

LA NORMALISATION DE L'AUDIOVISUEL ET DU MULTIMEDIA :

OPPORTUNITÉS CONCEPTUELLES, ACADÉMIQUES ET PATRIMONIALES POUR LE CINÉMA PÉDAGOGIQUE ET SCIENTIFIQUE

Henri Hudrisier

La lecture savante des matériaux audiovisuels tend à se généraliser pour les "audiovisualistes" (ceux qui depuis un demi-siècle s'intéressent aux sources audiovisuelles pour l'enseignement ou la recherche) comme pour les "non-audiovisualistes" (des littéraires, des sociologues, des historiens) que les catalogues bibliographiques multimédias comme celui de la BnF confrontent à la nécessité de prendre en compte les sources audiovisuelles au même degré d'importance scientifique que les sources textuelles. Cela nous incite bien sûr à faire évoluer les méthodes, tant pour les archivistes que pour les chercheurs, mais il est surtout remarquable que cette transformation participe d'une évolution structurelle et normative des NTIC, particulièrement celle de la couche normative de la description et de la requête audiovisuelle MPEG7. Celle-ci est elle-même conçue pour pouvoir interagir avec MPEG4, la couche de numérisation interactive du document ; fonctionnellement structurée elle permet de distinguer les différents composants de l'image et du son en autant de fichiers distincts sélectivement accessibles.

La progressive structuration de l'information numérisée

L'informatique, puis le multimédia et l'Internet ont considérablement fait évoluer nos modes de conservation et nos potentiels d'accès aux patrimoines audiovisuels. En moins de 20 ans nous avons connu deux phases successives (l'informatisation des références, puis l'édition multimédia et interactive des ressources tant sur supports - cédéroms, DVD - que sur réseaux). Aujourd'hui nous en abordons une troisième qui nous concerne particulièrement, comme enseignant et comme chercheur, celle des langages à balises et de l'information structurée.

Cette troisième phase avait débuté de façon confidentielle (militaro-industrielle) il y a une quinzaine d'années avec SGML, (norme générique des langages qui lui font suite TEI, HyTime, HTML, XML, MPEG7...). La TEI (Text Encoding Initiative) en avait été la première grande application savante : à partir de l'étude des textes fondée sur l'information structurée. HTML avait permis le développement du Web, ce qui avait popularisé la notion d'information structurée normalisée ainsi surtout que la notion très spectaculaire (et conceptuellement très puissante) d'hyperlien activable au niveau planétaire. Il faut cependant attendre XML (né des efforts conjugués de chercheurs informaticiens qui avaient conçus la TEI, HyTime ou l'Internet et le Web) pour disposer d'un contexte logiciel permettant de développer de façon commode (beaucoup plus simple qu'avec SGML) un cadre normatif multimédia permettant de traiter toutes les catégories d'informations et plus spécifiquement l'audiovisuel et le multimédia. Ainsi MPEG 7 fonde sa structure sur le langage XML.

Pour dire vite, le document audiovisuel organisé selon la norme MPEG7 sera structurellement lié à nombre de métadonnées qui aujourd'hui en sont normalement dissociées. Ainsi seront logiquement et informatiquement organisés le contenu numérique du synopsis, du scénario, du dialogues, du/des sous-titrage(s) éventuel(s), du découpage, des documents de production, des contenus textuels 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, 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).

Normes et sens

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 réseau, d'une numérisation des documents et leur préservation muséale comme œuvres 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, 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 œuvres, 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 mais aussi d'autres logiques normatives d'organisation qui rentrent en lice, notamment DAVIC (Digital Audio Visual Council).

Pour être complet il faut évoquer pour un futur plus lointain la couche MPEG12 qui pourrait être celle de l'organisation de l'échange de document audiovisuel et multimédia.

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 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.

Du bon usage des normes

HyTime, comme Hyper-(document) et Time (qui traduit le souci de pouvoir mettre en parallèle ou synchroniser plusieurs événements, plusieurs documents - ou fragments de documents, plusieurs logiques) est resté une norme très sophistiquée et très confidentielle mais elle a permis de développer des logiques informatiques normalisées indispensables pour l'information audiovisuelle nécessairement inféodée à un synchronisme et à des pluralités parallèles de significations.

La TEI nous intéresse dans notre propos (bien que son nom même suggère qu'elle a été prévue pour décrire et analyser des textes), parce que la communauté TEI a été le creuset de programmes de recherches très approfondis sur la " bibliothèque virtuelle " et la " lecture savante assistée par ordinateur ".

On peut notamment insister sur les caractéristiques propres à ces médias : plus grande diversité potentielle de modes, très large bande passante et très gros volume d'informations pour un seul document, hétérogénéité des catégories sémiotiques et en corrolaire, diversité des méthodologies d'approche. 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 sémiotiquement, structurellement et référentiellement (données catalographiques) dans le cadre de la post-production numérique du document audiovisuel.

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.

MPEG 4, dans son ciblage d'usage, est certes fortement liée 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 destinées 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). Pour ce qui est de l'aspect sémantique des documents, MPEG présente des redondances entre couches (2, 4, 7...) qui permettent d'assurer leur prise en compte par les utilisateurs au fur et à mesure de leur disponibilité industrielle. Ainsi, l'aspect juridique et commercial (copyright, droits patrimoniaux, droits d'auteurs, rétributions d'usage) est structurellement présent dans les couches MPEG2, 4 puis 7. Il est formellement prévu dans la description globale de MPEG que la couche MPEG12 sera celle de la normalisation de

l'échange audiovisuel et multimédia. L'ISO recommande cependant d'attendre prudemment 3 à 4 ans pour proposer l'organisation de commissions de normalisation en la matière. Le commerce électronique et les flux interbancaires et monétiques amorcent seulement leur processus de description ouverte mais structurelle et normative - par exemple OTP (Internet Open Trading Protocol), document de travail à vocation pré-normative, fondé sur XML bien sûr. A terme de 3 à 4 ans, la convergence des médias n'exigera peut-être pas une norme d'échange spécifiquement spécialisée pour l'audiovisuel et le multimédia. Il est cependant presque certain que les savoir-faire de la gestion de l'échange d'audiovisuel (marchand ou non marchand, peu importe) seront structurellement disponibles sur des couches logicielles normalisées sous l'appellation de MPEG12 ou sous le nom d'une " norme d'échange " plus générique.

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 par tout un chacun qu'ils sont complexes (et d'ailleurs extrêmement normalisés) dans leur détail de construction technique et de fonctionnement.