

# Une cinquième dimension pour les documents multimédia: La dimension annotation

Azze-Eddine Maredj\*, Madjid Sadallah et Lamia Hamouche

*Centre de recherche sur l'information scientifique et technique (CERIST), Alger, Algérie*

---

## Abstract

Un document multimédia intègre des entités de base de nature statique (texte, images, graphiques et tableaux) et de nature dynamique (vidéo, sons et animations), qui suivent une certaine organisation temporelle et spatiale. De nos jours, ces documents sont utilisés dans divers domaines d'application comme l'apprentissage à distance, la télémédecine, les visites virtuelles, la publicité, etc. Ces documents sont généralement modélisés par quatre dimensions: la dimension logique, la dimension temporelle, la dimension spatiale et la dimension hypermédia. Avec le temps, de nouveaux besoins sont apparus dans l'utilisation des documents multimédia, nous citons, l'adaptation des documents multimédias (où le document multimédia est présenté selon un profil utilisateur), la composition de document multimédia (qui permet de concevoir un document suite à une requête utilisateur), la recomposition de document multimédia (qui permet à un auteur de reconcevoir son document suite à l'analyse des traces de lecture), la génération de résumé automatique, etc. Les informations contenues dans ces quatre dimensions se révèlent insuffisantes pour répondre à ces besoins. D'autres informations sont donc nécessaires.

Dans ce travail, nous proposons d'étendre ce modèle à cinq dimensions, en définissant une nouvelle dimension: la dimension annotation. Cette dernière comportera des informations sur le document, ses entités, son auteur et ses lecteurs, qui contribueront à la prise en charge des applications précitées.

*Keywords:* Annotation ; Document multimédia ; Modèle de document multimédia ;

---

\* Corresponding author. Azze-Eddine Maredj. Tel. :+0-661-213-925-659  
E-mail address: [amaredj@cerist.dz](mailto:amaredj@cerist.dz)

## 1. Introduction

Un document multimédia est un document composé de médias de différents types (texte, image, vidéo, animation, audio) et dont la présentation comporte des composantes temporelle, spatiale et hypermédia. À ce jour, les documents multimédias sont décrits par un modèle à base de quatre dimensions:

- la dimension logique qui décrit le regroupement des médias sous forme d'entités,
- la dimension temporelle qui décrit le scénario temporel du document, c'est-à-dire l'ordonnancement des entités dans le temps,
- la dimension spatiale qui décrit l'emplacement spatial des entités visibles sur le support de présentation,
- et la dimension hypermédia qui décrit les liens de navigation des entités.

Avec la prolifération d'applications exploitant les documents multimédia tels que l'apprentissage à distance, la télémédecine, les visites virtuelles etc., de nouveaux besoins sont apparus tels que:

- l'adaptation des documents multimédias où le document multimédia est présenté selon un profil lecteur,
- la composition de document multimédia qui permet de concevoir un document suite à une requête et un profil lecteur,
- la recomposition de document multimédia qui permet à un auteur de reconcevoir son document suite à l'analyse des traces de lecture,
- la génération de résumé automatique, etc.

L'analyse des besoins suscités montre que leur prise en charge se traduit globalement par l'exécution de trois opérations. (1) l'opération de *suppression*, par exemple lors de la génération automatique d'un résumé, nous devons connaître le degré d'importance de chaque entité du document afin de ne pas supprimer les plus importantes. (2) l'opération d'*insertion*, par exemple pour mieux expliquer un concept non maîtrisé par le lecteur, on ajoute des liens vers d'autres entités contenant plus de détails. (3) l'opération de *remplacement* d'entités pour un besoin de personnalisation, par exemple pour répondre au style d'apprentissage d'un lecteur, on remplace un texte par une image ou inversement.

Les informations contenues dans le modèle actuel à quatre dimensions sont insuffisantes pour répondre à ces besoins et des informations supplémentaires sont donc nécessaires.

Afin de pallier à ces limites, nous proposons, dans ce travail, d'étendre le modèle de quatre à cinq dimensions en y ajoutant une nouvelle dimension: la dimension annotation. Elle permettra de définir et de gérer, les informations nécessaires (sur le document, ses entités, son auteur et ses lecteurs) pour la prise en charge des applications précitées.

