe d'initiative pour le balisage normalisé des textes est effectivement une norme de balisage, de notation et d'échange de corpus des documents électroniques

fondée sur SGML (et maintenant sur XML que la communauté des chercheurs TEI a largement contribué à définir technologiquement<sup>27</sup>).

Elle s'est élaborée pragmatiquement à partir des besoins de structuration, de conceptualisation et de mise en réseau de textes. À partir des premières hypothèses élaborées dans la réunion du Vassar Collège (Poughkeepsie, New York 1987), un long travail de recherche de consensus entre scientifiques intéressés par l'étude des textes (spécialistes de littérature, historiens, sociologues, linguistes, ethnologues, philosophes...) a permis de définir un ensemble de recommandations de codage : le « Guideline TEI<sup>28</sup> »

La TEI est une norme qui permet comme sur les stations de lecture assistée par ordinateur de faire des annotations, structurations, ou marquages mais qui, en outre, permet d'échanger en réseaux les résultats de ces lectures savantes. En allant vite, on pourrait dire que les travaux de l'Inathèque et de la BnF se sont surtout attachés à créer des interfaces hypermédia ergonomiquement séduisants pour la lecture savante sans préjuger des typologies du chercheur et en reportant à plus tard la problématique réseau. La TEI s'est attachée au contraire à répondre en priorité à ces deux dernières questions, mettant entre parenthèses la première question pourtant primordiale. Par principe la TEI ne préjuge d'aucun matériel : c'est ce qui fait son succès, mais aussi son talon d'Achille<sup>29</sup>.

### **Les nouveaux enjeux théoriques de la documentation, de la bibliothéconomie dans le contexte plus large d'une ingénierie du document :**

Si l'information canoniquement structurée devient progressivement la règle pour la totalité des informations susceptibles d'être éditées sur un support quelconque ou transmises notamment sur l'ensemble du réseau mondial Internet, cela entraîne des conséquences importantes pour tous ceux qui font profession de gérer des grands corpus d'informations.

- La totalité des références de documents de quelle nature ou de quelle origine qu'ils soient, doit pouvoir constituer un seul grand catalogue collectif non ambiguë. Ce qui nous pousse à approfondir et à systématiser des notions catalographiques qui nous sont familières mais qui doivent s'adapter pour correspondre aux nouveaux paradigmes.
- La totalité des moyens de « nommer » et au-delà de toutes les ressources descriptives doivent aussi pouvoir (ne serait-ce que formellement) fonctionner de façon non ambiguë. Ce dernier point implique aussi que nous approfondissions une autre activité que nous connaissons bien : faire fonctionner et animer des groupes de terminologie, y susciter des consensus, organiser des vocabulaires. En fait réactiver selon une autre philosophie le savoir-faire indispensable des thesaurus.

Ces notions se sont précisément fait jour au fur et à mesure des avancées théoriques. Ces deux derniers points impliquent que nous sachions redéployer une partie importante de notre activité en direction d'une collaboration internationale pour savoir créer des centres de ressources (qui peuvent être virtuels par la dispersion matérielle des informations) rassemblant des « ressources autorités » mondialement structurées, correspondant aux ressources de références catalographiques et aux ressources de descriptions qui peuvent être des noms, mais aussi quantité d'autres catégories de

---

<sup>27</sup> C'est dans cette lignée de bibliothèques virtuelles à valeur ajoutée que s'inscrivent les travaux de la TEI qui ne peuvent pas se dissocier des grands logiciels de balisage informatique du texte. Les mêmes chercheurs se retrouvent d'ailleurs à l'origine de la TEI, de HyTime, de XML. Ainsi par exemple :

- Steve De Rose, Inso corp. & Université Providence (TEI board & création des pointeurs et des liens dans XML)
- C.M. Sperberg Mc Queen, Université de Chicago (TEI board & spécification générale de XML)
- Henry Thompson, HCRC (Human Computer Research Center) Edinbourg (Ecosse), (TEI board & définition de XSL).

<sup>28</sup> Lou Burnard et C.M. Sperberg-Mc Queen, traduction François Role, La TEI simplifiée : une introduction au codage des textes électronique en vue de leur échange pp23 à 152 in Cahiers Gutenberg n° 24 juin 1996 ed. Irisa Rennes.

Ce numéro spécial TEI des Cahiers Gutenberg n° 24 est une excellente introduction en Français à la TEI.

<sup>29</sup> Silfide (Serveur Interactif pour la Langue Française, son Identité, sa Diffusion et son Étude) est un serveur de littérature francophone en ligne sous l'égide de l'AUFPELF et du CNRS. En s'appuyant sur les nouvelles technologies de l'information structurée (XML) le Loria développe actuellement une base francophone de littérature structurée selon la norme TEI selon des interfaces qui deviennent très conviviaux. Silfide est actuellement hébergé sur le serveur du Loria (Inria Lorraine) <http://www.loria.fr/projets/Sifide/>

données descriptives<sup>30</sup>. La norme RDF (Ressources Description Framework) qui propose un langage (ouvert mais normalisé) de description universelle de données devrait, sur ces points nous aider à organiser le cadre dans lequel nous pouvons ainsi redéployer notre savoir faire documentaire.

Les concepteurs de ces technologies ont longuement approfondi les notions théoriques et pratiques de la documentation. Ils se sont longuement appesanti pour trouver des solutions aux problèmes complexes de fonctionnement de plusieurs thesaurus en parallèle. Ils ont mis au point des logiques qui permettent de faire coexister des sémantiques concurrentielles.

Pour ce qui est du traitement de la référence le savoir-faire des spécialistes de l'information structuré s'appuie sur un savoir faire catalographique avancé qui est le résultat de la mise en synergie des approches savantes (TEI notamment) et des approches industrielles ou commerciales.

