



HAL
open science

Indexation de documents AV : ontologies, patrons de conception et d'utilisation

Antoine Isaac, Bruno Bachimont, Philippe Laublet

► **To cite this version:**

Antoine Isaac, Bruno Bachimont, Philippe Laublet. Indexation de documents AV : ontologies, patrons de conception et d'utilisation. IC - 16èmes Journées francophones d'Ingénierie des Connaissances, May 2005, Nice, France. pp.157-168. hal-01023959

HAL Id: hal-01023959

<https://inria.hal.science/hal-01023959v1>

Submitted on 15 Jul 2014

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.

Indexation de documents AV : ontologies, patrons de conception et d'utilisation^{*}

Antoine Isaac^{1,2}, Bruno Bachimont^{1,3}, Philippe Laublet²

¹ Institut National de l'Audiovisuel, Direction de la Recherche
4, Av. de l'Europe - 94366 Bry-sur-Marne
{aisaac,bbachimont}@ina.fr

² LaLICC, CNRS-UMR 8139, Université Paris IV-Sorbonne
Philippe.Laublet@paris4.sorbonne.fr

³ Université de Technologie de Compiègne

Résumé : Nous rappelons que l'indexation est nécessaire aux systèmes documentaires audiovisuels, et qu'elle exige ce que les langages documentaires existants ne peuvent apporter : une articulation convenable entre relations, contrôle et raisonnement. Dans le cadre de projets expérimentaux comme OPALES, nous avons pu constater qu'une indexation utilisant des ontologies pouvait remédier à cela. Néanmoins, il convient de diminuer la complexité de cette indexation pour l'indexeur ; nous proposons pour cela de généraliser l'usage de patrons d'indexation. Nous nous intéressons également à la manière dont on peut introduire ces patrons d'utilisation dès la conception des ontologies : on peut les articuler avec des patrons de conception de haut niveau, afin d'obtenir des ontologies qui soient légitimées tant sur le plan théorique que sur le plan applicatif.

Mots-clés : Indexation, Documents audiovisuels, Ontologies, Patrons de conception, Patrons d'utilisation

1 Introduction

L'indexation des documents audiovisuels (AV) est une activité ancrée dans le quotidien de l'INA. Le contexte de cette pratique est cependant en train d'évoluer : environnement technique, avec la numérisation et les outils afférents, mais aussi contexte organisationnel, avec l'émergence de petites communautés travaillant de manière précise, exigeante, à partir des documents indexés.

Pour parvenir à accorder les nouvelles techniques avec les usages existants ou pressentis, l'INA mène une réflexion s'appuyant à la fois sur sa connaissance des

^{*}Ce travail a été initié pour le projet RIAM OPALES. Nous tenons à en remercier l'ensemble des participants, et en particulier Véronique Malaisé, avec qui nous avons travaillé pour l'élaboration de certaines des ressources ontologiques présentées ici.

pratiques documentaires, sur des travaux fondamentaux et sur des projets expérimentaux. L'un de ceux-ci, OPALES, a eu pour ambition de tester la faisabilité d'une approche d'indexation utilisant des ontologies, dans un cadre restreint à des communautés ciblées (Isaac A. *et al.*, 2004). Notre participation à ce projet a été l'occasion de prolonger un effort, déjà ancien à l'INA, portant sur la conception d'ontologies pour l'indexation AV (Bachimont B., 2000), en nous concentrant davantage sur le volet formel et inférentiel de ces artefacts.

Nous allons d'abord voir comment les besoins en indexation AV amènent à se tourner vers la représentation de connaissances, et comment la complexité liée à un tel choix peut être réduite par le recours à des structures relationnelles prédéfinies, des *patrons d'indexation* . Nous discuterons ensuite de l'introduction de ces patrons d'utilisation dans la conception des ontologies, au regard de propositions récentes centrées sur la notion de *patron de conception* . Nous montrerons en particulier comment l'utilisation de connaissances de raisonnement permet d'articuler ces deux approches en un cadre méthodologique cohérent.

2 Vers l'indexation ontologique des documents AV

2.1 Documents AV et indexation : besoins et problèmes

La recherche et la réutilisation de documents AV se font suivant des critères plus liés à leur contenu qu'aux informations catalographiques sur leur cycle de vie – date de production, auteur, etc. A l'INA, on trouve par exemple des requêtes de documents concernant « Malraux, en support film couleur » ou des « images d'actualités illustrant des tempêtes violentes »... Ce besoin concerne tout autant le public des chercheurs qui utilisent ces objets comme support de leur travail que celui des spécialistes en production audiovisuelle.

