

THATCamp Paris 2012 Non-actes de la non-conférence des humanités numériques

Éditions de la Maison des sciences de l'homme

Utiliser l'ontologie CIDOC CRM pour l'information relative au patrimoine culturel

Proposé par Anne-Violaine SZABADOS (équipe LIMC-ArScAn / Archéologies et Sciences de l'Antiquité - UMR7041, CNRS), Katell BRIATTE (Ministère de la Culture et de la Communication) et Rosemonde LETRICOT (Irht)

Collectif

DOI : 10.4000/books.editionsmsmh.319
Éditeur : Éditions de la Maison des sciences de l'homme
Lieu d'édition : Paris
Année d'édition : 2012
Date de mise en ligne : 1 octobre 2012
Collection : La Non-Collection
ISBN électronique : 9782735115273



<http://books.openedition.org>

Référence électronique

COLLECTIF. *Utiliser l'ontologie CIDOC CRM pour l'information relative au patrimoine culturel* : Proposé par Anne-Violaine SZABADOS (équipe LIMC-ArScAn / Archéologies et Sciences de l'Antiquité - UMR7041, CNRS), Katell BRIATTE (Ministère de la Culture et de la Communication) et Rosemonde LETRICOT (Irht) In : *THATCamp Paris 2012 : Non-actes de la non-conférence des humanités numériques* [en ligne]. Paris : Éditions de la Maison des sciences de l'homme, 2012 (généralisé le 19 avril 2019). Disponible sur Internet : <<http://books.openedition.org/editionsmsmh/319>>. ISBN : 9782735115273. DOI : 10.4000/books.editionsmsmh.319.

Ce document a été généré automatiquement le 19 avril 2019.

Utiliser l'ontologie CIDOC CRM pour l'information relative au patrimoine culturel

Proposé par Anne-Violaine SZABADOS (équipe LIMC-ArScAn / Archéologies et Sciences de l'Antiquité - UMR7041, CNRS), Katell BRIATTE (Ministère de la Culture et de la Communication) et Rosemonde LETRICOT (Irht)

Collectif

Présentation

- 1 Il s'agit de présenter et de discuter d'un modèle, émanant du monde des musées (Icom : the International Council of museums), pour structurer et encoder les données sur les œuvres matérielles et immatérielles du patrimoine culturel (*Cultural Heritage*). Cette norme, surtout utilisée dans le monde des musées et peu dans la communauté scientifique, s'applique à des documents hétérogènes et pourrait servir de modèle pivot. Cependant, elle semble peu connue en France. Il serait souhaitable de voir émerger une communauté d'intérêt et de partage d'expériences.
- 2 Après une introduction au Modèle Conceptuel de Référence CIDOC (ISO 21127:2006), les domaines d'application (musées, sites archéologiques, gestion de collections et de fouilles, corpus d'objets), les notions clés (objet/concept, événement/temps, lieu, intervenant/agent de l'événement) et leurs interactions seront discutés.

Introduction

- 3 Le CIDOC CRM est un modèle conceptuel de référence pour l'information relative au patrimoine culturel qui émane du monde des musées¹. Il a acquis le statut de norme ISO en 2006, mais reste en constante évolution et continue d'être mis à jour. La dernière

version de l'ontologie (v.5.0.4, décembre 2011) est en ligne sur le site du CIDOC CRM SIG. La nouvelle version de la norme ISO (et de sa traduction en français) qui prend en compte les récentes évolutions du modèle est en cours de publication.