Les approches industrielles ou commerciales s'appuient, elles, sur le nécessaire développement d'une gestion systématique des documents destinée à accompagner toute chose ou toute action pour organiser l'automatisation des tâches intellectuelles et manuelles. En effet dans un contexte d'intégration des processus d'automatisation, toute activité industrielle ou commerciale ne se comprend que si l'ensemble des chaînes d'actions est accompagné par des documents dont il est fondamental de maîtriser les deux aspects référentiels et sémantiques, de savoir repérer et référencer avec précision un document<sup>31</sup>, de savoir mettre en œuvre sa « traçabilité<sup>32</sup> » dans les modifications de ses versions successives.

### **MPEG OU L'EMERGENCE D'UNE MACHINE GRAMMATOLOGIQUE DEPLOYEE DANS LA DIMENSION AUDIOVISUELLE :**

« La photographie est inclassable parce qu'il n'y a aucune raison de marquer telle ou telle de ses occurrences ; elle voudrait bien, peut-être se faire aussi grosse, aussi sûre, aussi noble qu'un signe, ce qui lui permettrait d'accéder à la dignité d'une langue ; mais pour qu'il y ait signe, il faut qu'il y ait marque ; privées de principe de marquage, les photos sont des signes qui ne prennent pas bien, qui tournent comme le lait. »

Roland Barthes

J'ai proposé le concept de machines grammatologiques<sup>33</sup> pour bien souligner non seulement l'émergence de ce phénomène de la mise en information structurée des informations présentes sur le réseau et sur toutes sortes de supports numériques mais aussi leur entrée en lisse dans notre nouvel environnement cognitif technoculturel. En introduisant le concept de machines grammatologie, je veux souligner le phénomène de mise en concordance globale de la grande diversité des langages et modes d'expression répandus sur la planète.

J'appelle cette machine grammatologique parce qu'elle perfectionne au maximum ce qui est envisageable aujourd'hui : la mise en concordance, l'accès immédiat de tous les éléments, de tous les grains d'information où qu'ils soient localisés sur un système quelconque informatique relié en réseau. On progresse ainsi notablement dans le domaine d'une grammatologie instrumentale de l'information. Je rappelle en effet que la grammatologie représente selon Derrida<sup>34</sup>, certes une

---

<sup>30</sup> Par exemple un graphe, une mesure, une formule mathématique, un son, une image...

<sup>31</sup> ... qui peut être un flux d'ordres industriels ou commerciaux et prendre la forme d'un bordereau, d'une facture, d'un plan, de la commande d'une machine outil, d'une commande de pilotage d'un engin, et bien sûr des commandes en retour permettant d'effectuer automatiquement ces chaînes d'actions : mesure de l'action réalisée en retour de l'ordre, comparaison avec l'ordre donné en commande ou vérification de la bonne réception d'une commande suivie des ordres de paiement etc., etc...

<sup>32</sup> Le mot à la mode de traçabilité n'est pas énoncé par hasard. Il s'agit bien d'un des savoir-faire fondamentaux de notre civilisation de communication virtuelle et réseautique aussi essentiel pour le destin civilisationnel de nos sociétés que le fut le développement des transports et des marchandises.

Du chef de rayon d'un grand magasin établissant ses stocks grâce au lecteur à code-barre de son terminal portable qu'il amène dans les rayonnages, au médecin hospitalier notant ses observations et interventions sur un dossier de malade qui communique avec sa carte à puce de Mutuelle, du fermier ouvrant tel sac d'engrais référencé, au chef de chantier calculant ses plans de charge, tous doivent aujourd'hui non seulement maîtriser des savoirs dont on sait de moins en moins distinguer l'aspect manuel et matériel, de l'aspect intellectuel mais aussi l'associer à une mise en documents structurés en réseaux.

<sup>33</sup> Henri HUDRISIER, SGML, HTML, XML : L'ère des machines grammatologiques, in « Passerelles » (Revue de l'Université de Paris 8) n°26, Dossier spécial Recherche, Octobre 1999.

<sup>34</sup> Jacques DERRIDA, De la grammatologie, éditions de Minuit, Paris ; 1967. Derrida propose ce terme défini dans le Littré mais resté sans aucun usage, il signale la réinvention de ce terme par I.-J. GELB en 1952 (I.-J. GELB, Pour une théorie de l'écriture, Flammarion, Paris ; 1973). La première édition originale (University of Chicago Press, Chicago ; 1952) paraît sous

interrogation philosophique sur la nature complexe et plurielle de l'information vue sous son aspect théorique et intellectuel, mais aussi une interrogation pragmatique sur son destin communicationnel en terme de dispositif physique ou matériel. Cette interrogation croise fondamentalement l'interrogation sur la disparité entre le son, l'image (fixe, animée ou associée au son), le texte, le geste (chorégraphie, langue des signes). Complémentairement à la machine grammatologique textuelle, une machine grammatologique déployée dans ses dimensions audiovisuelles et multimédia est en cours de développement. Elle est susceptible d'instrumentaliser le fonctionnement structurel et engrammatique de la globalité multimédia<sup>35</sup> et ainsi de déboucher, on peut le penser, sur une meilleure visibilité (en fait la première réelle visibilité) des interrelations grammatologiques de la totalité des catégories de documents. Elle peut recomposer en mosaïques significantes et cohérentes, la totalité des avancées théoriques et des savoirs professionnels jusqu'ici dispersés et impossibles à faire fonctionner facilement en synergie.

### **La mise en orbite des trois étages de la fusée MPEG**

La norme MPEG est généralement connue du grand public comme une simple norme technique de compression de l'image vidéo. Cette définition est parfaitement exacte, mais elle ne recouvre que la première phase de développement de cette norme (les couches MPEG 1 et 2) alors que dans ses couches en émergence, cette norme a vocation à devenir le langage structuré et balisé de l'information audiovisuelle : sa machine grammatologique.