Pour utiliser le document AV, on est ainsi obligé d'*interpréter* son contenu. En effet, la forme audiovisuelle, contrairement à la forme textuelle, ne repose pas sur un système fonctionnel d'assignation conventionnelle de sens aux unités de contenu, tel que la langue (Bachimont B., 1998). En fonction d'un point de vue particulier lui permettant de choisir parmi les interprétations potentiellement infinies d'une image, un opérateur doit donc effectuer une analyse du contenu documentaire pour en déterminer les aspects intéressants à fins de recherche ou de réutilisation. Il doit ensuite reformuler le résultat de cette analyse sous une forme plus adaptée à la réinterprétation et à l'exploitation : il *indexe*¹.

Un système documentaire AV se doit donc de consigner une interprétation du contenu des documents, et d'instrumenter cette interprétation de manière efficace. Si l'index sert de point de passage obligé entre l'utilisateur du système et le document, on doit faire particulièrement attention aux conditions de sa production et de sa réception. Se pose en particulier le problème de la *continuité sémantique* : les interprétations de celui qui indexe et de celui qui accède à

¹Dans le cas du texte, on peut se servir des éléments de contenu repérables physiquement, les mots, pour indexer un document. Cette approche, à la base de la recherche de contenu textuel, demeure inapplicable dans le cas de l'audiovisuel.

l'index doivent correspondre, et être respectées lors de l'exploitation des index par un système informatique. Par exemple, pour l'indexation d'images médicales, **opération** doit renvoyer à une action chirurgicale, et non à un calcul.

Pour pallier ces difficultés, on utilise des langages documentaires comme les *thesaurus*, qui contrôlent le vocabulaire des index et replacent les termes employés dans leur contexte d'usage *via* un réseau sémantique de spécialisations et d'équivalences. On utilisera **opération chirurgicale**, notion spécialisant **action-médecine**. Un système utilisant un thesaurus peut également, en renvoyant des documents traitant de notions liées à celles d'une requête, apporter une assistance cohérente par rapport à ce qu'attendent indexeur et chercheur.

Le recours au thesaurus ne peut cependant tout résoudre. Ainsi, la relation hiérarchique structurant les champs sémantiques mélange souvent subsomption *stricto sensu* et méréonomie, et peut causer de nombreuses imprécisions : **don du sang** spécialisant **opération chirurgicale**, par exemple. Ensuite se posent des problèmes d'expressivité : il faudrait utiliser dans les index des relations entre les descripteurs, relations qui dépendraient plus des notions du domaine d'application et de leur actualisation dans les documents que des liens sémantiques génériques trouvés dans un thesaurus.

Certains thesauri proposent bien un mécanisme permettant de préciser le rôle dans la situation décrite des entités auxquelles réfèrent les descripteurs : les *facettes*. A l'intérieur d'un même domaine, on va classer les notions suivant des catégories de haut niveau – *personne, action, etc.* – qui indiquent implicitement des relations entre ces notions. Ce mécanisme peut se combiner à celui des *formulaires*, où des champs correspondant à un schéma adapté à l'application structurent les descriptions : par exemple, un champ *participant* apparaîtra dans un formulaire dédié à la description d'une action. Néanmoins, ces liens restent souvent très généraux, et les mécanismes proposés ne sont pas suffisamment souples pour préciser leur type en fonction du cas descriptif rencontré.

2.2 Des index sémantiques relationnels

Il faut donc que le système documentaire puisse utiliser un langage d'indexation capable d'articuler des concepts et des relations de manière explicite et contrôlée. Une solution est de se tourner vers les techniques de représentation des connaissances, et en particulier vers celle s'appuyant sur des *ontologies*.

Expressivité

Les ontologies permettent de spécifier des primitives de connaissance – concepts, relations et règles associées – pouvant être employées pour construire des index structurés adaptés à la richesse du contenu documentaire. Ainsi, dans l'une des applications du projet OPALES, nous pouvons trouver un documentaire qui contient une séquence d'archives clarifiant un discours historique et une séquence d'interview clarifiant un aspect technique. Plutôt que de livrer pêle-mêle les termes correspondants, il sera évidemment préférable de créer l'index relationnel de la figure 1.