## 2. Contexte

### 2.1 Modèles de documents multimédia

#### 2.1.1 Modèles de référence

Un modèle de document fournit une description de l'organisation des données contenues dans ce document (Whitehead, 2002) et lui définit un cadre applicatif général. Ainsi, il précise la nature de données manipulées, la méthode de stockage et de mise en relation des éléments et comment le document peut être utilisé (Richard, 2002). Pour les documents hypermédia et multimédia, la plupart des modèles se basent sur les modèles de référence de Dexter (Halasz et al., 1994) et d'Amsterdam (AHM) (Hardman et al., 1994) qui ont permis d'asseoir les bases pour la définition de la structure de l'hypertexte, de l'hypermédia et du multimédia.

En termes de standards et de langage, l'évolution basée sur ces modèles allant de HyTime (ISO Standard, 1997), MHEG (Meyer-Boudnik et al., 2008 ) au langage et standard SMIL (Botterman, 2008) a permis de stabiliser la définition des formats, les cadres logiciels et les interfaces applicatives des documents multimédia tout en incorporant la conceptualisation des médias et leurs techniques de composition.

### *2.1.2 Placement et formatage spatial*

Deux familles de modèles spatiaux peuvent être distinguées: les modèles absolus et les modèles relatifs (Boll, 1999). Dans un modèle absolu, la description de l'espace et de l'emplacement des médias est faite en termes de coordonnées à deux ou trois axes par rapport à un référentiel spatial. Les modèles spatiaux relatifs, quant à eux, utilisent des formes continues et le placement se fait en spécifiant des relations spatiales entre les médias

### *2.1.3 Temporalisation et synchronisation*

Les approches de spécification de la synchronisation peuvent être regroupées en trois grandes familles (Maredj, 2009): les approches opérationnelles à base de programmation événementielle, les approches opérationnelles à base de structures d'arbres ou de graphes et les approches à base de contraintes d'intervalles ou d'instant. La définition du temps des documents multimédia est classifiée dans (Boll, 1999) en modèle à instants, modèles événementiels, modèles à intervalles et modèles à base de scripts.

## *2.2 Annotations*

### *2.2.1 Définition*

Selon Bringay et al. (Bringay et al., 2005) une annotation est une note particulière attachée à une cible. La cible peut être une collection de documents, un document, un segment de document (mot, groupe de mots, paragraphe, image ou partie d'image, etc.) ou une autre annotation. À une annotation correspond un contenu, matérialisé par une inscription qui est une trace de la représentation mentale que l'annotateur se fait de la cible. Nous appelons ancre ce qui lie l'annotation à la cible (un trait, un entouré, etc.).

### *2.2.2 Techniques d'annotation*

Deux techniques d'annotations non exclusives et souvent entrelacées peuvent être distinguées (Gayoso-Cabada, et al. 2018) (Zhai et al. 2005): l'association des métadonnées et l'enrichissement du contenu.

L'association de métadonnées utilise des modèles spécifiques pour la construction de structures sémantiques permettant des opérations diverses telles que la recherche dans le contenu. Cette approche requiert la compréhension du modèle d'annotation sémantique pour pouvoir associer aux média des métadonnées conformément au modèle.

La deuxième méthode, enrichissement de contenu, utilise d'autres éléments multimédia (images, graphiques, texte, audio, vidéo,...) pour enrichir les objets dans un flux audiovisuel générant un nouveau flux composé. L'utilisateur insère ses annotations interactivement en associant du texte, des images, des clips, etc., de sorte que le flux initial soit enrichi par un nouveau contenu.

### 3. La dimension Annotation pour un document multimédia

#### 3.1 Modèle de document multimédia proposé

Le modèle que nous proposons est à base de dimensions, nous définissons, en plus des quatre dimensions classiques, une nouvelle dimension appelée "dimension annotation" (Figure 1).