« La normalisation de la codification de l'information audiovisuelle (MPEG Moving Pictures Experts Groups) peut se comparer aux 3 étages d'une fusée. Les deux premiers étages, MPEG 1 et 2 sont voués à n'être que des technologies de transition. Bien sûr on nous les vend aujourd'hui sous l'appellation de « nouvelles technologies de la compression d'image ». Il est clair que ces 2 étages (MPEG 1 et 2) seront consommés par l'atmosphère et pour les centres de recherche industrielles ils sont considérés comme le moyen d'acquérir une culture et un savoir-faire industriel suffisant pour atteindre le but visé, MPEG 4 (l'étage de la télévision numérique interactive). A cet étage, il faudra bien sûr associer des niveaux de services :

- MPEG7, déjà en fin de définition de norme ISO qui répondra aux questions documentaires : décrire et retrouver des documents audiovisuels dispersés partout dans le monde,
- MPEG 21 qui devra répondre aux questions de la normalisation du mode d'échange (juridique, commercial, confidentialité, sécurité, etc.), et qui ne sera mis en chantier de normalisation que lorsqu'un certain nombre de prototypes et de propositions de recherche-développement pourront rentrer en lisse<sup>36</sup>. »

Cette métaphore de la fusée rend compte de la logique industrielle et marketing selon laquelle l'information structurée audiovisuelle se met progressivement en place.

Le niveau des ressources financières potentiellement disponibles pour les recherches et les développements liés à l'audiovisuel et surtout à la télévision est absolument fantastique et proportionné à l'immense chiffre d'affaires mondial que dégage ce secteur. Mais si les industriels ne veulent pas casser leur marché, ils doivent amener les consommateurs à appeler de leurs vœux la transformation progressive du paysage technique et les conséquences culturelles que cela entraîne.

MPEG 1 comme pour l'image fixe JPEG est une norme définissant une compression du signal à la fois spatiale, temporelle et adaptée aux conditions psycho-perceptives de l'image et du son. Sans la mise en chantier de cette définition arbitraire d'un code commun, les divers industriels pionniers de l'image numérique n'auraient pas su investir dans le long terme des recherches-développements, qui leur ont permis de produire des composants à même de générer un marché et d'induire nombre d'applications multimédia (notamment sur CDrom), qui composent le paysage numérique de la fin des années 90. Mais les très lourds investissements, que nécessitent la fabrication des composants vidéo-numériques à des niveaux d'intégration très poussée de miniaturisation, de fiabilité obligatoire pour

---

le titre A study of writing, the foundations of grammatology. Pour des raisons explicables mais qui sortent de notre problématique Gelb ne poursuit pas l'étude de la grammatologie dont il est l'inventeur anglophone.

<sup>35</sup> ... incluant bien sûr l'audiovisuel et à terme le texte considéré lui aussi comme un sous ensemble du multimédia (ce qui est un retournement techno-culturel)

<sup>36</sup> Jean-Michel Borde et Henri Hudrisier, Standardisation et normalisation : au cœur de la technologie du document, in Solaris n° 6, Revue électronique et ligne de l'URFIST, Paris, novembre 1999

survivre sur le marché grand public et a fortiori sur le marché des professionnels, ne se retrouvent pas en quelques mois. L'industrie et avec elle la recherche publique ont besoin de définir des étapes : celles des étages fonctionnels successifs de la fusée MPEG. Chaque étape représente un niveau cohérent conçu pour être l'objet d'une appropriation progressive certes, mais organisée par paliers reconstituant dans le domaine des objets techniques une évolution par mutations analogue à ce que connaissent les biologistes pour les espèces animales.

Ainsi nous nous sommes accoutumés aux divers paradigmes de l'image et du son numérique de niveau 1 adaptés à des usages faciles et compréhensibles sur nos micro-ordinateurs et, à leur stockage sur un support déjà ancien qui avait été créé pour le son mais non pour l'audiovisuel. Dans MPEG 1 la spécification a été, volontairement, bridée dans une classe de débits bien précise 1.5 Mb/s et avec pour vocation de servir à l'archivage sur support de distribution de masse (comprenez le CDROM.) Dans le monde de la microinformatique, on parle souvent de « killer application », c'est en effet dans un beau feu d'artifice que MPEG1 va orchestrer les obsèques du CD-ROM : il n'est pas assez volumineux pour ... contenir un film complet. Reportez-vous aux unes de vos journaux financiers, entre 1993 et 1994, relatives à la bataille des titans de l'électronique grand public de laquelle naîtra : le DVD.

MPEG 2 monte en puissance en terme de débit, de performance de la compression et des procédures de télécommunications mais ne change pas notablement de philosophie technique. Son seul rôle est de mettre en orbite le troisième étage de notre fusée. Sa vocation d'étage de transfert va lui donner une durée de vie industrielle sensiblement plus longue que les précédentes étapes : il lui faut survivre jusqu'à la mise à poste de la charge utile : le satellite MPEG4. »

#### **MPEG 4, une télévision numérique déployée dans l'interactivité**

MPEG 4 est une norme de télévision numérique vraisemblablement destinée à connaître une durée de vie équivalente à nos anciennes normes de télévision analogique (PAL, SECAM ou NTSC). Pour que cette prédiction se confirme, il faut :

1° que cette norme soit ouverte et évolutive

2° que l'audiovisuel et le multimédia dans son ensemble s'intègrent dans une culture globale du document structuré, transcendant en quelque sorte l'idée encore étroite que nous pouvons nous faire d'une télévision interactive.

La télévision numérique interactive, que l'on peut déduire des possibilités techniques de MPEG 4 en est la première brique, mais il faut véritablement savoir y associer les services que pourront rendre MPEG 7 et MPEG 21 et surtout voir dans la filière MPEG 4, 7 et 21 une dynamique d'achat d'équipement et d'usages professionnels à même de fédérer dans leur mouvement de convergence, un très grand nombre de composantes médiatiques ou modales : la médiation de l'achat, la visiophonie, la communication d'assistance au pilotage des véhicules et une quantité d'usages plus spécialisés (télésurveillance, téléconférence, enseignement, possibilité de dicter des énoncés en langues des signes, de disposer de systèmes d'analyse de scènes visuelles utilisables par les aveugles<sup>37</sup>, etc.).