- 4 La version SIG de l'ontologie, en anglais, est disponible gratuitement sur le site du CIDOC. En revanche, le document de la norme ISO accessible sur le site de l'ISO dans les versions anglaise et française, est payant. Une définition courte de la norme est disponible sur la page de l'ISO :
- 5 « ... Le domaine d'application de principe de l'ISO 21127 : 2006 peut être défini comme toutes les informations nécessaires pour l'échange et l'intégration de la documentation scientifique des collections de musée. [...] »
- 6 Le terme « collections de musées » englobe tout type de matériel rassemblé et exposé par des musées et des institutions apparentées, selon la définition de l'Icom. Ceci inclut des collections, des sites et des monuments en rapport avec des domaines tels que l'histoire sociale, l'ethnographie, l'archéologie, les beaux-arts, les arts appliqués, l'histoire naturelle, l'histoire des sciences et de la technologie.
- 7 La documentation des collections inclut la description détaillée d'objets individuels qui font partie des collections ainsi que des groupes d'objets et des collections dans leur ensemble. L'ISO 21127 : 2006 est spécifiquement censée couvrir des informations contextuelles (c'est-à-dire historiques, géographiques et théoriques qui donnent aux collections de musée leur signification culturelle et leur valeur). [...] »
- 8 Si le modèle émane du monde des musées, il n'est pas limité aux objets de musées ou de collections. Il traite de notions, de types de données, de domaines ou de champs disciplinaires très variés, dont le point commun est l'appartenance au « patrimoine culturel, aussi bien immatériel que matériel ».
- 9 Pouvant s'adapter à des documents et des notions très variés, le CIDOC CRM peut être un pivot pour traiter des données hétérogènes et pour permettre des approches interdisciplinaires. On peut, par exemple, l'employer à deux niveaux pour rassembler un ensemble hétérogène d'informations archéologiques issues d'un chantier de fouilles : d'une part pour structurer et encoder une grande partie de ces données (zones fouillées, géolocalisation, date des fouilles, date et identification des documents trouvés, liste et description des découvertes, état de conservation, référence à un taxon, lieu de conservation, intervenants, droits, etc.), d'autre part pour insérer des liens vers d'autres informations (reproductions iconiques, documents de fouilles, publication, thésaurus...).
- 10 C'est un modèle souple, qui peut être utilisé partiellement ou qui peut être augmenté d'ajouts pour des besoins spécifiques. Le modèle est en effet extensible, pourvu que les extensions respectent les règles de l'ontologie. Patrick Le Boeuf (BnF, membre actif du CIDOC) a en particulier travaillé sur l'extension FRBR *object-oriented* (FRBRoo)[8], une extension du CIDOC pour le monde des bibliothèques. L'articulation entre FRBR et le CIDOC permet de différencier l'œuvre de son contenu, de faire dialoguer le livre en tant qu'objet patrimonial et le livre en tant que contenu. Ainsi s'articulent les différents domaines métier. Une tentative d'extension dans le monde des archives semble pour le moment abandonnée à cause de problèmes de réception au sein de la communauté. Les

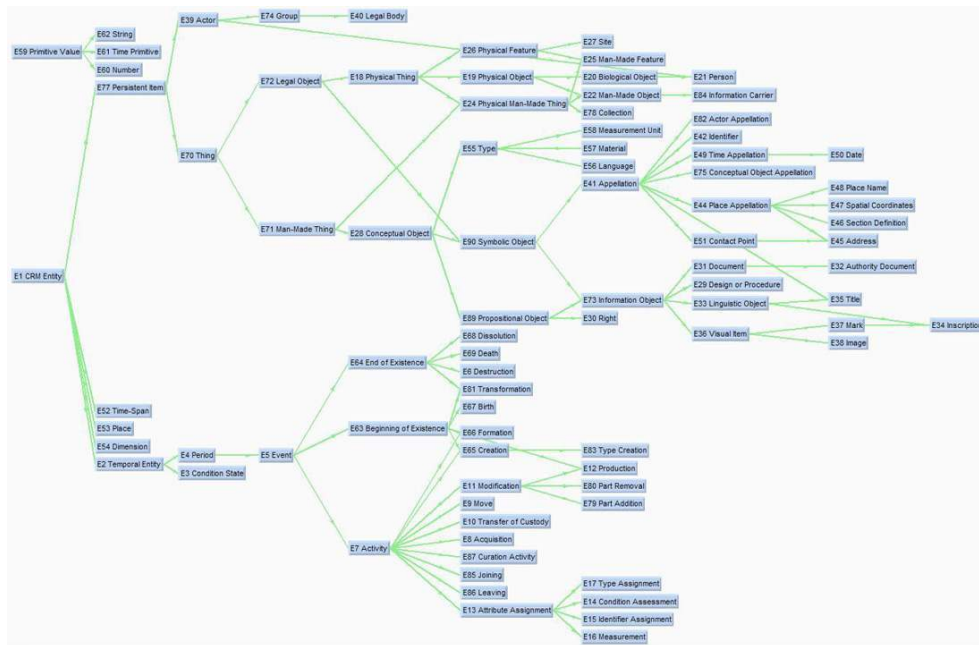
archives sont en train de créer leur propre modèle conceptuel, mais le risque est de voir proliférer les ontologies et de perdre en interopérabilité.

11 Quelques principes permettent de définir le modèle :

- Il est fondé sur cinq notions-clés qui sont l'objet (*Thing*) matériel et immatériel, l'agent (*Actor*) ou l'acteur de l'événement, le lieu (*Place*), le temps/événement (*Time/Event*) et le concept.
- Il inclut des éléments permettant de traiter des ensembles d'informations comme, par exemple, le droit, la documentation, la gestion de l'objet, etc.
- Il combine deux types d'éléments : les « classes », qui sont les notions et les concepts (au nombre de 90) et les « propriétés » (prédicats exprimés par un verbe) qui correspondent à une action, un verbe ou une interaction possible entre les classes (149 ont été définies)
- Il existe une hiérarchie entre les classes ainsi que des interactions classes/propriétés. Par exemple, la combinaison entre classes et propriétés permet de reformuler une information du type : « la statue est actuellement localisée dans tel lieu, tel musée ». La notion « est actuellement localisée » correspond à la propriété « P55_has current location (currently holds) » ; ou encore, « l'objet est en marbre » devient *l'objet* - P45 consist_of - E57 Material *marbre* (P45 étant une propriété et E57 une classe).
- La notion d'événement (Event) est essentielle et, combinée avec les autres éléments proposés dans le modèle, permet la description du contexte historique et des étapes de la vie de l'« objet ».

