
Description des documents audiovisuels : s'affranchir des limitations de MPEG-7

FDL, vers un langage de description par objets, extensible et modulaire

Marc Caillet^{*,**} — Jean Carrive^{*} — Cécile Roisin^{**}

^{*}*Institut National de l'Audiovisuel (INA)
Thème de recherche Description des Contenus Audiovisuels
4 avenue de l'Europe, 94366 Bry-sur-Marne Cedex, France
{mcaillet,jcarrive}@ina.fr*

^{**}*INRIA Rhône-Alpes - Équipe WAM
655 avenue de l'Europe, Montbonnot 38334 St-Ismier Cedex, France
{marc.caillet,Cecile.Roisin}@inrialpes.fr*

RÉSUMÉ. L'Institut National de l'Audiovisuel (INA) est concerné au premier chef par les conséquences de la numérisation des contenus audiovisuels. Parmi celles-ci, l'émergence de nouveaux modes d'accès à ces contenus définit une classe d'applications multimédias nouvelles, centrées sur la description des contenus. Cet article confronte le langage normalisé de description de documents multimédias MPEG-7 à nos besoins en termes d'expressivité. Concluant qu'il ne les satisfait pas tous, nous proposons notre propre langage, FDL, en alternative.

ABSTRACT. The National Institute of Audiovisual (INA) has to face the consequences of audiovisual contents digitalization. Among those consequences, new access modes to audiovisual content are defining a class of new multimedia applications, focused on content description. This article evaluates the standard multimedia description language MPEG-7 as an applicant to fulfill our needs of expressiveness. Coming to the conclusion that it's not fulfilling them all, we propose our own language, FDL, as an alternative.

MOTS-CLÉS : langage de description des documents audiovisuels et multimédias, MPEG-7, DDL, XML Schema, FDL, langage orienté objets, FERIA.

KEYWORDS: multimedia description language, MPEG-7, DDL, XML Schema, FDL, object oriented language, FERIA.

1. Introduction

La numérisation des contenus audiovisuels a non seulement pour conséquence l'évolution des pratiques professionnelles autour de leur production, elle fait également émerger de nouveaux usages en termes d'accès à ces contenus. Parmi ceux qui concernent au premier chef les détenteurs de fonds audiovisuels – et plus généralement de fonds multimédias, l'Institut National de l'Audiovisuel (INA) en particulier – nous pouvons citer : l'accès du grand public à leur offre commerciale, à différents niveaux de granularité (*e.g.* une collection d'émissions, une émission en particulier, un extrait d'une émission...); la publication multisupport — apparue comme un besoin urgent il y a de cela quelques années et dont les solutions restent à inventer — qui permettrait de produire un document multimédia à destination de plusieurs plateformes, en tenant compte de leur capacités interactives respectives, sans qu'il soit pour cela nécessaire de mettre en œuvre un coûteux processus de production pour chacune d'elles. Ces usages émergents ne sont pas figés : ils s'inscrivent aujourd'hui encore dans une dynamique exploratoire, avec pour conséquence la difficulté d'optimiser de façon pérenne une application pour un usage donné. Il apparaît ainsi indispensable de disposer d'une technologie d'organisation et de production de contenus audiovisuels suffisamment souple, capable d'accompagner, voire même de susciter, l'évolution des usages. C'est l'objet du projet FERIA ([BRU 05]).

Autre conséquence de la numérisation : la virtualisation de la manipulation des contenus audiovisuels. Manipuler virtuellement des contenus est une pratique courante dans le domaine de la production audiovisuelle. Un exemple de l'intérêt de cette opération est la possibilité d'effectuer de multiples essais de montage, tout en minimisant le nombre d'opérations sur la pellicule. Nous proposons d'étendre ce procédé de manipulation virtuelle à la classe des applications multimédias nouvelles, dont deux représentants ont été évoqués dans le paragraphe précédent, et qui reposent sur d'autres applications parmi lesquelles la navigation multimodale dans une collection de contenus, la recherche d'information dans ces contenus, ou encore l'indexation automatique. Cette orientation a pour conséquence de centrer ces applications sur la notion de description.