Techniquement MPEG 4 se présente comme un système de multiplexage-démultiplexage<sup>38</sup> très sophistiqué. L'image vidéo est décomposée non seulement dans ses divers canaux constituants (son et vidéo), mais ceux-ci se décomposent eux-mêmes en objets multimédia de base et en graphes décrivant la composition des scènes. Un objet peut être un personnage, un décor, un objet sonore. Il peut lui-même se décomposer et, n'importe lequel des sous-ensembles d'un objet (par exemple un visage d'un

<sup>37</sup> Ce sujet pourrait lui aussi faire l'objet d'un développement rédactionnel important. Il n'est pas impensable, cela se fait déjà, de proposer à des aveugles de s'équiper de combinaison de données tactiles (délivrant des sensations tactiles en d'autres endroits du corps que l'extrémité des doigts), d'aides audio ou d'autres interfaces à même de leur rendre compte en temps réel de l'évolution de leur environnement visuel. On sait par ailleurs que cela correspond à un désir déjà largement exploré dans cette communauté et chez ses formateurs spécialisés : aviation de tourisme (les miraux volants, football, visites touristiques et culturelles, etc.)

<sup>38</sup> Il est essentiel de ne pas se laisser impressionner par ce terme d'allure technique et barbare. Par multiplexage on peut entendre que les différentes composantes d'un langage aussi composite, donc complexe, que le cinéma ou la vidéo sont susceptibles d'être mis en bandes composantes parallèles, potentiellement indépendantes mais mixées virtuellement ensemble. Par démultiplexage, il faut entendre que la réciproque est vraie, ce qui signifie qu'à tout moment un élément quelconque du composite pourra être démixé, soustrait du document et éventuellement remplacé par un autre.

personnage, un tableau dans un décor, le dialogue dans une bande sonore) peut être remplacé par un autre. On accède ainsi à un audiovisuel véritablement interactif. On comprend dès lors, que dans certaines applications, on puisse parvenir à des taux de compression extrêmement forts mais aussi à ce qu'on puisse réaliser des applications interactives jusqu'ici inenvisageables. Par exemple un film d'art contenant une grande majorité de plans constitués en majeure partie en banc-titre sera numérisé comme une série d'images fixes associées à leurs graphes de déplacement x, y et z et à une ou plusieurs bandes-son puisque, on le comprend, la dimension sonore peut se déployer dès lors dans le multilinguisme ou la diversité des niveaux d'usage.

Autre exemple : un catalogue de marchand de meubles de cuisine pourra, à partir d'une seule vidéo de présentation et de montage des meubles, proposer de multiples versions de cette séquence y compris en intégrant dans le décor les photos numérisées du décor et des volumes de la maison du futur acheteur. On pourrait même substituer les visages et les habits des acheteurs aux têtes des comédiens qui auraient joué la scène.

Application professionnelle, extrêmement utile : un film ne sera pas véritablement mixé comme en analogique mais le mixage construira un graphe de mixage des différentes voies et objets sonores. Il sera ainsi beaucoup plus simple qu'aujourd'hui d'obtenir une VI ( Version Internationale, c'est à dire une version comprenant la musique, les bruitages, les ambiances sans les dialogues, indispensables pour pouvoir doubler un film)<sup>39</sup>.

Autre conséquence, qu'il faut voir comme une brique, composante essentielle de la réalisation d'indexation d'images et de requête d'images : la décomposition en objets multimédia de base et en graphes est un apport jusqu'ici jamais égalé pour la description documentaire systématique des images, des sons, des plans, des séquences tant dans leur aspect formel que sémantique (connaissance de certains objets dans le décor ou dans les corps, visages, costumes...). D'où la nécessité de disposer de MPEG4 pour rendre possible MPEG7. En effet, ceux à qui on présente le niveau MPEG 7, sans qu'ils aient véritablement intégré et compris le niveau MPEG 4, ne comprennent pas, ne croient pas qu'une documentation normalisée de l'audiovisuel à un niveau mondial puisse voir le jour. C'est parce qu'ils raisonnent avec les anciens paradigmes de techniques documentaires dans lesquels, il était impossible de fournir commodément une description des objets audiovisuels autres que textuelle. Le niveau MPEG 4, par construction, produit (en sous-produits si on peut dire) une analyse détaillée plan par plan, image après image et objet après objet de la totalité des documents. On comprend dès lors que si on sait fabriquer un méta-niveau organisant cette description locale dans une logique de bibliothèque virtuelle sophistiquée (MPEG7) on puisse parvenir par l'alliance des deux niveaux à une performance documentaire audiovisuelle jusqu'ici absolument impensable.

Dès lors MPEG 4 se comporte véritablement comme un satellite sorti du troisième étage d'une fusée : pour fonctionner, il doit déployer ses antennes (l'équivalent de la couche MPEG7 qui lui permet d'assurer la communication sous forme d'indexation et de requête) et ses panneaux solaires (l'équivalent de la couche MPEG 21) qui lui donneront son énergie essentielle : l'augmentation de ressources par l'automatisation de l'échange marchand.

### **MPEG7, la bibliothèque virtuelle de l'audiovisuel**

MPEG7 est une étape fondamentale des nouveaux champs de possibles ouverts à l'image numérique dans le contexte du document structuré. MPEG7 répondra avec une efficacité qui n'a pas fini de nous étonner à des questions laissées encore aujourd'hui sans solution vraiment satisfaisante : « si on sait faire sans trop d'impasses méthodologiques un catalogue et un fichier textuel pour la bibliothèque, comment réaliser homothétiquement, c'est-à-dire avec des fichiers sonores et visuels celui d'une cinémathèque, d'une phonothèque, d'une photothèque ? »

Roland Barthes indique bien ce désarroi de l'intellectuel à l'aise dans le monde des idées portées et pensables par la langue et le texte. Pour lui, la photographie et le cinéma glissent sous l'analyse, pour

---

<sup>39</sup> Pour le son comme pour l'image, un certain nombre d'applications professionnelles préfigurent depuis deux ou trois ans ce qui sera la norme dans MPEG 4. Il est indispensable de bien distinguer des effets rendus possibles en préfiguration de la norme MPEG 4 elle-même, qui généralisera à très bas coût et pour tout usager, ce qui aujourd'hui est de l'ordre du professionnel à des coûts incompatibles avec l'usage grand public ou institutionnel.

reprendre ses propres mots : « tournent comme le lait »... «pour qu'il y ait signe, il faut qu'il y ait marque ; privées de principe de marquage, les photos sont des signes qui ne prennent pas bien ».