12 Un exemple de schéma de la hiérarchie des classes montrant l'articulation entre les différents niveaux de l'ontologie est disponible sur le site du CIDOC.

Le site du CIDOC



Le schéma donne une idée des différents types de concepts traités. Sont également disponibles sur cette page internet des schémas thématiques présentant des interactions classes/propriétés, et des modèles d'organisation de l'information (acquisition d'un objet de collection ; phases de création / modélisation d'un objet ; documentation et références ; relation temps / espace ; etc.).

Interopérabilité

- 13 L'Aphrodite Braschi, statue conservée à Munich au Staatliche Antikensammlungen und Glyptothek [inv. Gl 258] est une réplique romaine de l'Aphrodite de Cnide de Praxitèle. Une notice de description en DC (15 métadonnées du *Dublin Core* non qualifié) a été établie. Cependant, le *Dublin Core* « aplanit l'information » et met tout au même niveau comme par exemple, les dates de création de la statue ou de création du document numérique. Se limiter à l'emploi du DC oblige à choisir entre les informations sur l'enregistrement et celles sur le document traité (la statue) au lieu de pouvoir les hiérarchiser dans des types de balises différentes. Cette statue, comme beaucoup de documents du patrimoine culturel, génère beaucoup d'informations pertinentes pour les requêtes des internautes, mais le DC ne permet pas de toutes les traiter.

Exemple de l'Aphrodite Braschi et des métadonnées DC (volontairement trop renseignées)



```
<dc:title>Aphrodite Braschi</dc:title>
<dc:title>Munich, Staatliche Antikensammlungen und Glyptothek Gl 258
</dc:title>
<dc:creator>Praxitèle</dc:creator>
<dc:creator>inconnu</dc:creator>
<dc:subject>Aphrodite, Vénus, Antiquité, sculpture, archéologie...</
dc:subject>
<dc:description>statue en marbre d'Aphrodite (type Aphrodite de
Cnide)</dc:description>
<dc:description>Aphrodite nue et debout près d'un vase</
dc:description>
<dc:publisher>LIMC</dc:publisher>
```

```

<dc:contributor>équipe LIMC-France</dc:contributor>
<dc:date>IIe s. av. J.-C.</dc:date>
<dc:date>époque hellénistique récente</dc:date>
<dc:date>2011/08/20</dc:date>
<dc:date>2012/09/27</dc:date>
<dc:type>objet physique, statue</dc:type>
<dc:type>dataset</dc:type>
<dc:format>ressource électronique</dc:format>
<dc:identifiant>http:// http://www.limc-france.fr/objet/14767</
dc:identifiant>
<dc:source>http://www.limc-france.fr</dc:source>
<dc:language>fr</dc:language>
<dc:relation>http://www.limc-france.fr/image/1TN0683B</dc:relation>
<dc:coverage>Rome, Fiumicino, Munich, Glyptothèque , Glyptothek,
Antiquité, IIe s. av. J.-C., 2011/08/20</dc:coverage>
<dc:rights>copyright...</dc:rights>

```

Extrait d'une version en CIDOC CRM (prototype IDA)

```

<?xml version="1.0" encoding="UTF-8"?>
...
<crm:E24_Physical_ManMade_Thing> <!-- extrait de l'enregistrement
concernant la statue : -->
...
<crm:P129_is_subject_of> <!-- extrait de la déclaration de
l'enregistrement -->
<crm:E73_Information_Object>
<crm:P48_has_preferred_identifiant>
<crm:E42_Identifiant>http://www.limc-france.fr/objet/14767</
crm:E42_Identifiant>
</crm:P48_has_preferred_identifiant>
</crm:P129_is_subject_of>
...
<crm:P1_is_identified_by> <!-- extrait des informations concernant
l'objet -->
<crm:E41_Appellation>Aphrodite Braschi</crm:E41_Appellation>
</crm:P1_is_identified_by>
<crm:P45_consists_of> <!-- matériaux -->
<crm:E57_Material>marbre</crm:E57_Material>
<crm:P127_has_broader_term>
<crm:E55_Type>pierre</crm:E55_Type>
</crm:P127_has_broader_term>
</crm:P45_consists_of>
...
<crm:P62_depicts> <!-- mots-clés sur l' iconographie -->
<crm:E28_Conceptual_Object>
<crm:P2_has_type>
<crm:E55_Type>Aphrodite</crm:E55_Type>
<crm:E55_Type>Aphrodite de Cnide</crm:E55_Type>
<crm:P127_has_broader_term>
<crm:E55_Type>déesse</crm:E55_Type>
<crm:E55_Type>divinité</crm:E55_Type>
</crm:P127_has_broader_term>