Cette orientation fait aussi surgir la nécessité de disposer d'un langage permettant de décrire ces contenus, langage qui, entre autres exigences, doit être doté de la capacité d'extensibilité, de façon à pouvoir répondre à tout nouveau besoin. Cette exigence forte, ainsi que les autres (elles sont décrites dans la section suivante), nous les retrouvons en grande partie dans le cahier des charges du langage de description normalisé MPEG-7 ([MAR 02a] et [MAR 02b]). MPEG-7 se trouvait donc, *a priori*, en position de candidat naturel pour remplir le rôle de langage de description qui répondrait aux besoins que nous venons d'exprimer. Nous montrons cependant pourquoi MPEG-7 échoue à répondre à nos exigences (section 3). Nous décrivons ensuite (section 4) notre proposition d'un langage de description des documents audiovisuels et multimédias : FDL (pour *Feria Description Language*), puis rendons compte (section 5) de sa mise à l'épreuve dans le cadre du projet FERIA. Enfin, nous concluons et discutons des perspectives d'évolution de FDL (section 6).

2. Exigences pour un langage de description des documents audiovisuels et multimédias

Nos propres exigences pour un langage de description des documents audiovisuels et multimédias recouvrent peu ou prou celles qui ont présidé à la définition de la norme MPEG-7 ([PER 01]), pour la plupart tout à fait pertinentes de notre point de vue :

- **Schémas de description et descriptions** : le langage doit permettre l’expression de schémas de description et de descriptions. Un schéma de description définit une classe de descripteurs. Une description est contrôlée par un schéma de description et instancie la classe de descripteurs qui y est définie.

- **Hierarchie de descripteurs** : les descripteurs doivent être organisés en taxonomie selon, donc, la relation de subsomption *is-a*¹. Cette hiérarchisation taxonomique fournira ainsi un moyen d’expression de la sémantique des descripteurs.

- **Extensibilité** : le langage doit fournir un mécanisme de définition de nouvelles classes de descripteurs.

- **Modularité** : le langage doit être modulaire de façon à permettre réutilisation aisée et exploitation partielle de schémas de description et de descripteurs.

- **Composition** : les descripteurs peuvent être composés entre eux de manière temporelle, spatiale, structurelle ou conceptuelle.

- **Lien au média et indépendance par rapport au média** : les descripteurs peuvent référencer les contenus qu’ils décrivent, ce lien devant se faire indépendamment du codage et de l’emplacement physique des médias.

- **Identification** : le langage doit comporter un mécanisme d’identification unique des médias, schémas de description et description.

- **Types de base** : le langage doit fournir les types de base couramment utilisés, ainsi que des types additionnels permettant la localisation temporelle et spatiale.

- **Syntaxe XML** : le langage doit être muni d’une syntaxe XML, pour les schémas de description et descriptions, de façon à faciliter leur échange et leur stockage, ainsi que pour la prise en charge d’une part non négligeable de leur validation syntaxique par des outils XML existants.

3. MPEG-7 face à nos exigences

Nous montrons dans cette section pourquoi MPEG-7 échoue à satisfaire à certaines de ces exigences, et notamment celles de hiérarchisation taxonomique des descripteurs, d’extensibilité et de modularité. Le principal responsable, selon notre analyse : le choix de XML Schema pour langage de définition de descriptions (DDL, *Description Definition Language*), dans lequel sont exprimés schémas de description (DS) et descriptions ([ISO01b]).

1. *is a hyponym of*.