La norme MPEG7 n'est pas limitée au cinéma et à la vidéo mais elle s'ouvre évidemment à des stocks de photographies d'objets ou d'oeuvres d'art en trois dimensions et de bibliothèques sonores. Elle ne peut véritablement se déployer que si elle dispose de « marques » : les langages à balises (Markup Languages), ont précisément cette qualité.

MPEG 7 est donc une couche décrivant une méthode normalisée d'indexation et de requête dans une banque d'images. C'est aujourd'hui une norme non encore entièrement définie. Il est donc impossible de décrire aujourd'hui ses fonctionnalités précises. A partir des compte-rendus des différents meetings MPEG 7 et vu l'imminence des échéances<sup>40</sup>, on peut faire un certain nombre d'hypothèses sur les fonctionnalités qui seront effectivement proposées dans la logique MPEG 7 version zéro.

Ce qui est certain c'est qu'on sera dans un contexte d'informations structurées fondé sur XML, qui permettra de construire des bibliothèques virtuelles multimédia décomposant les documents audiovisuels et multimédia associés à leurs multiples métadonnées (document de production, scénario, partition musicale, contrat juridique, plan de décor, plan de tournage etc...) en grains d'informations susceptibles de nourrir des systèmes d'indexation et de recherche automatique à partir de très nombreuses facettes, des plus traditionnelles aux plus innovantes (recherche par les sonorités, recherche d'image par la forme, la couleur ou le rythme des séquences et des plans etc...).

Bien que le format MPEG 7 soit prévu pour fonctionner avec tout type de matériel audiovisuel normalisé, il ne sera évidemment efficace que s'il est intégré dans une chaîne numérique MPEG complète (entendez MPEG 4 et 7 puis à terme MPEG 21).

On se retrouve dès lors dans des situations de bibliothèque virtuelle comparable à celles que l'on commence à connaître pour le texte. La recherche dans de gros corpus structurés de littérature permet des requêtes sémantiques subtiles, rendues possibles par la mise en ligne de très nombreuses occurrences des mots, des tournures stylistiques, des phrases-types pouvant ressortir de « patrons linguistiques ». Pour l'audiovisuel, des heuristiques aussi riches seront possibles parce qu'on pourra constituer des banques structurées de décors, de sonorités, de visages. En ce qui concerne les visages, ils sont dans de très nombreux cas cités dans les génériques, dans les commentaires, voire dans les « cavaliers » indiquant le nom des journalistes à l'antenne. De ce fait, une banque documentaire de visages permet de retrouver et de reconnaître ceux qui ne l'auraient pas été. Il en sera de même pour les banques de donnée de paysages, d'objets, de tableaux, d'animaux, de botanique etc...

Le document audiovisuel organisé selon la norme MPEG7 sera structurellement lié à nombre de métadonnées qui aujourd'hui en sont normalement dissociées. Ainsi sera logiquement et informatiquement organisé le contenu numérique du synopsis, du scénario, du dialogue, du/des sous-titrage(s) éventuel(s), du découpage, des documents de production, du contenu textuel du générique, des images clefs, des partitions musicales, des références catalographiques... mais seront aussi liés, toujours dans la même structure logicielle normalisée, des données intrinsèques à l'image, au son et à leur structure temporelle de déroulement en plan ou séquences, qui peuvent être issues, mais pas exclusivement<sup>41</sup>, d'une reconnaissance automatique des formes (vocales, musicales, visuelles) mais aussi à la reconnaissance de structures de plus haut niveau (la macrostructure de la séquence peut se construire automatiquement par recoupement de cooccurrences formelles dans la structure plus atomique du plan que l'on sait isoler automatiquement avec un bon degré de fiabilité).

---

<sup>40</sup> Le « Comity Draft » de MPEG7 doit être présenté à la réunion internationale de MPEG en octobre 2000 et distribué officiellement aux industriels pour qu'ils développent des applications industrielles normalisées en janvier 2001

<sup>41</sup> Les informations sémantiques et structurelles peuvent provenir aussi d'une analyse et d'une description humaine ou d'une « déclaration numérique » de la signification audiovisuelle par le producteur du document. Cette notion est très importante parce qu'elle permet d'envisager que nombre de documents seront décrits sémantiquement, structurellement et référentiellement (données catalographiques) dans le cadre de la post-production numérique du document audiovisuel. On comprend dès lors l'importance de ces moteurs d'images expérimentaux qui cherchent l'image par l'image, ou le son par le son. Ils ne trouvent toute leur place que dans le contexte plus large de l'information structurée.



Ce nouvel environnement informatique a des conséquences cognitives et épistémologiques qu'il est nécessaire de prendre en compte au niveau académique. La dynamique patrimoniale, documentaire, catalographique et l'organisation matérielle d'une mise en réseaux, d'une numérisation des documents et leur préservation muséale comme oeuvres originales doivent se mettre en synergie avec les dynamiques de lecture savante ou d'usage pédagogique.

Pour ce qui est de la mise en valeur de la signification (explicite ou savante), les balises sémantiques et structurelles de la TEI (accessibles en XML et en MPEG7) représentent notamment (mais il doit s'en créer d'autres spécifiques à l'audiovisuel) une ressource conceptuelle normalisée mais ouverte<sup>42</sup>, qui est déjà le fruit de consensus entre chercheurs. Ainsi nombre de grands instituts de recherches (tant industriels que scientifiques) contribuant à la finalisation du cadre normatif MPEG7 s'investissent dans l'étude savante (sous formats TEI) de textes théâtraux mais aussi de scripts ou de commentaires de cinéma et de télévision avec lesquels l'entité de format « TEI drama » est compatible.