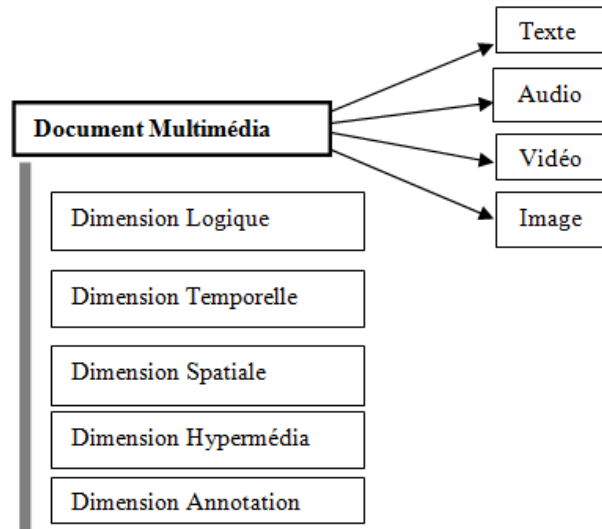


Fig. 1. Modèle de document multimédia enrichi avec la dimension Annotation

#### 3.1.1 La dimension logique

La dimension logique permet de décomposer un document multimédia en parties. Dans un processus de conception d'un document, l'auteur sélectionne les informations qu'il souhaite communiquer puis choisit comment il présentera l'information d'une façon cohérente. La dimension logique d'un document, doit donc refléter le point de vue de l'auteur.

#### 3.1.2 La dimension temporelle

Elle exprime l'enchaînement des entités dans le temps de manière à définir un scénario temporel cohérent. Dans la littérature, cette dimension est souvent représentée par un modèle à base de relations temporelles.

#### 3.1.3 La dimension spatiale

Elle décrit le placement spatial, spécifié par l'auteur du document, d'une entité sur un dispositif de présentation du document multimédia.

### 3.1.4 La dimension hypermédia

Elle exprime les liens de navigation entre les entités du document. Contrairement au document classique, lors de la présentation d'un document multimédia l'aspect temporel des médias (apparition et disparition d'un média, changement progressif de l'état d'un média, etc.) lui confère une certaine complexité de perception. De nouveaux besoins comme un retour sur un média particulier, un détail passé ou un saut vers un instant donné doivent être considérés pour maîtriser cette complexité et permettre par la même, une meilleure lecture du document. La réponse réside dans les capacités d'interaction offertes par le système.

### 3.1.5 La dimension annotation

À travers cette nouvelle dimension, nous allons permettre l'ajout d'informations sur le document, sur une partie du document, sur l'auteur et sur le lecteur. Ces informations peuvent être extraites automatiquement ou fournies manuellement. Nous intégrons cette dimension à la définition même du document multimédia de sorte à rehausser son importance et de permettre ainsi l'exploitation des informations qu'elle contient pour différents besoins applicatifs.

## 4. Modèle d'annotation proposé

Après l'analyse des besoins en termes d'informations nécessaires pour la prise en charge des applications mentionnées, nous proposons la structure d'annotation ci-dessous, représentée par la figure 2.

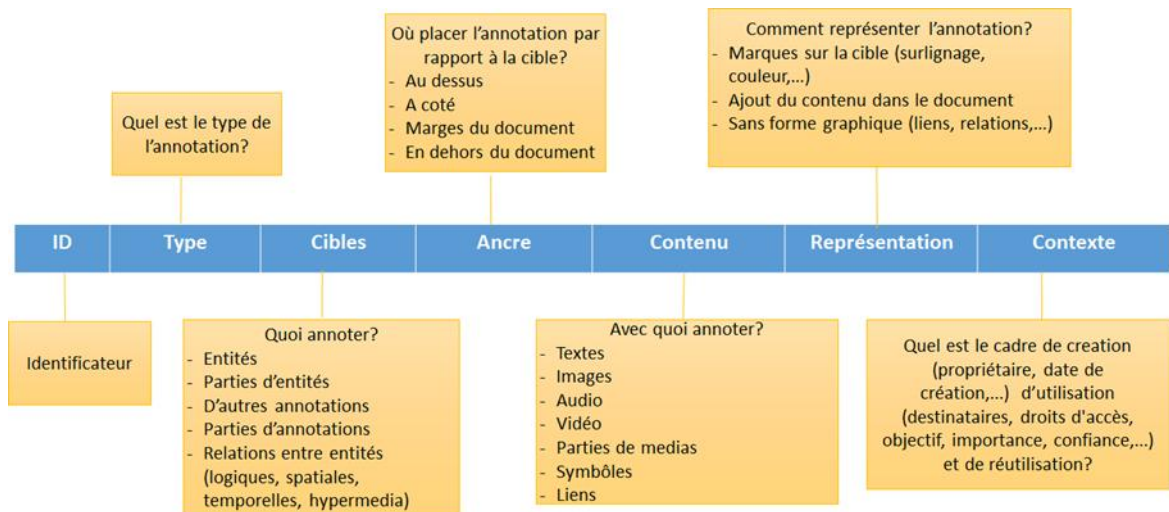


Fig. 2. Structure d'annotation

Nous caractérisons l'annotation par un ensemble d'attributs issus des interrogations suivantes:

- Quel est le type de l'annotation?
- Quelle est sa cible?
- Où la placer par rapport à la cible?
- Avec quoi l'annoter?