3.1. À la recherche d'un langage de définition de descriptions

À Lancaster, en février 1999, se tient une commission chargée d'évaluer les différentes propositions de DDL. [HUN 00] propose une synthèse de l'évaluation de chacune des propositions ; ce papier décrit plus en détail la proposition retenue pour servir de base au futur DDL. Un consensus apparaît clairement quant à la syntaxe XML du DDL de MPEG-7 et quant à la nécessité de proposer, d'une part, des procédés de validation structurelle et de types, d'autre part, une modélisation orientée objet pour l'expression de la sémantique des descripteurs. Ce dernier point se révèle être critique : il est remarquable qu'aucune des propositions ne parvienne à concilier tout à fait ces deux aspects, validation et expression de la sémantique ; il est cependant clair, dès cette époque, que le DDL doit fournir un mécanisme d'héritage. L'une des propositions est alors retenue ([HUN 99]) pour servir de base à des développements collectifs, guidés par le but de résoudre ce point critique.

Ces travaux complémentaires aboutissent finalement au choix de XML Schema² ([THO 04] et [BIR 04]) pour DDL. Ce choix paraît, en première analyse, parfaitement fondé. XML Schema vient alors tout juste d'être normalisé (avec la contribution des membres du consortium MPEG-7) et s'impose dès lors comme une évolution importante et nécessaire de DTD en termes d'expression de contraintes, structurelles et de types, de validation d'un document XML. De plus, XML Schema partage — mais de façon trop limitée, selon nous, dans le cadre de son emploi en tant que DDL de MPEG-7 — des caractéristiques communes avec les langages à objets.

XML Schema comme DDL de MPEG-7 apparaît ainsi comme un compromis entre les deux aspects validation et expression de la sémantique, compromis qui finalement, ne satisfait pas toutes nos exigences, ni même toutes celles de MPEG-7.

3.2. MPEG-7 et l'expression d'une hiérarchie taxonomique

3.2.1. Propriétés de la relation *is-a*

La relation *is-a*, véritable colonne vertébrale d'une hiérarchie taxonomique, possède, entre autres, deux propriétés essentielles³ :

- **Transitivité** : si la classe p est une descendante de la classe q et q une descendante de la classe r , alors p est aussi une descendante de r .
- **Inclusion de classes descendantes** : si p est une instance de la classe q et q une descendante de la classe r , alors p est aussi une instance de la classe r .

2. Augmenté de types relatifs à l'expression d'un intervalle temporel (`basicTimePoint` et `basicTimeDuration`) et de types `Array` et `Matrix` ([ISO01b]).

3. Elles peuvent être exprimées formellement, selon la *F-Logic*, abréviation de *Frame-Logic*, formalisme permettant de raisonner sur les objets et les langages à base de *frames* ; c.f. [KIF 95].

3.2.2. Dérivation de types en XML Schema

Un schéma XML permet de décrire la structure de documents XML d'une façon assez proche de celle des langages orientés objets : on y retrouve les notions d'héritage de type, de type abstrait ou encore de type terminal. Il est constitué de définitions de types, ainsi que de déclarations d'éléments et d'attributs. Il existe principalement deux catégories de types :

- **simple types** : ce sont des types de données qui peuvent être "affinés" (*e.g.* par restriction du domaine de définition, par énumération), mais ne peuvent ni dériver d'autres types, ni comporter d'attributs.

- **complex types** : au contraire des types simples, les types complexes peuvent comporter des attributs et dériver d'autres types.

L'opération de dérivation consiste à définir un type à partir de la définition d'un autre type. Elle n'est ainsi autorisée que pour les types : un élément ne peut dériver d'un autre élément⁴. Deux mécanismes permettent la dérivation de type :

- **extension** : ajoute de nouveaux éléments à la fin d'un type donné ;
- **restriction** : restreint le domaine de définition de certains éléments d'un type donné.

Les types ainsi définis sont liés entre eux par une relation de sous-typage, à laquelle il est souvent fait référence sous la désignation d'héritage. Parler ici d'héritage nécessite d'avoir une acceptation large de ce terme : il s'agit, dans le cas de la dérivation de types, de réutilisation de la définition d'un type et non de spécialisation entre éléments, comme le note [EUZ 03].