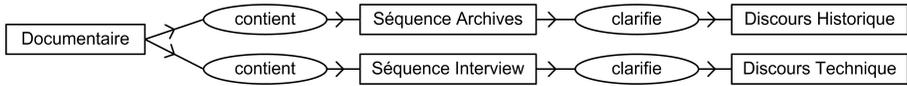


FIG. 1 – Index ontologique

Contrôle de la cohérence sémantique

Les ontologies contiennent des spécifications formelles, obéissant à une sémantique bien établie et exploitables par des systèmes de raisonnement. Ceci permet de contrôler automatiquement le vocabulaire et la structure des index : on ne peut pas utiliser n'importe quelle notion, et on ne doit pas articuler n'importe comment les notions entre elles². Les définitions formelles des notions présentes, plus précises que celles autorisées par la relation hiérarchique d'un thesaurus, peuvent être couplées à des spécifications langagières (Bachimont B., 2000). Il est ainsi possible de replacer correctement chaque concept et relation dans son contexte d'usage, et ce par des moyens immédiatement compréhensibles par des utilisateurs humains. La continuité interprétative est donc assurée d'un bout à l'autre de la chaîne documentaire.

Assistance à la recherche d'index

Plus que ne le permet un thesaurus, une ontologie peut accroître la pertinence d'un système d'information, en facilitant la correspondance entre les requêtes effectuées et les index présents. On peut en effet ajouter, aux côtés des hiérarchies de concepts et de relations, des axiomes logiques, connaissances de raisonnement qui s'appliquent à ces primitives. Par exemple, dans une ontologie dédiée à la représentation des rapports entre les éléments de contenu des documents AV, nous avons introduit une règle de composition entre les relations **contient** et **clarifie**. Muni de cette règle, un système d'inférence déduira de l'index de la figure 1 que le documentaire indexé éclaircit à la fois des questions historiques et techniques (figure 2). L'information peut sembler triviale, mais une prise en compte pertinente, même élémentaire, de ces relations propres aux domaines d'application est précisément reconnue comme un point d'amélioration des systèmes documentaires existants (Tudhope D. *et al.*, 2001).

La recherche est ainsi rendue plus robuste, ce qui permet d'alléger une partie de la charge liée à la complexité de l'indexation. En effet, même dans un contexte applicatif limité, un indexeur ne peut pas tout expliciter : il doit se reposer sur un fonds de connaissances qu'il espère partagé par le chercheur. Or réinterpréter un index conceptuel est toujours plus difficile que comprendre une liste de mots-clés. Et reformuler une requête en cas de résultat négatif devient plus délicat. Un système de recherche ontologique, de par sa nature formelle et logique, a justement pour but de renvoyer un index s'il implique la requête

²Les domaines de relations, la hiérarchie de subsomption et des axiomes comme l'exclusion entre concepts sont des connaissances qui permettent de contrôler les assertions relationnelles.

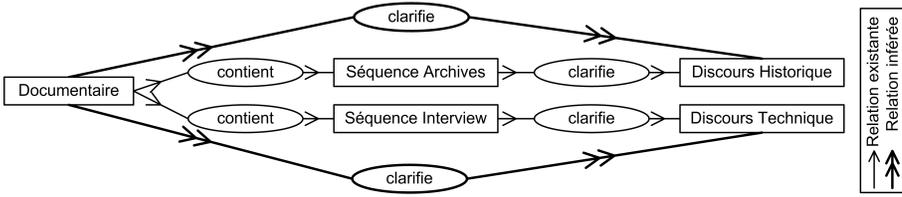


FIG. 2 – Application d’une règle d’inférence à un index

d’après les connaissances consignées dans l’ontologie, et non plus uniquement en cas de correspondance exacte. Utiliser des calculs d’inférence aide grandement à diminuer les risques induits par l’utilisation d’index conceptuels dans les processus de recherche.

2.3 Des patrons d’indexation pour la construction et l’utilisation de descriptions complexes

Le recours à des index ontologiques, s’il permet d’optimiser le fonctionnement d’un système d’information, va donc de pair avec une complexification de ce système : l’accès aux index peut être plus délicat, leur création aussi. Il devient alors nécessaire de guider l’indexeur dans la tâche difficile de sélection et d’articulation des notions rendant compte de son analyse. L’ontologie permet de contrôler la structure des descriptions, mais elle ne prescrit rien : les domaines des relations apportent bien une forme de contrainte locale, mais ils ne donnent pas la structure générale d’un index. . . L’expérience dans le cadre d’un projet appliqué comme OPALES montre pourtant qu’il existe une grande régularité en ce qui concerne le contenu des annotations, tant du point de vue conceptuel que structurel. Cette observation a favorisé l’utilisation de *graphes patrons* mobilisant les concepts et les relations apportés par l’ontologie.