```

```

</crm:P2_has_type>
</crm:E28_Conceptual_Object>
</crm:P62_depicts>
<crm:P108B_was_produced_by> <!--événement "fabrication de l'objet" --
>
<crm:E12_Production>
<crm:P32_used_general_technique>
<crm:E55_Type>ronde bosse</crm:E55_Type>
<crm:P127_has_broader_term>
<crm:E55_Type>sculpture</crm:E55_Type>
</crm:P127_has_broader_term>
</crm:P32_used_general_technique>
<crm:P4_has_time-span> <!--datation formulée par fourchette
chronologique ) -->
<crm:E52_Time-Span>
<crm:P82_at_some_time_within>
<crm:E61_Time_Primitive>
<start>-150</start>
<end>-100</end>
</crm:E61_Time_Primitive>
</crm:P82_at_some_time_within>
</crm:E52_Time-Span>
</crm:P4_has_time-span> <!--datation formulée par appellation
traditionnelle -->
<crm:P10_falls_within>
<crm:E4_Period>
<crm:P1_is_identified_by>
<crm:E49_Time_Appellation>époque hellénistique récente</
crm:E49_Time_Appellation>
</crm:P1_is_identified_by>
</crm:E4_Period>
</crm:P10_falls_within>
<crm:P14_carried_out_by> <!-- artiste -->
<crm:E39_Actor>atelier grec</crm:E39_Actor>
</crm:P14_carried_out_by>
</crm:E12_Production>
</crm:P108B_was_produced_by>
...
<crm:P12B_was_present_at><!--informations sur la découverte -->
<crm:E7_Activity>
<crm:P2_has_type>
<crm:E55_Type>découverte</crm:E55_Type>
<crm:P7_took_place_at>
<crm:E53_Place>
<crm:P87_is_identified_by>
<crm:E44_Place_Appellation> Fiumicino</crm:E44_Place_Appellation>
<crm:E47_Spatial_Coordinates>
<latitude>41.7667</latitude>
<longitude>12.2333 </longitude>
</crm:E47_Spatial_Coordinates>
</crm:P87_is_identified_by>
<crm:P89_falls_within>

```

```

<crm:E53_Place>Italie</crm:E53_Place>
</crm:P89_falls_within>
</crm:E53_Place>
</crm:P7_took_place_at>
...
<crm:P130B_features_are_also_found_on><!--reproductions numériques
de l'objet -->
<crm:E38_Image>
<crm:P149_is_identified_by>
<crm:E75_Conceptual_Object_Appellation>URL de l'image (http://
www.limc-france.fr...1DE19367)</
crm:E75_Conceptual_Object_Appellation>
</crm:P149_is_identified_by>
<crm:P104_is_subject_to>
<crm:E30_Right>
<crm:P3_has_note>
<crm:E62_String>notice de copyright</crm:E62_String>
</crm:P3_has_note>
</crm:E30_Right>
</crm:P104_is_subject_to>
</crm:E38_Image>
</crm:P130B_features_are_also_found_on>
...
<crm:P70B_is_documented_in><!-- bibliographie, webographie...
concernant l'objet -->
<crm:E31_Document>Klein, W., Praxiteles (1898) 262-264</
crm:E31_Document>
<crm:E31_Document>Furtwängler, Beschreibung der Glyptothek (1910)
266 n° 258</crm:E31_Document>
<crm:E31_Document>Vierneisel-Schlörb, KatSkulptMünchen II 323-336 n
° 31, fig. 158-164</crm:E31_Document>
<crm:E31_Document>LIMC, vol. II, Aphrodite 399*</crm:E31_Document>
<crm:E31_Document>http://...</crm:E31_Document>
</crm:P70B_is_documented_in>
</crm:E24_Physical_ManMade_Thing>

```

- 14 Dans ce cas, l'avantage du CIDOC CRM par rapport au DC réside dans le fait qu'il permet d'introduire plus de finesse et de sémantique dans la structuration de l'information et donc de traiter davantage d'informations, de données. L'emploi du CRM permet de structurer une information plus riche et plus détaillée : différencier l'objet documentaire (statue) et sa reproduction (photo), différencier/séparer/isoler diverses notions de lieux pertinentes (lieu de fabrication de l'objet, lieu de trouvaille, lieu de conservation actuel, lieux antérieurs de conservations/collections...), de date (date de création de la statue, date d'entrée au musée, date de l'enregistrement). Sa principale différence par rapport au Dublin Core est qu'il permet de distinguer la description de la ressource de celle de l'objet dont parle la ressource.
- 15 La structuration plus fine élargit les possibilités d'utilisation du fichier de données produit. L'utilisation de la donnée a une forte influence sur le modèle de structuration de l'information : si les métadonnées DC peuvent correspondre à une interrogation (type champ Google), un schéma plus fin permet l'élaboration d'une interface Web de recherche différenciant davantage de critères de recherche (facettes...). Un outil

complexe servant à recueillir des informations variées et détaillées (une base de données relationnelle, un fichier RDF...) peut également reposer sur l'emploi de cette norme qui semble adaptée à cette problématique.