Aussi, dans le but de simuler la relation de descendance entre éléments, XML Schema fournit l'attribut optionnel `substitutionGroup`. Le pouvoir d'expression de la relation de sous-typage — pour les types — et de descendance — pour les éléments — est cependant assez limité : si la première respecte la propriété de transitivité de la relation `is-a`, la seconde enfreint la propriété d'inclusion de classes descendantes, l'attribut `substitutionGroup` ne s'appliquant pas systématiquement à tous les éléments d'un type donné, mais au cas par cas. De plus, [BRU 03] montre que ces relations ne permettent pas de modéliser des hiérarchies parallèles covariantes⁵, pourtant monnaie courante en modélisation objet. Enfin, [TRO 04a] propose un scénario face auquel MPEG-7 échoue en raison des limitations discutées ici.

Le choix de XML Schema pour DDL de MPEG-7 met le langage normalisé de description de contenu multimédia dans l'incapacité d'organiser les descripteurs selon une hiérarchie taxonomique.

4. On aurait pu envisager, par exemple, la possibilité de définir un élément par dérivation d'un autre élément, défini, lui, par un type implicite.

5. Bâties selon la technique de covariance : modification du type des arguments lors des redéfinitions [MEY 97]

3.3. MPEG-7 et l'extensibilité

MPEG-7 spécifie la définition d'un nombre considérable de descripteurs : en ce qui concerne les seuls Multimedia Description Schemes (MDS, [ISO01a]), il en existe plus de six cents. Cet ensemble est cependant loin d'être exhaustif — mais est-il seulement possible de définir un ensemble exhaustif de descripteurs ? — et ne permet pas de satisfaire tous les besoins, notamment ceux des outils d'analyse automatique. Il n'existe pas, par exemple, de descripteur pour exprimer les résultats d'un outil de transcription sous la forme de simples mots « timecodés ». Disposer de moyens d'étendre le langage est alors crucial.

Dans l'intention, MPEG-7 se veut extensible. Selon [PER 01] : « *The Description Definition Language (DDL) is a language that allows the creation of new Description Schemes and, possibly, Descriptors. It also allows the extension and modification of existing Description Schemes.* » [ISO01b] caractérise définitivement MPEG-7, par le truchement de son DDL, comme un langage extensible : « *[The DDL] specifies the language for defining the standard set of description tools (DSs, Ds, and datatypes) and for defining new description tools.* »

La section 3.2.2 montre clairement qu'il est possible de créer de nouveaux schémas de description (DS) par dérivation de types normalisés par MPEG-7 — à moins que ceux-ci ne soient marqués `final`. Les nouveaux schémas ainsi créés appartiennent-ils pour autant à la norme ? Non, semble-t-il : MPEG-7 ne prévoit aucun mécanisme spécifique qui permettrait de contrôler l'extension de ses DSs, reportant cette tâche sur son DDL. [ISO02] précise, concernant les *Multimedia Description Schemes*, que la validation des éléments et attributs définis en dehors de la norme n'est pas requise. Ainsi, si l'on peut affirmer sans risque que toute extension de MPEG-7 fait partie de l'ensemble des schémas exprimables en XML Schema, rien ne permet d'avancer qu'une extension d'un DS de MPEG-7 fait partie de MPEG-7. En particulier rien ne garantit que l'extension d'un descripteur audiovisuel est un descripteur audiovisuel.

Il est certes possible d'étendre un descripteur MPEG-7 grâce à son DDL. Le mécanisme d'extension est cependant trop laxiste, selon nous, car il rend hors-norme toute extension du langage. Notre niveau d'exigence vis-à-vis de la norme préconise que toute extension d'un descripteur de MPEG-7 soit un descripteur de MPEG-7. De ce point de vue, les possibilités d'extension de MPEG-7 sont nulles.

3.4. MPEG-7 et la modularité

MPEG-7 n'est pas modulaire : la validation d'une description nécessite l'importation de la totalité de la norme. Cette nécessité rend problématique le développement d'applications embarquées qui disposent de ressources très limitées, et pour lesquelles on pourrait imaginer n'intégrer que les DSs qui définissent les descripteurs utilisés dans l'application. [BRU 03] rapporte une tentative de modularisation de MPEG-7. Sans succès.