Pour ce qui est de l'organisation matérielle des oeuvres, de leurs fragments (plans ou séquences d'archives), de leur logique de corpus (fonds), de leur mise en réseaux, c'est d'une part la logique MPEG 4<sup>43</sup> mais aussi d'autres logiques normatives d'organisation qui rentrent en lice notamment DAVIC (Digital Audio Visual Council) un standard de structure qui constitue une « recommandation<sup>44</sup> » d'un groupement de professionnels de la télévision interactive leur permettant de normaliser leur procédure d'organisation des documents et des équipements d'émission et de réception. »

Admettons cependant que les professionnels<sup>45</sup> comme le grand public, auront besoin de quelques années pour apprivoiser les applications rendues possibles par cette extraordinaire mutation des technologies de l'audiovisuel et du multimédia.

### **MPEG 21: l'étage de l'échange marchand audiovisuel.**

Certes MPEG 21 est prévu comme la couche de l'échange de la norme MPEG<sup>46</sup>. Cependant les organisations de normalisation (sous la recommandation de l'ISO) se sont données consigne d'attendre

---

<sup>42</sup> Un chercheur de la communauté TEI cherchera à produire son analyse en s'appuyant sur les balises de sa (ou de ses) discipline(s) d'appartenance, mais il peut toujours, soit proposer de nouvelles balises (sémantiques ou structurelles) soit créer des sous catégories sur des balises existantes.

<sup>43</sup> MPEG 4 dans son ciblage d'usage est certes fortement lié à l'aspect matériel et au potentiel interactif du document audiovisuel lui-même, mais MPEG dans sa globalité fonctionne en préservant toujours d'une couche à l'autre des zones d'interface et de recoupement destinés à assurer la montée en charge temporelle des couches (MPEG 4 générera vraisemblablement ses premiers produits industriels dans 1 ou 2 ans alors que MPEG7 n'en générera que dans 3 ou 4 ans) en attendant à terme leur fonctionnement en synergie. Ainsi dès MPEG 2, puis dans MPEG4 des zones logiques (container) permettent l'association de métadonnées relatives au document audiovisuel (copyright, mots clefs, notice catalographique voire analytique). Il est à noter que dans MPEG4 ces métadonnées sémantiques, structurelles et référentielles sont primordiales pour assurer le fonctionnement de l'interactivité que vise ce standard qui se veut celui de la télé-interactive (et au-delà de l'audiovisuel ou du multimédia interactif en général)

<sup>44</sup> une recommandation est un stade intermédiaire entre une norme et un standard. Une recommandation est ouverte, elle a vocation à assurer le codéveloppement de ceux qui l'éditent. Presque toutes les « normes » du W3C, le comité des développements sur Internet, sont abusivement appelées normes et ce sont souvent des recommandations.

<sup>45</sup> .. et pour ce qui nous concerne les iconothécaires non directement intéressés par la vidéothèque ou la cinémathèque en ligne : muséographes, bibliothécaires, historiens d'art

<sup>46</sup> On se pose souvent la question de savoir quelle logique se cache derrière la progression mathématique difficile à identifier (1, 2, 4, 7, 21). Au début la progression est claire 1, 2, il exista un groupe MPEG 3 qui s'est spécialisé sur les aspects strictement sonores et qu'il faut donc mettre entre parenthèse de la progression des couches MPEG à vocation générale (on trouve assez souvent cette référence chez les spécialistes du son) et nous arrivons normalement à MPEG 4. Au delà les explications divergent certains prétendent avoir laissé de la place pour des groupes d'expertises qui auraient du s'intercaler, d'autres avancent l'explication plus plausible que cela tient à une certaine coquetterie du groupe MPEG qui fait ainsi planer un certain mystère sur ses travaux. MPEG 7 succède ainsi à MPEG 4 dans la progression historique des couches et maintenant MPEG 21 est prévu pour être la couche qui lui succède historiquement. Pour la petite histoire, cette couche destinée au commerce et à l'échange, a été nommée sous différents numéros d'identification dans ses toutes premières définitions dans la terminologie du groupe MPEG : ainsi on évoqua d'abord MPEG 11, puis MPEG 12 et le groupe qui doit officiellement lancer sa mise en chantier en octobre 2000, semble s'être accordé pour la dénommer MPEG 21 ! Attendons la suite du feuilleton, mais de source autorisée (comme on dit) cela semble définitif pour cette couche fonctionnelle. Gageons que le groupe MPEG aura encore fort à faire pour intégrer les couches entres elles, finaliser une intégration qui serait plus multimédia et textuelle... Les prochains épisodes risquent d'être aussi passionnants et certains happy few de la technologie

la fin de l'année 2000 pour amorcer la mise en chantier des projets de la norme MPEG 21. Cela tient à ce qu'il est « urgent d'attendre » que la technologie de l'échange informatisé et numérisé normalise un grand nombre de ses procédures et langage pour ne pas faire porter sur la recherche audiovisuelle, la mise au point de segments qui ne lui sont pas spécifiques.

Une convergence bien pensée et bien conçue exige que chaque métier convergent, chaque média convergent, chaque modalité convergente accomplisse la part de tâche qui lui incombe. MPEG 21 se devra de concevoir toutes les situations d'échange, de location, de vente, de mise à disposition, de redistribution de droits, de préservation de droits moraux etc...

MPEG 21 devra intégrer les paramètres juridiques et commerciaux propres à tel ou tel réseau, tel ou tel producteur, tel ou tel type de client. Il devra intégrer les législations et les usages de copyright ou de droits d'auteurs propres à tel ou tel pays.

MPEG 21 doit rendre possible la réalisation de vente microscopique d'extraits qui actuellement coûteraient beaucoup plus chers que le prix objectif pour retrouver un document dans une banque d'images à travers une requête humaine non-informatisée ou informatisée de façon traditionnelle (c'est le niveau MPEG 7 obligatoirement associé à MPEG 21).

MPEG 21 permettrait ainsi la mise à disposition commerciale, entièrement automatisée, de produits ou d'œuvres audiovisuelles, dans des conditions juridiques et commerciales absolument contrôlées. Ainsi s'ouvriraient des usages véritablement inouïs, qui mettront la quasi-totalité des grands catalogues audiovisuels à portée de l'oeil et de la bourse d'une clientèle grand public désormais directement accessible.