- 16 Quelques schémas types fondés sur le CIDOC CRM existent, par exemple LIDO (*Lightweight Information Describing Objects*) destiné aux collections de musées.

Communauté en France

- 17 La documentation mise en ligne sur le site du CIDOC CRM augmente (communications, schémas, liens vers applications, etc.) mais on y constate peu de références françaises. Des travaux ont cependant été amorcés par Pierre Le Bœuf à la Bibliothèque Nationale de France (BnF) et des réflexions sont en cours au Ministère de la Culture. Quant à la communauté scientifique française, elle ne semble pas avoir tellement utilisé ou même travaillé sur ce modèle. Pourtant, il semble interpeller de plus en plus d'acteurs de la recherche, notamment dans les champs disciplinaires liés à l'art et à l'archéologie, comme en témoignent les interrogations (« Vous allez parler du CIDOC CRM ? Justement, je me posais des questions à son propos... ») exprimées lors de rencontres scientifiques, suivies d'exclamations (« Ah ! Mais cela pourrait résoudre mon problème !), vite suivis de nouvelles d'interrogations (« Mais comment ça marche ? »).
- 18 Parmi les initiatives émanant du monde de la recherche, on peut signaler l'élaboration, actuellement, d'un schéma pivot pour un portail dédié aux objets de l'Antiquité (programme IDA avec le TGE Adonis). L'équipe française du LIMC (CNRS UMR7041-ArScAn), impliquée dans ce programme, a également participé à la création du portail international CLAROS dont l'interopérabilité est fondée sur l'emploi du CIDOC CRM et dont le schéma RDF est présenté et commenté sur un site dédié, CLAROSwiki. Sur ce dernier, on peut consulter des exemples de fichiers RDF avec les données/valeurs/instances.
- 19 Si l'on souhaite voir un exemple de site fondé sur l'emploi du CIDOC, on peut aller consulter la réalisation allemande Arachne/iDAI
- 20 On pourrait soumettre l'idée de la création d'une communauté d'intérêt qui fait particulièrement défaut en France. Un accompagnement ou un espace d'échange permettrait de partager des expériences dans ce domaine mais aussi entre spécialistes du domaine, spécialistes du traitement de l'information et informaticiens. Rosemonde Letricot (Institut de recherche et d'histoire des textes, IRHT) a approché le CIDOC dans le cadre du projet IDA. Elle porte actuellement une réflexion autour de l'utilisation d'ontologies pour rendre compte de la complexité du livre médiéval à partir des données de la recherche. Son expérience d'utilisation du CRM l'amène au constat d'un besoin d'un apprentissage de la norme (explications, documents pratiques, schémas, publications de modèles, de schémas incluant les instances/valeurs/données...) et d'accompagnement des utilisateurs. Elle suggère de créer un groupe d'utilisateurs et d'échange pour ne pas démultiplier les créations de modèles ontologiques. Elle souligne l'intérêt de l'ontologie pour des données hétérogènes car celle-ci permet de recontextualiser l'information dans son domaine et de mettre à plat les postulats implicites de ce dernier. Quand on s'attaque à une ontologie, il devient nécessaire d'être critique par rapport aux données car il faut tout formaliser : les concepts, les caractéristiques de ces concepts, leurs relations, leurs rôles, leurs conditions et leurs restrictions.

- 21 Il faut également se demander dans quel but on structure les données (réflexion autour de l'application qui va restituer l'information).
- 22 Il serait aussi intéressant de se positionner par rapport à d'autres propositions, et d'autres normes comme EDM d'Europeana.