4. FDL : une alternative à MPEG-7

Ainsi MPEG-7 échoue à satisfaire à toutes nos exigences. Partant, le cadre applicatif de la norme est restreint. En particulier, [FAT 04] met en exergue son incapacité à exprimer différents points de vue sur un même document : la recherche de contenu audiovisuel, décrit par MPEG-7, nécessite alors de passer par une modélisation abstraite intermédiaire, en dehors de la norme. En cause, comme l'a montré la section précédente, l'un de ses composants fondamentaux : son DDL. Ce qui coupe court, de notre point de vue, à toute velléité de proposition d'une évolution de la norme. Cette section décrit, à partir des points d'achoppement de MPEG-7, notre proposition d'un langage de description des documents audiovisuels et multimédias, FDL⁶.

4.1. FDL : une modélisation objet et une syntaxe XML

Nos exigences nous ont conduits à faire reposer FDL sur une modélisation abstraite de type objet — et non sur une modélisation syntaxique — tout en le dotant d'une syntaxe XML, de façon à reporter une partie significative de la tâche de validation sur des outils existants et éprouvés, et à faciliter l'échange et le stockage de schémas de description et descriptions exprimés en FDL.

Afin de doter FDL de capacités d'organisation hiérarchique des descripteurs selon des relations ontologiques et mérologiques, de capacités d'extensibilité et de modularité, nous l'avons conçu sur la base d'un noyau le plus restreint possible qui définit les invariants du langage. Parmi ceux-ci :

- **un nombre très réduit de descripteurs génériques** : `Descriptor`, `TemporalDescriptor` (spécialise `Descriptor`) et `SpatialDescriptor` (spécialise `TemporalDescriptor`, il s'agit donc d'un descripteur spatio-temporel).

- **des structures de composition** : deux types de structures simples sont actuellement proposés par FDL : `ListStructure` permet d'aggréger des descripteurs de type, ou descendant de, `Descriptor` ; `TemporalStructure` des descripteurs temporels, de type, ou descendant de, `TemporalDescriptor`.

- **des mécanismes d'extension** : ils permettent de définir un nouveau descripteur à partir de descripteurs existants.

Ces invariants — et d'autres concernant la définition documentaire de collections de documents audiovisuels, le lien au média, la définition de types élémentaires, de types relatifs à la localisation spatiale et temporelle et de type matrice de dimension quelconque — permettent d'exprimer, avec une syntaxe XML, des :

- **Schémas de Description (DS)** : un DS est une classe qui définit un nouveau descripteur par : (1) héritage d'un descripteur générique de FDL ou d'un descripteur défini dans un autre DS, (2) composition de descripteurs.

6. *Feria Description Language*. Anciennement dénommé AVDL, une version initiale de ses principes est exposée dans [TRO 04a] et [TRO 04b].

– **Descriptions (D)** : une description est un objet, instance d’une classe définie par un DS.

Aussi FDL ne fige-t-il qu’un tout petit nombre de descripteurs d’un très haut niveau de généralité. Libre à l’utilisateur de définir — par extension contrôlée de descripteurs existants — les descripteurs dont il a besoin. Cette conception a notamment pour avantage de doter le langage de la capacité d’expression de points de vue spécifiques à un contexte applicatif donné, ouvrant ainsi la voie à des applications de recherche d’information dans des documents audiovisuels et multimédias.

4.2. FDL : l’extensibilité

L’objectif de FDL étant de fournir un moyen de décrire des objets audiovisuels et multimédias, il est centré sur la notion de descripteur (`Descriptor`), un descripteur étant défini par une classe d’objets (`Descriptor class`). FDL autorise les deux modes classiques d’extension — l’exigence d’extensibilité est ainsi satisfaite — de classes d’objets, et donc de descripteurs :

– **l’héritage de propriétés** (`Property`), il s’agit de l’héritage classique des langages à objets ;

– **la composition**, permise par FDL par le biais de structures (`Structure`), qui permet d’exprimer des relations méréologiques⁷.