## **PERSPECTIVES PATRIMONIALES**

Tout autour de moi, c'était la même furieuse clameur, sur ce thème de l'anéantissement de l'art : la liquidation de son symptôme central - la figuration - par le matériau et le document ; de son sens par l'absence de sujet ; de ses lois organiques par la construction ; de son existence même, abrogée et remplacée par une reconstitution pratique et réelle de la vie, sans l'intermédiaire des fictions et des fables.  
S.M. Eisenstein<sup>47</sup>

Ainsi commença ma double activité artistique, qui ne cessa d'unir le travail créateur au travail analytique : tantôt en commentant ma production par l'analyse, tantôt en vérifiant sur elle telle ou telle de mes hypothèses théoriques.  
S.M. Eisenstein<sup>48</sup>

Dès les premiers pas, celui du classement (il faut bien classer, échantillonner, si on veut constituer un corpus), la photographie se dérobe. Les répartitions auxquelles on la soumet sont en effet ou bien empiriques (professionnels/amateurs), ou bien rhétoriques (paysages/objets/portraits/nus), ou bien esthétiques (réalisme/pictorialisme), de toute manière extérieures à l'objet, sans rapport avec son essence, qui ne peut être (si elle existe) que le Nouveau dont elle a été l'événement ; car ces classifications pourraient très bien s'appliquer à d'autres formes, anciennes, de représentation. On dirait que la Photographie est inclassable. Je me demandais alors à quoi pouvait tenir ce désordre.  
Roland Barthes

Un film ou une vidéo (pour s'arrêter à cette instance particulière de l'audiovisuel) c'est d'abord un document qui grammatologiquement tient plus du volumen que du codex (c'est un objet séquentiel) ce qui induit d'emblée quantité de spécificités des méthodes, c'est ensuite une œuvre qui d'une part est généralement éditée sous de multiples formats, en diverses langues et qui fait appel à un très grand nombre d'intervenants auteurs ou ayant droit. Les bibliothécaires de l'audiovisuel ont du de ce fait très tôt prendre des attitudes très innovatrices en matière de catalogage des références et rejeter les méthodes de leurs collègues du livre qui traitaient en terme d'exception la pluralité des auteurs et

---

n'en finiront pas certains soirs, dans les veillées d'après colloque au milieu du Pacifique, de s'étouffer de rire en se racontant les numéros comme autant d'histoires drôles !

<sup>47</sup> S.M. Eisenstein, Préface, traduction et notes de Jacques Aumont, Mémoire 1, Comment je suis devenu réalisateur, coll 10/18, Union Générale d'Éditions, Paris, 1978, p 239

<sup>48</sup> S.M. Eisenstein, Préface, traduction et notes de Jacques Aumont, Mémoire 1, Comment je suis devenu réalisateur, coll 10/18, Union Générale d'Éditions, Paris, 1978, p 244

ayant droit, les mentions d'éditeurs (transformation, reprise, refontes, traduction, ajouts d'éléments exogènes, etc...).

Un film c'est ensuite de l'image et une sonorité, deux composantes qui à l'évidence ne se laissent pas facilement nommer avec des mots. De ce fait les bases d'images ont toujours été la pierre de discorde des projets qui centralisaient dans un même univers documentaire des livres et des audiovisuels. Les spécialistes des banques d'images ne se sont jamais sentis complètement solidaires de leurs collègues du livre et du texte. Les spécialistes des banques d'images et d'audiovisuels se sentaient plutôt soudés entre eux par une inquiétude méthodologique qui fut pour eux un perpétuel aiguillon d'innovation. Par définition pourrait-on dire, tout disfonctionne dans leurs systèmes de données tant du côté du catalogage qui pour le moins « mal fonctionne » que du côté de l'analyse du sens qui elle pêche toujours par incomplétude.

Ce paysage en apparence très technique et complexe n'est, à bien y regarder, que le résultat d'une prise en compte progressive dans la nouvelle culture numérique des contraintes et du potentiel de liberté créative<sup>49</sup> du « sens » tel qu'il s'était élaboré dans la culture analogique de l'audiovisuel. C'est pourquoi, il est essentiel, que des institutions patrimoniales audiovisuelles (a fortiori lorsqu'elles ont vocation académique) s'impliquent de très près et intègrent ces normes tant dans leur politique de conservation et de mise à disposition de leur fonds que dans l'étude savante, documentaire et pédagogique de leurs contenus et significations qui peuvent être un des points de départ possible pour le redéploiement fondamental de notre future technoculture de l'audiovisuel et du multimédia.

Les documents constitués d'images animées sonorisées (cinéma, vidéo), les images fixes (notamment les photographies et les sons (phonogrammes) posent depuis longtemps un problème théorique complexe aux spécialistes, qui comme les historiens, les sociologues, les critiques d'art, les chercheurs spécialisés dans l'étude de l'adaptation littéraire audiovisuelle veulent analyser ces types de documents. Mais les spécialistes de l'étude savante des médias ou de l'audiovisuel ne sont pas les seuls à devoir étudier méthodiquement des photographies, des enregistrements sonores, des images filmiques. De nombreux professionnels, notamment les journalistes, les policiers, les enseignants, les formateurs pratiquant l'autoscopie<sup>50</sup>, les militaires, les professionnels du cinéma, tous sont confrontés à une même question « comment inventer une méthode à même de permettre l'appréhension rationnelle des documents audiovisuels.

Nous n'avons encore rien dit du son, l'autre composante de l'audiovisuel plus facilement liée à la construction théorique grammatologique, si tant est que ses deux instances princeps (la musique et la parole) sont chacune l'objet d'une très ancienne tradition d'écriture qui les rendent plus facilement appréhendables par notre culture. Cependant par construction, pourrait-on dire, la théorie musicale communique étroitement avec la connaissance des phénomènes acoustiques (de la même façon que la peinture et la photographie communiquent avec l'optique théorique et la perspective). Pour des raisons liées à la fois au volume d'information relativement plus faible pour le son que pour l'image et à la formalisation mathématique du signal musical, le son numérique disponible pour le grand public a devancé d'une dizaine d'années l'image numérique, photographique, cinématographique et télévisuelle. Cela contribue à donner une certaine antériorité à l'étude d'aspects plus formels et plastiques de la sonorité jusqu'ici difficilement réductible à une notation musicale<sup>51</sup>.