Discussion

- 23 La question de la place de l'historien dans la structuration des données et l'implication du chercheur dans la construction d'ontologies a été évoquée.
- 24 **Q :** Le CIDOC CRM peut-il être utilisé dans le cadre d'un projet sur une collection pour créer des liens entre les entités dans les textes et les collections et les outils. Comment au musée de Nuremberg peut-on adapter le CRM, et mélanger des ontologies ?
- 25 **R :** Le Göttingen Centre for Digital Humanities (GCDH) fédère les ontologies produites par différentes institutions².
- 26 **Q :** Dans le cadre d'études architecturales et l'instauration d'un vocabulaire spécifique à la période moderne, le CRM peut-il traiter la gestion des droits de diffusion des prises de vues ? ou répondre à des problématiques concernant l'interopérabilité et la conservation sur le long terme (pérennité), et par rapport aux autres vocabulaires, normes, etc. ?
- 27 **R :** Le CIDOC CRM permet d'insérer de façon précise et détaillée les informations concernant les droits (détenteur des droits, notice de *copyright*... en rattachant précisément ces informations aux documents concernés, par exemple objet documentaire principal [statue], ses reproductions photographiques signalées...). Si on crée des bases de données avec des ontologies spécifiques il est ensuite nécessaire de faire des mappings pour les rendre interopérables.
- 28 Il ne faut pas générer de la différenciation là on cherche l'interopérabilité. Il est nécessaire d'articuler les différents modèles. Par exemple, il faut réfléchir à l'articulation entre ISO 21127 et la norme ISO de gestion des vocabulaires (ISO 25964). La classe E55 Type, utilisée pour porter les valeurs issues de nomenclatures ou de vocabulaires normalisés, doit être rapprochée de la classe ThesaurusConcept d'ISO 25964.
- 29 Au ministère de la Culture et de la communication, un travail est en cours pour articuler ensemble des fonds issus de domaines très divers. Le dialogue passe aussi par le partage de vocabulaires. Un schéma de données (ou une ontologie) décrit bien comment on doit manipuler les concepts (les classes), mais ne dit pas avec quelles valeurs on peut ou on doit les instancier. Il est donc important d'avoir des vocabulaires partagés pour renforcer l'interopérabilité sémantique.
- 30 **Q :** A-t-on connaissance d'ontologies utilisées en SHS ? Le CIDOC CRM permet-il de traiter l'histoire sociale ? Dans le cadre d'une plateforme collaborative et une base de données ouverte³, une équipe lyonnaise a, depuis cinq ans, mis en place une plateforme multi-utilisateurs et les utilisateurs ont créé une ontologie du domaine, en dialogue avec les informaticiens, et ont choisi un modèle ouvert : objets/rôles/informations. Les ontologies sont relationnelles et non RDF car il y a une prédéfinition des prédicats. Ils auraient besoin d'un mapping pour exposer ces données dans une ontologie commune pour partager l'information via le web sémantique (schéma pour l'interopérabilité). Le CIDOC semble ajouter une couche de description (contextualisation), et avoir des possibilités d'extension.

- 31 **R** : C'est un problème pour la description des textes qui sont très importants dans le travail de l'historien (le CRM propose juste un conteneur).
- 32 Concernant la dimension sociale, les notions d'acteur et d'appartenance à un collectif sont développées dans le CIDOC CRM, mais cela peut-il répondre aux besoins ? La notion d'évènement permet de rendre compte du cycle de vie des objets. Les évènements peuvent être typés (par exemple par la classe E55 Type), comme chaque entité/classe du modèle. Le CRM permet de formuler finement des concepts, notamment pour exprimer l'évènement « fabrication de l'objet » et les informations qui s'y rattachent (lieu de fabrication, artiste, matériaux).
- 33 Le CIDOC CRM est organisé en hiérarchie de classes et de propriétés, avec la notion d'héritage. En particulier, la classe principale (d'entrée) « E1 Entité » porte la propriété « P2 a pour type » (qui s'associe avec la classe E55 Type), qui est donc héritée par toutes les autres classes.
- 34 Si le CRM permet d'exprimer des informations complexes ou riches et d'approfondir leur granularité, certains choix de mise en œuvre peuvent traiter ces mêmes informations en privilégiant l'utilisation, plus facile, d'éléments comme la classe E55 Type (voir le document de la norme) pour catégoriser directement les concepts traités.
- 35 Par exemple, dans le cas du portail CLAROS, le schéma est simple et « plat ». Il repose principalement sur l'utilisation de la classe E55 type. La notion « artiste » est réglée avec la création d'un type E55 = artiste⁴ alors qu'elle aurait pu s'insérer dans une structure réunissant de façon logique plusieurs interactions classe/propriété appropriées (encapsulation d'informations) plaçant, par exemple, l'artiste comme intervenant dans un évènement défini comme « processus de création », etc. Cela pose le problème de l'interopérabilité. Si on propose plusieurs schémas pour une même notion, est-ce vraiment interopérable ? Ce choix répond bien au besoin d'interopérabilité du portail : les ressources numériques des partenaires « interopèrent », un nouveau partenaire peut facilement mapper ses données..., mais le nouveau gisement de ressource répond-il vraiment à ce que l'on attend de l'interopérabilité, du web sémantique, etc. ?
- 36 Faut-il privilégier des choix, préconiser des utilisations de classes ou de propriétés ?
- 37 **Q** : Dans le cas d'une plateforme mise en place autour de fichiers son, avec un vocabulaire propre à la discipline, où les notions d'œuvre, de créateurs des ressources, etc. sont importantes, quel format choisir pour les métadonnées ?
- 38 **R** : Dans le cadre du projet européen Caspar sur l'archivage (Jérôme Barthélémy, IRCAM) l'ontologie CIDOC CRM a été utilisée pour l'archivage et la pérennisation, en vue de restituer complètement l'œuvre.
- 39 **Q** : Dans quelle mesure le CIDOC CRM peut-il être souple dans son utilisation et applicable à d'autres types de sources ? Y-a-t-il une norme commune pour les objets hétérogènes ?
- 40 **R** : Il n'y a pas de réponse. La norme prétend couvrir un large spectre de domaines. Il faut explorer la norme pour voir les besoins d'évolution, la confronter à des cas spécifiques. Il y a une nécessité à fédérer les acteurs, qui pourraient alors devenir une force de proposition pour le groupe de réflexion du CRM SIG.
- 41 **Q** : Y a-t-il une modélisation des ontologies ?
- 42 **R** : Plusieurs classes peuvent être adaptées à une même notion. Il est parfois difficile de choisir, de percevoir toutes les nuances de la définition. On peut opter pour une classe dont la définition est plus large. Il y a un jeu dans la hiérarchie des classes, et il faut