Ces mécanismes sont entièrement contrôlés par FDL : l’héritage de propriétés est contrôlé par les mécanismes classiques d’héritage des langages à objets ; concernant la composition, FDL contrôle la conformité des descripteurs aux contraintes de la structure à laquelle ils sont ajoutés. Ils assurent ainsi que toute extension d’un descripteur FDL résulte en un descripteur FDL.

4.3. FDL : les relations ontologiques et la modularité

Dans le paradigme objet, les classes peuvent être vues à la fois comme module et comme type [MEY 97] :

– Selon le point de vue « module », un descendant définit une extension du module parent. Un module définit un ensemble de services. L’héritage permet de définir un nouveau module en ajoutant de nouveaux services à des modules existants. Par exemple, si q hérite de r alors tous les services de r sont disponibles dans q .

– Selon le point de vue « type », l’héritage définit la relation *is-a*.

Le mécanisme d’héritage d’un langage orienté objet permet ainsi de satisfaire aux exigences de hiérarchisation taxonomique — et, partant, d’expression de la sémantique des descripteurs — et de modularité.

7. De méréologie : théorie des relations entre tous et parties.

5. FDL à l'œuvre

FDL est au cœur du projet FERIA⁸ qui a pour objectif de faciliter le développement de la classe d'applications identifiée en introduction. FERIA est générique et implémente les fonctionnalités élémentaires nécessaires à la représentation et à la manipulation de contenus par l'intermédiaire de leurs descriptions ([BRU 05]).

Le langage est utilisé dans ce projet pour exprimer les résultats des outils d'analyse automatique, lesquels sont exploités soit par d'autres outils dans le cadre d'une chaîne de traitement automatique, soit directement par les applications multimédias élaborées dans le cadre du projet FERIA : PACE, une application de navigation dans une collection d'émissions ; FIDELIO/ALTO, applications, respectivement, d'enrichissement d'une captation d'opéra et de lecture des contenus ainsi produits.

6. Conclusion et perspectives

Une analyse des conséquences de la numérisation des contenus audiovisuels a permis d'identifier une classe d'applications multimédias nouvelles, centrées sur la notion de description. Nullement figée, cette classe constitue aujourd'hui encore un vaste terrain d'exploration. Les applications nouvelles qui la constituent font naître le besoin d'un langage de description des documents audiovisuels et multimédias qui soit notamment modulaire et extensible, et qui permette l'expression de relations ontologiques et méréologiques.

Nous avons montré dans cet article pourquoi la norme MPEG-7, candidate naturelle *a priori*, ne répond pas à toutes nos exigences. FDL, notre proposition, y répond — partiellement pour l'instant, les mécanismes de composition étant à perfectionner — et se positionne ainsi en alternative à MPEG-7. Il est doré et déjà exploité dans le cadre du projet FERIA.

Les efforts à venir porteront essentiellement sur l'extension des capacités de composition de descripteurs de FDL. Afin de prendre en compte la spécificité temporelle — et spatio-temporelle — des descripteurs, nous travaillons à établir un formalisme à la fois bien fondé et calculable qui permette l'expression de compositions spatiales et temporelles. A la suite de quoi nous envisageons de pourvoir le langage en mécanismes d'inférence de relations de subsomption entre descripteurs composés, problème identifié comme ardu ([CAR 00]).

8. Framework pour l'Expérimentation et la Réalisation Industrielle d'Applications multimédias. Ce projet RIAM, en fin de développement, est mené par un consortium qui réunit l'INA, Arte France, les équipes de recherche TexMex et Vista de l'IRISA, Samova et Sig de l'IRIT, ainsi que les sociétés Vecsys, NDS Technologies et C-S.

