

OPALES

L'INDEXATION DE DOCUMENTS TEXTUELS ET AUDIOVISUELS QUELQUES PROBLEMES ET BESOINS

ID document	OPALES-WP1-MSH-RT-301000.doc
Auteur(s)	Peter Stockinger – MSH
Date de la réunion	30 octobre 2000
Distribution	<input checked="" type="checkbox"/> Public <input type="checkbox"/> Restreint <input type="checkbox"/> Interne
Description	Note sur l'indexation de documents audiovisuels

1. LA PROBLEMATIQUE

Un document audiovisuel ou textuel peut servir ou répondre à des besoins/objectifs très variés et donc être objet de « lectures » multiples, pour reprendre une expression de Barthes. Et cela, même à l'intérieur d'une communauté de chercheurs (chercheurs-enseignants).

Par exemple, un chercheur :

- Peut avoir besoin (envie) d'exploiter uniquement un sujet particulier développé au sein d'un document audiovisuel. Par exemple, dans un documentaire ethnographique sur une ethnie, il peut avoir besoin des informations relatives aux rites de passage ou aux techniques de tissage, etc.
- Peut avoir besoin (envie) de comprendre/de s'exprimer sur la valeur (l'authenticité, le « sérieux », ...) de l'information fournie dans un document audiovisuel. Par exemple, en consultant un documentaire audiovisuel, il peut se demander si les informations données sont vraiment sérieuses, mises à jour, etc. (et il peut vouloir s'exprimer à ce sujet).
- Peut avoir besoin (envie) de comprendre le point de vue, la vision qui détermine (semble déterminer) le traitement d'un sujet dans un document audiovisuel. C'est, par exemple, le cas de l'anthropologie visuelle mais aussi de l'iconographie, de restituer à partir d'indices multiples (plastiques, compositionnels, iconiques, rythmiques, ...) une certaine vision du monde, une certaine idéologie, voir encore des motifs ou lieux communs. Par ailleurs, une interrogation similaire porte souvent sur « l'école », le « mouvement » (artistique, politique, ...) dans lequel s'inscrit tel ou tel document audiovisuel ou textuel (cf. « cinéma direct », « cinéma vérité », ..., « anthropologie participative », ...).
- Peut avoir besoin de mieux comprendre le traitement discursif d'une information dans un document audiovisuel : organisation en scènes et séquences ; manifestation d'une information dans telle ou telle autre scène ou séquence ; procédures d'introduction d'une information ; condensation/expansion d'une information ; information explicite vs information implicite ou présupposée ; etc.

- Peut avoir besoin d'analyser un document audiovisuel plan par plan en tenant compte surtout des indices visuels et sonores (musique, bruitage, parole) – indices visuels tels que type de plan, cadre, prise de vue, mouvement de caméra, etc.
- Etc.

1.1 Différentes formes d'indexation

Ces exemples montrent que **l'indexation du «contenu»** (au sens de «sujet» ou «thème») n'est qu'un cas particulier. Par ailleurs, elle-même peut varier très fortement d'un chercheur à un autre par rapport au :

- Degré de spécialisation (cherché par le chercheur ou imposé par le document lui-même)
- Niveau de précision (uniquement au niveau global du document, au niveau des séquences/scènes ; au niveau des plans)
- Points de vue théoriques

Il faudrait disposer de plusieurs canevas d'indexations qui permettrait à chacun de tenir compte d'un aspect particulier du «traitement» d'un document audiovisuel ou textuel par le chercheur - canevas d'indexation :

1. du sujet (du «contenu»)
2. du contexte (de production, d'exploitation, ...)
3. de la valeur du document
4. de l'organisation syntagmatique interne (i.e. (objet) - plan – scène – séquence)
5. du genre (modèle typique)
6. du traitement discursif de l'information audiovisuelle (degré d'expansion thématique ; explicite/implicite ; montage ; cadrage ; prise de vue et focus ; ...)

En ce qui concerne l'élaboration d'index, je crois qu'il faut bien distinguer entre une optique «documentaliste-bibliothécaire» et une optique «chercheur» dans le «traitement» d'un document audiovisuel. Un documentaliste est certainement plus enclin à utiliser des index ou thesaurus très élaborés, cela n'est pas le cas du chercheur. La raison est simple : le chercheur ne veut pas exclusivement classer un document, il cherche surtout à l'exploiter dans le cadre de ses recherches ou en vue d'une publication, d'un colloque, d'un enseignement, etc.

Conséquence (notamment en ce qui concerne l'indexation du « sujet », i.e. l'indexation thématique) : mieux que de proposer des vocabulaires (descripteurs, ...) très poussés, il faudrait donner au chercheur (mais aussi au documentaliste) la possibilité de construire son propre index.

2. LES BESOINS

A mon avis, cela signifie qu'il faudrait définir une « *base minimale* », un « *canon* » pour chacun des canevas identifiés ci-dessus et laisser à l'utilisateur la liberté de les adapter à ses besoins.

Les référents du contenu d'un documentaire (ethnographique, ...) sont typiquement des types récurrents d'objets tels que :

- Actions
- Participants
- Instruments
- Lieux
- Moments,
- Intentions/visées
- Ressources,
- Résultats
- Données préalables
- Etc.

Chacun de ces objets peut être caractérisés par des attributs, se spécialiser en des objets plus spécifiques et se combiner avec d'autres objets afin de constituer des sortes de « scènes » ou configurations sociales stéréotypées.

Mais, surtout, ces objets peuvent être utilisés pour spécifier la base d'un index approprié à une thématique de recherche. Pour cela, il faut pouvoir, par exemple, renommer les termes ci-dessus, etc.

Je pense qu'il faudrait donner au chercheur (documentaliste, ...) la possibilité :

- De se créer ses propres « scènes »
- De renommer les objets ci-dessus dans son propre langage
- De spécialiser les objets de la base
- D'introduire de nouveaux objets dans la base.

Ainsi, le chercheur, pourra schématiser (« indexer » ; « décrire ») à son gré le contenu d'un document audiovisuel – soit au niveau global (du document lui-même), soit au niveau local (de la séquence, de la scène, voir du plan).

La définition d'une base minimale ou canonique est tout aussi possible pour les autres canevas (indexation contextuelle, discursive, selon le genre, de la valeur d'information, ...).

3. L'ANNOTATION

Enfin, il semble que le problème de l'annotation ou, plutôt, de l'organisation et de la structuration de l'annotation peut être traité d'une manière similaire à l'indexation :

- annotation selon différents points de vue que l'on peut adopter par rapport au document (cf. les 6 « approches ») : thématique, genre, discours, valeur, contexte, ...
- annotation au niveau global (du document), de la séquence, de la scène, du plan, de l'objet proprement dit ou de l'une de ses facettes
- annotation sous forme de :
 - définition,
 - exemplification
 - lien hypertextuel
 - etc.