rappeler qu'une sous-classe respecte la définition de ses classes supérieures et hérite de ses propriétés. Cela permet une définition plus ou moins fine des concepts, et aussi des définitions plus complexes.

- 43 Il est noté l'importance de la collaboration avec les informaticiens. Le travail/dialogue avec eux est essentiel pour la mise en pratique et cet échange profite à l'élaboration de la norme, du schéma, et de la façon de l'appliquer concrètement. C'est un enrichissement mutuel, pour les participants, le projet mais aussi la norme.
- 44 L'idée de privilégier la simplicité dans la conception de schémas, sans pour autant renoncer à la spécificité et la richesse du CIDOC CRM, pourrait faciliter l'émergence de schémas communs à des besoins, des disciplines (interopérabilité, web sémantique, émergence de besoins et discussion au sein de la communauté d'intérêt...). Il faut trouver un équilibre dans les modèles, car on ne peut pas tout traiter. Même si la norme le permet, faut-il aller vers trop de complexité ? Le sens critique est nécessaire sur des concepts que recouvre l'information à traiter, par exemple les notions de lieux, de datation, d'institutions, de propriété, de localisation, de nommage, etc. Se confronter à la norme oblige à réfléchir davantage et plus précisément aux informations/concepts que l'on souhaite traiter. Cela déclenche l'évaluation critique sur sa propre approche. Une structuration évidente depuis des années peut apparaître inadaptée ou inadéquate lors de ce processus de critique, présenter des distorsions, des lacunes. Cette confrontation avec les définitions du CIDOC CRM est profitable et intéressante à faire même s'il n'y a pas de réalisation pratique prévue. Il se dégage également des notions plus délicates à traiter et à rendre interopérables. Celles-ci apparaissent notamment lorsque l'on élabore un schéma pour un portail visant à rendre interopérables plusieurs ressources numériques semblant pourtant proches dans leur conception : par exemple les datations, les styles, les aires « géo-culturelles » de création, etc.
- 45 Son utilisation est encore peu répandue. Du point de vue d'un chercheur, on ressent la nécessité de réfléchir sur les normes, car il n'y a pas de norme qui permette de gérer l'ensemble des informations produites par la recherche historique.
- 46 Il serait intéressant de voir les différents traitements ou schémas proposés et les vocabulaires utilisés. Comment les choisir et comment formuler dans le schéma ?
- 47 On pourrait fédérer un réseau autour du CIDOC CRM, comme la TEI a su le faire, en associant des spécialistes des sciences de l'information et de l'informatique en plus des spécialistes des champs disciplinaires traités (équipes de recherche).
- 48 **Q** : Est-ce que le Louvre utilise cette norme ?
- 49 **R** : Le CIDOC CRM est peu utilisé dans les musées français. Une des explications parfois avancées est que ce modèle émane de la communauté des documentalistes et non de la communauté des conservateurs qui n'y adhèrent pas, peut-être en raison de sa complexité. Dans le cadre de projets dans la nébuleuse Europeana (Athena), on assiste à un rapprochement de plusieurs modèles dans le schéma XML LIDO (CIDOC CRM, Spectrum (gestion des collections), et Museumdat⁵).
- 50 **Q** : Est-il difficile de s'approprier cet outil, cette norme ?
- 51 **R** : Il y a besoin d'un dialogue pour lever des ambiguïtés possibles et partager les pratiques. À titre d'expérience, les données du LIMC (<http://www.limc-France.fr>) sont dans une base de données dont la première version remonte à 1981 (c'est-à-dire au début du programme d'informatisation). Dans ce cas, la base LIMC n'a pas été remise à plat,