Dès les débuts de la linguistique, se posait aussi la question théorique de savoir ce que pourrait être une sémiologie englobant dans ses méthodes la linguistique mais adaptée à l'étude d'autres langages comme l'audiovisuel.

---

<sup>49</sup> ... et paradoxalement de plus en plus humaniste; cette notion de plus grande liberté de l'utilisateur, rendue possible par la sophistication technique, nous pouvons intuitivement la comprendre quand on observe l'usage des environnements techniques complexes (comme l'automobile ou l'aviation) qui sont d'autant moins contraignants à utiliser qu'ils sont complexes (et d'ailleurs extrêmement normalisés) dans leur détail de construction technique et de fonctionnement.

<sup>50</sup> On a appris à l'occasion de la coupe du monde de rugby 1999 que l'entraîneur des « Old Black Australiens » avait mis en oeuvre un système d'autoscopie par vidéonumérique extrêmement performant. Ce système semble être construit à partir d'une version expérimentale de MPEG 4 associée à des fonctionnalités de type MPEG 7. L'entraîneur qui l'utilise est capable à la mi-temps de montrer aux joueurs concernés des phases de jeu, l'étude de leurs gestes comparés à une animation idéale, de comparer des archives sur un attaquant et d'analyser image par image les points faibles permettant une contre-attaque.

<sup>51</sup> Voir sur ce point les très nombreuses recherches notamment à l'IRCAM.

En effet, une fraction agissante des intellectuels ont longtemps pensé que tout objet d'étude (tout au moins dans les Sciences humaines) devait pouvoir se réduire in fine à un texte théorique et seulement un texte que l'on supposait pouvoir épuiser la totalité de tout questionnement théorique.

Or d'évidence, la photographie et le cinéma sont deux catégories de documents qui admettent mal cette réduction. ce que nous permettent les NTIC depuis déjà quelques années c'est de pouvoir sortir progressivement de cette contradiction logique.

Pour rédiger une légende strictement textuelle de la photo de groupe d'un conseil des ministres sur le perron de l'Elysée, on peut toujours recourir à la classique description stéréotypée de gauche à droite et du premier rang au dernier rang. Mais cette méthode montre vite ses limites et ses difficultés : comment traiter le sous-secrétaire d'état aux anciens combattants qui a mal pris le « créneau » et qui flotte entre le deuxième et le troisième rang. Une image plus sophistiquée quant à son organisation et dont on veut mettre en évidence la structure formelle et sémantique (tel le célèbre filet à provisions Panzani étudié par Roland Barthes<sup>52</sup>) sera beaucoup plus rétive à se traduire en texte. Dès lors l'association d'une logique de coordonnées (suffisante pour localiser à coup sûr le nom des ministres) ou d'une logique graphique au discours d'analyse se révélera indispensable. On pourrait multiplier les exemples dans le domaine de l'étude musicale, sonore ou cinématographique. Ainsi pour M. Merleau Ponty : « C'est le bonheur de l'art que de montrer comment quelque chose se met à signifier, non par allusion à des idées déjà formées ou acquises, mais par l'arrangement temporel et spatial des éléments<sup>53</sup> »

Si la machine grammatologique est utile pour le document textuel, on peut dire que pour l'audiovisuel, elle est non seulement utile mais indispensable et urgente. Pour le document audiovisuel, sonore ou visuel, l'information structurée au format MPEG n'apporte pas seulement comme pour le texte un surcroît de performance à des études savantes, grammaticales, lexicales, bibliothéconomiques qui sont connues pour le texte depuis l'Antiquité ; la machine grammatologique rend possible l'étude structurelle et sémantique des langages de l'image et du son jusqu'à aujourd'hui très difficiles à analyser rigoureusement.

Comme pour le texte on peut faire l'hypothèse que nous sommes confrontés à deux approches complémentaires et synergiques :

- *L'approche pragmatique industrielle et commerciale* qui s'appuie sur l'idée qu'on peut faire fonctionner un audiovisuel de la télévision et du cinéma porteur de promesses en terme d'activité professionnelle (donc de chiffre d'affaire). C'est dans ce contexte que grandit l'idée non encore techniquement aboutie mais technologiquement définie de grande vidéothèque ou de grande cinémathèque en ligne : en terme de norme d'information structurée normalisée ce niveau est MPEG7, et son niveau ultérieur de mise en ligne commerciale normalisée c'est MPEG21.
- *L'approche savante d'une bibliothèque virtuelle de l'audiovisuel et des arts.* Qu'on se rassure ces grands projets mondiaux de normalisation de l'information structurée ne se font pas sans une très lourde mobilisation de la recherche fondamentale universitaire, militaire et industrielle. Il est en outre inéluctable que, selon un schéma que notre communauté audiovisuelle connaît bien, les archives filmiques et télévisuelles soient l'objet de la première et plus grande attention et que la phonothèque soit négligée dans un premier temps puis que l'on se rende compte que c'est une brique composante fondamentale (et donc plus rapidement réalisable que la cinémathèque). On s'intéressera aussi aux photothèques qui représentent avec les agences photographiques un enjeu commercial non négligeable mais on se rendra compte aussi que le savoir théorique (bibliothéconomique, muséographique, sémiotique, d'historien d'art) est largement présent dans les bibliothèques, les médiathèques, les musées, les centres de recherche spécialisés. C'est précisément dans ce contexte que se situent nos intérêts de chercheurs spécialisés en documentation audiovisuelle.

---

<sup>52</sup> Roland Barthes, *Rhétorique de l'image*, in Communications n° 4, Seuil, Paris 1964

<sup>53</sup> M. Merleau Ponty, *Conférence à l'IDHEC le 13 mars 1945* reprise dans *Sens et non sens* et citée par Christian Metz in *Le cinéma langue ou langage ?*, in Communications n° 4, Seuil, Paris 1964