Un patron d’indexation est une construction relationnelle adaptable, qui présente, dans un contexte d’indexation typique, l’articulation entre les concepts et les relations ontologiques les plus caractéristiques, *structurants*, d’une application. Grâce à lui, un indexeur dispose d’une structure de connaissances accessible, explicite, globalement pertinente pour son point de vue, qu’il peut aisément reconfigurer pour rendre compte de ses interprétations. Le graphe de la figure 3 montre l’un de ces patrons, tiré d’une des applications d’OPALES : l’utilisateur peut copier un tel graphe dans son index, puis le modifier en spécialisant les concepts et relations présentes, en en ajoutant d’autres ou bien en en retirant, suivant les éléments d’information reconnus pertinents. Il est à noter que les patrons d’indexation peuvent également être présentés lors de la construction d’une requête pour la guider – et en optimiser le résultat – en illustrant la manière dont la base d’index est construite. Ils jouent alors un rôle semblable à celui de formulaires d’interrogation.

L’utilisation de guides de description n’a rien de fondamentalement nouveau.

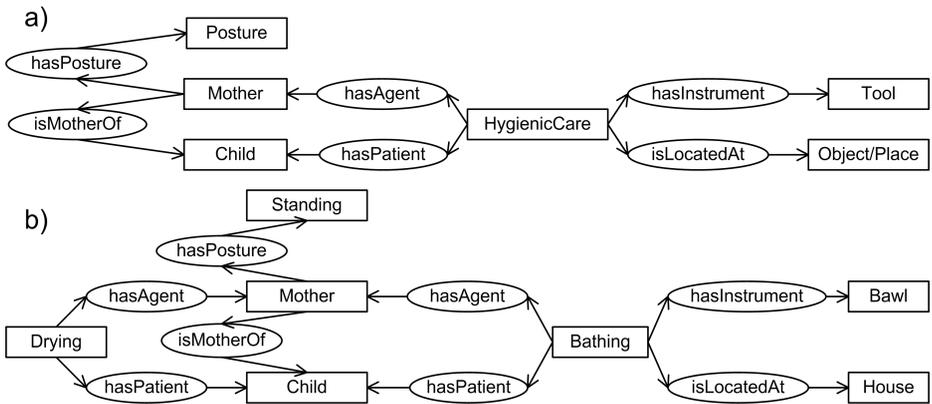


FIG. 3 – (a) Un patron d’indexation pour la description de soins sur des enfants, et (b) une spécialisation possible

Déjà, dans le cadre de l’indexation textuelle, on utilise couramment des références – *grilles* – d’indexation recommandant, pour des thèmes ou des événements précis, des descripteurs pertinents. On peut même, à la manière d’un formulaire, faire apparaître les notions préconisées dans des champs. De telles idées ont fait l’objet d’adaptations plus ou moins explicites dans des expériences de création d’annotations à base de connaissances (Schreiber A. Th. *et al.*, 2001; Hyvönen E. *et al.*, 2004). Mais, curieusement, toutes tendent à figer les descriptions produites en proposant des interfaces de type formulaire – où l’on peut par exemple choisir les concepts joints par une relation sans pouvoir modifier celle-ci – alors qu’il paraît tout à fait possible de concilier cohérence et souplesse lors de l’utilisation des patrons d’indexation. Les observations tendent en effet à exhiber des patrons plutôt simples, tant en termes qualitatifs – le niveau, basique pour l’application, auquel appartiennent concepts et relations – que quantitatifs – un patron contient généralement un petit nombre de notions.

Il est intéressant de constater que le patron a également un rôle à jouer dans la spécification du fonctionnement inférentiel du système à base de connaissances. En effet, sa configuration et les notions qu’il présente sont par définition proches des index produits, ce qui autorise à le considérer comme un structure pivot, à laquelle on peut comparer les requêtes et les descriptions de manière à faciliter leur rapprochement. De fait, dans un système ontologique riche, de nombreuses connaissances de raisonnement – définitions de concepts, axiomes relationnels – accompagnent les notions qu’il mobilise. Autour d’un patron, on peut trouver des règles qui vont soit déduire de nouvelles connaissances à partir de la structure existante (pour répondre à des questions dont les besoins ne seraient pas directement traités par le patron, figure 4a), soit ramener des connaissances « non standard » à un format plus proche de celui qui est généralement utilisé pour les requêtes (figure 4b).