mais des mapping ont été créés, d'une part pour préciser les notions traitées – la norme CIDOC CRM devient alors un document de référence pour les définitions, un outil « fort » pour documenter la ressource numérique – et des mapping plus succincts destinés à des extractions pour différents projets (portails IDA et CLAROS, Projet d'OAI-PMH). On peut aussi envisager d'élaborer directement une ressource sur l'emploi de la norme (norme, puis modélisation puis schéma d'implantation, etc.). L'opération de mapping à partir de données existantes n'est peut-être pas aussi facile que les concepteurs du modèle l'ont parfois dit. La communauté aurait besoin d'exemples et de modèles d'utilisation. On rencontre aussi des problèmes pour la traduction d'une norme. Par exemple, on est en attente de la nouvelle version ISO et de sa traduction française, bien qu'une traduction française de l'ancienne version soit déjà disponible.

- 52 **Q** : Dans le contexte de la réalisation d'un programme pédagogique en archéologie où l'on souhaite imbriquer virtuel et réel (terrains, objets) avec un objectif pédagogique et créer des scénarios, est-il possible d'utiliser cette ontologie pour permettre au public de renseigner des informations sur les objets ?
- 53 **R** : La complexité des ontologies impose la mise en place d'interfaces utilisateurs qui masquent cette complexité et qui permettent des restitutions très riches (exemples : timeline, géolocalisation, facettes d'interrogation). Le modèle CIDOC est présenté comme « ouvert » et permet des discours contradictoires sur un même objet.
- 54 **Q** : Comment gérer le modèle dynamique ?
- 55 **R** : C'est une question d'interface : on doit définir les besoins des utilisateurs et se limiter à présenter les « champs », les propriétés qu'ils auront à renseigner.

Conclusions, perspectives

- 56 Il est important de pouvoir présenter et faire connaître cette norme. Il faut aussi structurer un groupe de réflexion sur les modèles d'ontologies ce qui permettrait de faire remonter les usages et les difficultés.
- 57 Il ressort de la discussion qu'il serait bien de faire émerger une communauté d'intérêt en France, mais également ouverte à des collègues étrangers, dans un esprit trans-disciplinaire et inter-institutionnel.
- 58 Pour commencer, on pourrait créer une liste de diffusion et, dans ce but, récupérer les adresses mails des personnes intéressées. Dans un deuxième temps, on pourrait peut-être songer à un carnet de recherche par exemple sur Hypotheses.org pour accompagner cette communauté et lui donner une vitrine dynamique permise par le carnet.

NOTES

1. CIDOC /Comité international pour la documentation de l'Icom.
2. <http://erlangen-crm.org/>. Le CRM Erlangen / OWL est une implémentation OWL-DL 1.0 du Modèle Conceptuel de Référence du CIDOC (CIDOC CRM).

3. (LARHRA – Université de Lyon – Projet SyMoGIH – http://larhra.ish-lyon.cnrs.fr/Pole_Methodes/symogih_accueil_fr.php)
 4. http://explore.clarosnet.org/wiki/index.php?title=CIDOC_CRM_Objects
 5. Harvesting Format for Providing Core Data from Museum Holdings ;LIDO (LIDO is an XML harvesting schema) ; <http://network.icom.museum/CIDOC/working-groups/data-harvesting-and-interchange/what-is-lido/> ; Norme Spectrum http://www.pro.rcip-chin.gc.ca/GetForumRecord.do?type=sd&lang=fr&id=FORUM_23068
-

RÉSUMÉS

L'idée de l'atelier est de présenter et de discuter d'une norme pour structurer et encoder les données sur les œuvres matérielles et immatérielles du patrimoine culturel. Norme qui concerne des documents hétérogène ce qui la rend intéressante pour servir de modèle pivot, mais peu connue en France. Parler de l'émergence d'une communauté d'intérêt et de partage d'expériences.

Cette norme est surtout utilisée dans le monde des musées, pas tellement dans la communauté scientifique

Introduction au Modèle Conceptuel de Référence CIDOC (ISO 21127:2006), émanant du monde des musées (ICOM) et utilisé pour structurer, encoder les données sur les œuvres, matérielles ou immatérielles, du patrimoine culturel (*Cultural Heritage*): présentation et discussion sur les domaines d'application (musées, sites archéologiques, gestion de collections et de fouilles, corpus d'objets, etc.), les notions clés (objet/concept, événement/temps, lieu, intervenant/agent de l'événement...) et leurs interactions.

INDEX

Mots-clés : ontologie, patrimoine culturel, code, programmation, collection