- Quelle est sa représentation visuelle?
- Quel est son contexte de création?

#### 4.1 Types d'annotation

À ce niveau, nous proposons un ensemble de types d'annotations que nous regroupons en quatre classes:

- 1) Données générales: Cette classe est réservée à l'auteur pour décrire le contexte de création, d'utilisation et de réutilisation du contenu de l'élément annoté. Elle est composée des types suivants:
  - Métadonnées:
    - Auteur: nom, prénom, email et ses domaines de compétences.
    - Date de création et date de la dernière mise à jour de l'élément annoté.
    - Une synthèse du contenu de l'élément annoté.
  - Indexation: mots clés caractérisant de l'élément annoté.
  - Pondération: calcul/affectation d'une valeur d'importance de l'élément annoté par rapport au reste des éléments.
- 2) Accessibilité: c'est un ensemble d'informations requis à la compréhension du contenu de l'élément annoté, on a retenu les types suivants:
  - pré-requis (langue et connaissances).
  - niveau de formation requis.
  - niveau de connaissance requis (débutant, intermédiaire, avancé, expert.).
- 3) Contexte pédagogique: composé des types suivants:
  - Discipline du document (informatique, mathématique, etc.).
  - Domaine (langage de programmation, base de données, etc.).
  - Sous domaine.
  - Thématique.
- 4) Objectifs pédagogiques: pour indiquer s'il s'agit d'un acquis d'un savoir ou d'un savoir-faire.
  - Connaissances.
  - Aptitudes.

#### 4.2 Annotations de commentaires

Elles sont généralement réservées au lecteur, nous en avons défini trois types d'annotations:

- 1) Commentaire: donner un avis sur le contenu de l'élément annoté.
- 2) Discussion: débattre de son contenu, poster des question/réponses.
- 3) Révision: suggérer des modifications de son contenu, nous en avons défini:
  - La correction: erreurs de langue (orthographe, grammaire, etc.).
  - La mise à jour: reformulation et évolution.
  - La restructuration: réorganisation de l'élément annoté.

#### 4.3 Annotations d'enrichissement

Cette classe permet d'ajouter du contenu à l'élément annoté.

- 1) Enrichissement: documenter le contenu de l'élément annoté.
- 2) Définition de liens et de relations:
  - Liens de navigation.
  - Relations sémantiques (par exemple en contradiction, démonstration, antithèse, etc.).

#### 4.4 Annotations d'appréciation

- 1) Niveau de compréhension: attribution d'une valeur/mention du degré de compréhension de l'élément annoté.
- 2) Niveau de difficulté: attribution d'une valeur/mention du degré de difficulté de l'élément annoté.
- 3) Niveau d'intérêt: attribution d'une valeur/mention du degré d'intérêt de l'élément annoté.

#### 4.5 Annotations de mise en forme

Cette classe permet de donner une forme à l'élément annoté:

- Marque (forme et couleur): entourage, surlignage, soulignage, surbrillance, barrer, rayer, etc.
- Effet (modification de la présentation): typographie, taille, couleur, etc.

#### 4.6 Cible d'une annotation

La cible de l'annotation désigne l'élément annoté sur lequel porte l'annotation. Cet élément peut être:

- Le document.
- Une ou plusieurs entités du document.
- Une ou plusieurs annotations.

#### 4.7 Contenu d'une annotation

Peut-être un contenu textuel, une image, une audio, une vidéo, un dessin ou conventionnel (smiley ou symbole). Une annotation peut ne pas avoir de contenu, on utilise dans ce cas le soulignage, le surlignage ou la mise en surbrillance, par exemple.

#### 4.8 Représentation visuelle d'une annotation

Elle dépend du type de l'annotation et peut prendre diverses formes:

- 1) Marquage sur le contenu:
  - Mise en forme/évidence: surlignage, soulignage, entourage, barrage, typographie, alignement, changement de couleur, etc.
  - Symbole et icônes: croix, cercle, flèche (binaire), étoiles (ratings), points d'interrogations, etc.

- 2) Ajout de contenu: texte et multimédia.
- 3) Sans forme: par exemple pour les liens et les relations. Alternativement, des conventions de matérialisation peuvent être élaborées pour rendre compte de l'annotation.