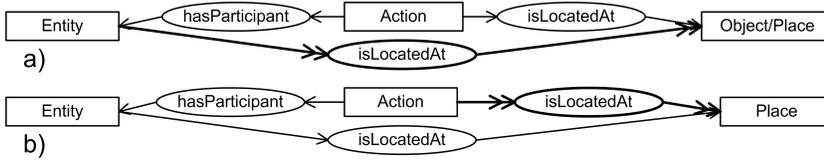


FIG. 4 – Connaissances de raisonnement et patron

2.4 Utilisation des patrons dans le projet OPALES

Au cours du projet OPALES, nous avons mis en œuvre cette approche pour deux points de vues applicatifs : la description de vidéos relatives à la petite enfance, et l’analyse de documents pédagogiques mettant en lumière l’articulation entre procédés AV et présentation d’un discours scientifique.

Ces expérimentations ont nécessité la construction de deux ontologies comprenant chacune quelques centaines de concepts et dizaines de relations, ainsi que quelques dizaines de règles d’inférences. Ici, c’est le formalisme des graphes conceptuels (Sowa J., 1984) qui a été retenu : il permet une visualisation immédiate des ressources, et des raisonnements – reposant sur des mécanismes de projection – satisfaisants au vu de nos besoins³.

Les représentants des domaines concernés (six enseignants et documentalistes pour les deux domaines) ont été impliqués dans la création et l’évaluation de ces ressources. Pour l’évaluation, une dizaine d’heures de vidéo a été finement indexée (une petite centaine d’index) à partir des graphes patrons d’indexation, ce qui a permis de valider l’adéquation de notre approche, mais aussi sa robustesse. La souplesse offerte par la combinaison des graphes patrons et des règles d’inférence a en effet permis de se limiter à un patron par point de vue applicatif, ce qui a facilité l’appropriation de cette démarche par des utilisateurs non-experts en représentation de connaissances.

3 Ingénierie ontologique et patrons d’indexation

Les patrons d’indexation sont, on l’a vu, intimement liés à l’efficacité du système d’information. Dès lors, la question de leur introduction dès l’étape de conception des ontologies se pose. Cette activité cruciale, difficile, doit offrir, au moindre coût, des garanties d’efficacité pour le système qui utilise son produit. Nous allons montrer comment des *patrons de conception* peuvent être introduits, de sorte que la conception bénéficie à la fois d’une légitimité théorique liée à l’adhésion à des principes propices à consensus et d’une reconnaissance « métier » apportée par une prise en charge correcte des structures de description typiques.

³D’autres langages pourraient convenir, comme le langage RDF du w3c, pourvu qu’on les utilise avec des outils d’inférence suffisamment élaborés.

3.1 Patrons de conception

De nombreuses réflexions méthodologiques ont été conduites pour proposer des cycles de vie généraux ou des critères de cohérence « nettoyant » les ontologies, comme ceux de (Guarino N. & Welty C., 2002). Cependant, ces propositions sont plus axées sur la qualité du développement de l'ontologie lui-même que sur les compétences du système d'information qui l'utilisera : on sait comment construire, mais on ne sait pas ce qu'il faut spécifier afin que le système soit pertinent pour une application donnée. Pour cela, des travaux essaient d'initier la conception avec des composants issus d'ontologies de haut niveau satisfaisant des critères d'utilisabilité : les connaissances génériques doivent être facilement mobilisables lors de la production des ontologies, et leur contenu doit être adapté à un éventail applicatif considéré à présent comme prioritaire. On compte ainsi rendre plus explicite l'engagement ontologique plaçant la conceptualisation de l'application dans une vision « propre » du monde qui l'environne.

C'est une telle vision qui sous-tend les travaux actuels sur les patrons de conception ontologiques, ainsi nommés en référence aux *design patterns* de l'ingénierie logicielle. Gangemi, dans (Gangemi A. & Mika P., 2003), propose de réutiliser les notions de haut niveau de l'ontologie DOLCE (Gangemi A. *et al.*, 2002), obéissant aux principes de cohérence formelle énoncés par Guarino, au sein d'une structure relationnelle qui présente ces notions dans un contexte d'utilisation – ici, les « Descriptions et Situations ». Le patron qui en résulte (*D&S*, cf. figure 5) présente des *descriptions* de *séquences d'événements* qui ordonnent des entités temporelles (*perdurants*), des rôles que des entités physiques (*endurants*) peuvent jouer dans ces événements, ainsi que des *paramètres* qui sont utilisés pour décrire rôles et événements, et prennent leurs valeurs dans des *régions* plus ou moins abstraites. Cette structure générique cherche à s'abstraire d'un domaine particulier tout en résolvant un problème de description ontologique donné. Elle est aisément adaptable à la conceptualisation d'un domaine particulier, pourvu que l'usage visé par cette nouvelle conceptualisation soit similaire à celui qui a dicté l'élaboration du patron. Il suffit de rattacher aux notions qu'il introduit celles, plus spécifiques, du domaine d'application envisagé. Les objets résultant de cette adaptation forment une modélisation réduite aux notions les plus générales du domaine, dite *noyau*, qui pourra être à son tour être spécialisée pour introduire les éléments de connaissances nécessaires à des besoins plus précis.