7. Bibliographie

- [BIR 04] BIRON P. V., MALHOTRA A., « XML Schema Part 2 : Datatypes (Second Edition) », Recommendation W3C, 2004.
- [BRU 03] BRUNIE V., TRONCY R., « Essai d'utilisation de MPEG-7 », rapport, 2003, Institut National de l'Audiovisuel (INA).
- [BRU 05] BRUNIE V., CARRIVE J., VINET L., « Ingénierie des documents audiovisuels : le projet FERIA », soumis à TSI, 2005.
- [CAR 00] CARRIVE J., PACHET F., RONFARD R., « CLAViS - A Temporal Reasoning System for Classification of Audiovisual Sequences », *Proceedings of RIAO*, Paris, France, 2000.
- [EUZ 03] EUZENAT J., NAPOLI A., BAGET J.-F., « XML et les Objets (Objectif XML) », *L'Objet*, vol. 9, n° 3, 2003, p. 11–37.
- [FAT 04] FATEMI N., LALMAS M., RÖLLEKE T., « *Semantic Web and Information Retrieval* », chapitre How to Retrieve Multimedia Documents Described by MPEG-7, 2004.
- [HUN 99] HUNTER J., « DSTC, A Proposal for an MPEG-7 DDL, P547 », MPEG-7 AHG Test and Evaluation Meeting, Lancaster, Février 1999.
- [HUN 00] HUNTER J., NACK F., « An Overview of the MPEG-7 Description Definition Language (DDL) Proposals », *Signal Processing : Image Communication Journal, Special Issue on MPEG-7*, vol. 16, 2000, p. 271–293.
- [ISO01a] ISO/IEC JTC1/SC29/WG11 MPEG01/M7009, Singapore, « Text of ISO/IEC 15938-5 FCD Information Technology - Multimedia Content Description Interface - Part 5 : Multimedia Description Scheme », Mars 2001.
- [ISO01b] ISO/IEC JTC1/SC29/WG11/N4288, Singapore, « Text of ISO/IEC 15938-2/FDIS Information Technology - Multimedia Content Description Interface - Part 2 : Description Definition Language », Septembre 2001.
- [ISO02] ISO/IEC JTC1/SC29/WG11, « Text of ISO/IEC 15938-7 Information Technology - Multimedia Content Description Interface - Part 7 : Conformance Testing », 2002.
- [KIF 95] KIFER M., LAUSEN G., WU J., « Logical Foundations of Object-Oriented and Frame-Based Languages », *Journal of ACM*, vol. 42, 1995, p. 741–843.
- [MAR 02a] MARTÍNEZ J. M., KOENEN R., PEREIRA F., « MPEG-7 : The Generic Multimedia Content Description Standard, Part 1 », *IEEE Multimedia*, vol. 9, n° 2, 2002, p. 78–87.
- [MAR 02b] MARTÍNEZ J. M., KOENEN R., PEREIRA F., « MPEG-7 : The Generic Multimedia Content Description Standard, Part 2 », *IEEE Multimedia*, vol. 9, n° 3, 2002, p. 83–93.
- [MEY 97] MEYER B., *Conception et Programmation Orientées Objet*, Eyrolles, 1997.
- [PER 01] PEREIRA F., « MPEG-7 Requirements Document V.16 », ISO/IEC JTC1/SC29/WG11/N4510, Pattaya, Décembre 2001.
- [THO 04] THOMPSON H. S., BEECH D., MALONEY M., MENDELSON N., « XML Schema Part 1 : Structures (Second Edition) », Recommendation W3C, 2004.
- [TRO 04a] TRONCY R., CARRIVE J., LALANDE S., POLI J.-P., « A Motivating Scenario for Designing an Extensible Audio-Visual Description Language », *Proceedings of CORIMEDIA*, Sherbrooke, Canada, Octobre 2004.
- [TRO 04b] TRONCY R., CARRIVE J., « A reduced yet extensible audio-visual description language », *DocEng '04 : Proceedings of the 2004 ACM symposium on Document engineering*, Milwaukee, Wisconsin, USA, 2004, ACM Press, p. 87–89.