#### 4.9 Ancre d'une annotation

Elle peut être à l'intérieur du document ou en dehors du document.

#### 4.10 Contexte d'une annotation

Il décrit le cadre de création, d'utilisation et de réutilisation d'une annotation. Il contient les données suivantes:

- Identification de l'auteur de l'annotation,
- Date de création de l'annotation,
- Date de modification de l'annotation,
- Date d'expiration de l'annotation,
- Sphère : privé, groupe ou public.
- Droits d'accès,
- Portée: renseigne si l'annotation peut être réutilisée dans un autre document contenant la même cible.

## 5. Conclusion

La production et l'exploitation des documents multimédia requièrent l'élaboration d'un modèle de données qui décrit les entités constituant le document et les relations entre elles. Un tel modèle permet de préciser les règles selon lesquelles les informations sont organisées dans l'espace informationnel et les mécanismes et paradigmes permettant les diverses manipulations des documents s'y conformant. Néanmoins les modèles existant n'arrivent pas à couvrir tous les besoins que font émerger les nouveaux contextes applicatifs.

Ce travail a pour but d'enrichir, par des annotations, le modèle de document multimédia afin de prendre en charge les nouveaux besoins. Par ailleurs, les annotations proposées restent insuffisantes pour appréhender l'information sémantique (qu'elle soit locale ou globale). Disposer de l'information sémantique permet de comprendre la raison d'existence de chaque entité du document et les relations qui existent entre-elles. Afin d'en disposer, une des perspectives est d'ajouter à la dimension logique des relations sémantiques, la théorie RST (Bill Mann, 1999) s'avère dans ce contexte une option intéressante car permettant de définir d'une façon déclarative les intentions ou les buts de l'auteur en termes d'effets sur les croyances du lecteur.

## Références

- Whitehead, 2002 : Uniform comparison of data models using containment modeling. In Proceedings of the thirteenth ACM conference on Hypertext and hypermedia, ACM Hypertext '02, pages 182–191, New York, NY, USA.
- Bertrand Richard, 2002. Lecture active de documents audiovisuels : de la gestion de connaissances personnelles au soutien de l'attention du lecteur actif. Thèse de doctorat en informatique, INSA de Lyon.
- Frank Halasz, Mayer Schwartz, 1994. The dexter hypertext reference model. Commun. ACM, 37(2) :30–39.
- Lynda Hardman, Dick C. A. Bulterman, Guido van Rossum, 1994. The Amsterdam hypermedia model : adding time and context to the dexter model. Commun. ACM, 37(2) :50–62,



- International Organization for Standardization, 1997. Information technology –hypermedia/time-based structuring language (hytime). ISO Standard.
- Thomas Meyer-Boudnik, Wolfgang Effelsberg, 2008. Mheg explained. *IEEE MultiMedia*, 2:26–38.
- Dick Botterman : Synchronized multimedia integration language (SMIL 3.0). W3C recommendation, W3C, <http://www.w3.org/TR/SMIL3/>.
- Susanne Boll, Wolfgang Klas et Utz Westermann, 1999. A comparison of multimedia document models concerning advanced requirements. Rapport technique.
- Azze-Eddine Maredj, 2009. Spécification et validation des relations temporelles et spatiales dans un document multimédia interactif. Thèse de doctorat en informatique, Université des Sciences et de la Technologie Houari Boumediène, Faculté d'Electronique et d'Informatique.
- Sandra Bringay, Catherine Barry, Jean Charlet, 2005. Les annotations pour gérer les connaissances du dossier patient. IC - 16èmes Journées francophones d'Ingénierie des Connaissances, May 2005, Nice, France. pp.73-84. fhal-01023761f
- Gang Zhai, Geoffrey C. Fox, Marlon Pierce, Wenjun Wu et Hasan Bulut, 2005. eSports : Collaborative and synchronous video annotation system in grid computing environment. In *ISM '05 : Proceedings of the Seventh IEEE International Symposium on Multimedia*, pages 95–103, Washington, DC, USA. IEEE Computer Society.
- Joaquín Gayoso-Cabada, Antonio Sarasa Cabezuelo, José-Luis Sierra, 2018. Document annotation tools: annotation classification mechanisms. 9th International Workshop on Software Engineering for E-learning (ISELEAR'18).
- Bill Mann, 1999. An introduction to rhetorical structure theory (RST).