Nous avons tenté de reproduire cette approche pour une ontologie dédiée à la description des documents AV (Isaac A. & Troncy R., 2004) : le patron D&S cible en effet un usage qui pouvait convenir à nos besoins. Pour adapter ces notions au domaine AV, nous avons appliqué le patron aux deux activités structurant le cycle de vie du document audiovisuel de télévision, avant son archivage : la production et la diffusion. Du point de vue de la diffusion, la description d'un document AV suppose un enchaînement d'événements de diffusion, tels que l'*émission* d'un programme sur un *canal de diffusion*. Des rôles de diffusion, comme *diffuseur* ou *récepteur*, sont joués par des entités qui peuvent être des *organisations* ou des *personnes*. Et, dans la description de ces événements, nous trouvons des paramètres comme la *date de diffusion* ou le *taux d'audience*,

valués respectivement par des *dates* et des *taux*. La figure 5 montre comment ces concepts spécialisent ceux du patron de conception pour obtenir un premier patron d'utilisation, que nous appelons de haut niveau.

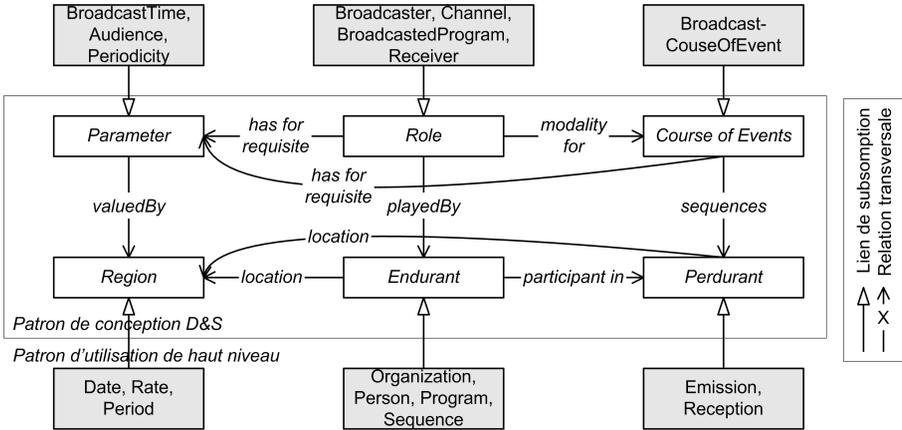


FIG. 5 – Introduction de concepts généraux de l’audiovisuel suivant le patron de conception D&S de (Gangemi A. & Mika P., 2003)

3.2 Vers une solution articulant patrons de conception et patrons d’utilisation

Rattacher une ontologie de domaine à un patron de conception de haut niveau améliore indéniablement sa qualité intrinsèque : on clarifie l’engagement ontologique en adhérant explicitement à une conceptualisation qui a bénéficié d’une effort théorique important, et qui de plus apporte des connaissances de raisonnement se répercutant, *via* la relation de subsumption, aux notions du domaine. On peut envisager que l’ontologie sera plus facilement partageable, et réutilisable par les modélisateurs d’autres domaines. Pour autant, sera-t-elle utilisable par l’indexeur ou l’expert du domaine ? La complexité des notions impliquées par des considérations abstraites peut masquer la vue applicative sur le domaine, et limiter la pertinence de l’ontologie. Spécialiser un patron de conception ne suffit pas toujours à obtenir un véritable patron d’utilisation. Pour cela, on doit faire la part entre les décisions théoriques d’un patron de conception quant à l’organisation du monde à décrire, et le domaine qui peut exiger une modélisation plus pragmatique. Dans un cadre applicatif concret, il faut veiller à se placer au niveau des notions typiques de l’application – qui ne sont pas nécessairement les notions générales du domaine – et chercher à obtenir une structure vraiment proche du besoin descriptif.

Par exemple, une étude des besoins en matière de description documentaire AV (reposant en grande partie sur l’examen des notices produites à l’INA et

concernant à ce titre autant le catalogage que l'indexation) aboutit à l'obtention d'un patron d'utilisation dont la structure relationnelle, illustrée en figure 6, est beaucoup plus simple que celle esquissée en figure 5. Les descriptions dont nous avons réellement besoin sont centrées plus sur les documents que sur les situations de production ou de diffusion : certaines subtilités informationnelles du patron de conception ne seront donc pas nécessairement à expliciter pour notre application. Par exemple, savoir qu'un document donné joue le rôle de programme diffusé dans une séquence d'événements de diffusion est relativement inutile pour nous. On préférera plutôt des propriétés simples permettant de relier directement le programme à sa date de diffusion, sa mesure d'audience, etc.

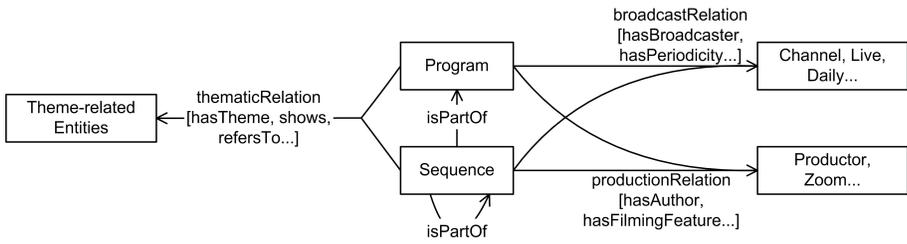


FIG. 6 – Patron de description d'un document AV

Nous proposons, pour faire coexister patron d'utilisation de haut niveau et patron applicatif au sein d'un cadre méthodologique cohérent, d'articuler les deux formes d'expression par des connaissances de raisonnement formelles. L'objectif est d'autoriser un système d'inférence à passer de l'une à l'autre à chaque fois que cela est faisable. En particulier, il faudra présenter les liens « simples » demandés par le patron d'utilisation comme autant de *raccourcis relationnels* de chemins présents dans le patron de haut niveau. Par exemple, on peut considérer qu'une relation *was broadcasted at* entre un programme et une date est utile si l'on ne veut pas dire que le programme joue le rôle de message dans une séquence d'événements qui admet pour paramètre une date de diffusion évaluée par ladite date. Il nous faut alors introduire un axiome (cf. figure 7) qui permettra de gérer simultanément les concepts et les relations des deux points de vue⁴.

De telles connaissances de raisonnement, exprimées de manière formelle – et donc exploitables de façon automatique – autorisent un système à gérer simultanément un point de vue adéquat cognitivement et opérationnellement avec les usages du domaine, et un autre se référant plus à des principes de modélisation de haut niveau. Cette situation a deux avantages. D'abord, la connaissance est utilisable de manière satisfaisante pour l'application visée : on peut spécifier de manière naturelle des connaissances de raisonnement pertinentes, et on

⁴On doit remarquer qu'une équivalence parfaite n'est pas forcément atteignable. Dans l'exemple donné, une information exprimée selon le patron de haut niveau contiendra des assertions qu'on ne pourra pas déduire précisément – on utilisera des quantificateurs existentiels – de connaissances issues du patron d'utilisation.

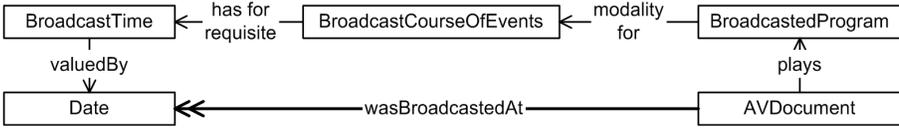


FIG. 7 – Déduction d’une relation à partir d’assertions respectant le patron D&S

prescrit bien une description adaptée aux préoccupations applicatives. Ensuite, cette connaissance acquiert un statut plus consensuel et peut être réutilisable par d’autres applications construites autour du même noyau ontologique, ce qui ne peut qu’améliorer l’interopérabilité de ces systèmes.

4 Conclusion

L’indexation, nécessaire pour exploiter les documents AV, peut profiter de techniques permettant de créer et de manipuler des index en tant que véritables connaissances structurées. Pour cela, il faut des ontologies riches, capables d’assister efficacement le processus de recherche, mais complexes à concevoir et à utiliser. Nous avons donc utilisé des patrons d’indexation, conçus ici comme des structures relationnelles typiques d’un domaine d’application. Couplée à celle de connaissances de raisonnement, leur utilisation diminue la charge cognitive pesant sur les utilisateurs.

Poursuivant cet effort de légitimation des systèmes ontologiques, nous avons montré que dans certains cas on peut rattacher les patrons d’utilisation ontologiques à des patrons de conception de haut niveau apportant une consolidation théorique bienvenue. Pour ne pas trop s’écarter des besoins de description réels, souvent très simples, nous proposons d’utiliser des connaissances de raisonnement dédiées au passage d’un mode d’expression de la connaissance à un autre. Légitime en termes d’usages, l’ontologie le sera aussi sur le plan théorique.

Des travaux partageant nos préoccupations (Rector A. *et al.*, 1999) notent cependant les problèmes de calculabilité et d’expressivité relatifs à de telles solutions. Dans cette approche, il s’agit bien en effet d’introduire des représentations intermédiaires entre l’utilisateur et la connaissance exprimée formellement, mais celles-ci sont pas gérées par les mécanismes de raisonnement ontologiques standards. Ce choix est dû à la décision fondamentale d’exclure ces « connaissances d’encodage » des ontologies, mais aussi aux capacités parfois restreintes des systèmes de raisonnement concernés.

Si nous-mêmes considérons que les patrons et leur articulation avec des connaissances plus abstraites font bien partie de la spécification d’une conceptualisation *située*, nous ne devons cependant pas négliger le problème de l’opérationnalisation des connaissances de raisonnement. A un travail qui a validé une hypothèse relative à l’utilisation des ontologies – les patrons d’indexation – puis détaillé des éléments de conception et d’exploitation de ces connaissances – le couplage

entre patron de conception et d'utilisation – il conviendrait d'ajouter des tests opérationnels à plus grande échelle. Il importe de valider un choix de formalismes et d'outils dont les possibilités expressives et l'efficacité calculatoire permettront de tirer pleinement parti d'une approche dont les retombées en termes d'usage sont indéniables.

Références

- BACHIMONT B. (1998). Bibliothèques numériques audiovisuelles : des enjeux scientifiques et techniques. *Document numérique*, **2**(3-4).
- BACHIMONT B. (2000). Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en ingénierie des connaissances. In *Ingénierie des Connaissances : Evolutions récentes et nouveaux défis*. Eyrolles.
- GANGEMI A., GUARINO N., MASOLO C., OLTRAMARI A. & SCHNEIDER L. (2002). Sweetening ontologies with DOLCE. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW'2002)*, Sigüenza, Spain.
- GANGEMI A. & MIKA P. (2003). Understanding the semantic web through descriptions and situations. In *International Conference on Ontologies, Databases and Applications of Semantics (ODBASE'03)*, Catania, Italy.
- GUARINO N. & WELTY C. (2002). Evaluating ontological decisions with OntoClean. *Communications of the ACM*, **2**(45).
- HYVÖNEN E., SAARELA S. & VILJANEN K. (2004). Application of ontology techniques to view-based semantic search and browsing. In *Proceedings of the First European Semantic Web Symposium (ESWS 2004)*, Heraklion, Greece.
- ISAAC A., COUROUTNET P., GENEST D., MALAISÉ V., NANARD J. & NANARD M. (2004). Un système d'annotation multiforme et communautaire de documents AV : OPALES. In *Journée sur les Modèles Documentaires de l'Audiovisuel, SDN 2004*, La Rochelle, France. <http://archivesic.ccsd.cnrs.fr/>.
- ISAAC A. & TRONCY R. (2004). Designing an audio-visual description core ontology. In *Workshop on Core Ontologies in Ontology Engineering, 14th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2004)*. CEUR online Proceedings, <http://ceur-ws.org/Vol-118/>.
- RECTOR A., ZANSTRA P.E., SOLOMON W. D., ROGERS J. E., BAUD R. & AL. (1999). Reconciling user's needs and formal requirements : Issues in developing a re-usable ontology for medicine. *IEEE Transactions on Information Technology in BioMedicine*, **4**(2).
- SCHREIBER A. TH., DUBBELDAM B., WIELEMAKER J. & WIELINGA B. (2001). Ontology-based photo annotation. *IEEE Intelligent Systems*, **16**(3).
- SOWA J. (1984). *Conceptual structures : information processing in mind and machine*. Addison-Wesley, Reading (MA US).
- TUDHOPE D., ALANI H. & JONES C. (2001). Augmenting thesaurus relationships : Possibilities for retrieval. *Journal of Digital Information*, **1**(8).