



HAL
open science

L'ontologie CIDOC CRM appliquée aux objets du patrimoine antique

Anne-Violaine Szabados, Rosemonde Letricot

► **To cite this version:**

Anne-Violaine Szabados, Rosemonde Letricot. L'ontologie CIDOC CRM appliquée aux objets du patrimoine antique. 3e Journées d'Informatique et Archéologie de Paris - JIAP 2012, Jun 2012, Paris, France. halshs-00752996

HAL Id: halshs-00752996

<https://shs.hal.science/halshs-00752996v1>

Submitted on 19 Feb 2013

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.

L'ontologie CIDOC CRM appliquée aux objets du patrimoine antique

Anne-Violaine SZABADOS et Rosemonde LETRICOT¹

L'intérêt porté, à l'occasion des *Journées d'Informatique et Archéologie de Paris*, au rôle des métadonnées et référentiels standards dans l'interopérabilité des données, nous permet de poursuivre le discours, amorcé lors des *JIAP* précédentes, sur l'utilisation d'ontologies², en particulier sur l'application de la norme de référence CIDOC CRM à des documents antiques. La capacité de cette norme à couvrir la diversité des notions attachées à un objet du patrimoine culturel en fait un outil possible pour rassembler des informations aussi hétérogènes que celles relatives aux documents archéologiques ou muséaux. Quelques exemples appropriés nous permettront de rappeler que l'association de choix techniques, d'une structuration fine de l'information et de l'emploi de vocabulaires de référence favorise la sémantisation et l'interopérabilité des données.

La ressource au cœur du paradigme numérique.

Avec l'arrivée de l'Internet, la description des ressources issues du domaine patrimonial et de la recherche s'est complexifiée et internationalisée. Les applications web permettent en effet une visibilité accrue des contenus et une valorisation du travail de contextualisation de l'information produite par les institutions patrimoniales, muséales ou de recherche. Auparavant ces applications étaient principalement destinées à des utilisateurs « humains » qui devinaient, par l'utilisation de moteurs de recherche, le contenu potentiel des bases de données, mais aujourd'hui les robots sont techniquement capables d'interpréter ces informations pour les redéployer dans d'autres services. La « page » web en tant que document laisse progressivement place à un espace calculé où les données sont générées à la volée selon les paramètres de l'utilisateur. Quotidiennement nous manipulons donc de plus en plus de contenus computationnels plutôt que de contenus documentaires³.

Or, si l'humain est capable de déduire les relations implicites entre plusieurs informations relatives à un même objet, les machines ont, elles, besoin d'un balisage et d'une sémantisation de l'information pour pouvoir raisonner. C'est le rôle que jouent les métadonnées dans l'environnement du numérique, leur finalité étant de décrire la donnée pour améliorer son indexation et la valoriser auprès des moteurs de recherche mais aussi de rendre possible l'interopérabilité entre différents systèmes. L'interopérabilité est une notion-clé dans ce qui est appelé aujourd'hui le « Web de données » car c'est en libérant les données du carcan des applications qu'il est possible de naviguer « au flair » et de rebondir sur des informations

¹ A.-V. SZABADOS (chargée de la gestion, l'interopérabilité et la valorisation des ressources numériques, CNRS UMR7041 – ARSCAN, équipe LIMC / Fondation internationale pour le LIMC. www.mae.u-paris10.fr/arscan/-ArScAn-Lexicon-Iconographicum-.html) : anne-violaine.szabados@mae.cnrs.fr. R. LETRICOT (chargée de projet numérique, IRHT, Service Développement et Interopérabilité. www.irht.cnrs.fr) : rosemonde.leticot@cnrs-orleans.fr.

² SZABADOS, A.-V., « Du Système documentaire du LIMC au portail CLAROS. Interopérabilité et optimisation de l'information archéologique grâce à l'usage de normes », p.11-25 ; CIEZAR, P., « Acquisition et enregistrement de données à grande échelle en archéologie préventive » p.224-226, in GILIGNY, F./ COSTA, L./ DJINDJIAN, R./ CIEZAR, P./ DESACHY, B. (dir.), *Actes des 2^e Journées d'Informatique et Archéologie de Paris – JIAP 2010 (Paris, 11-12 juin 2010), Archeologia e calcolatori*, suppl. 3, 2012.

³ DELAFORGE, N./ GANDON, F./ MONNIN, A., « L'avenir du web au prisme de la ressource », in *Le document numérique à l'heure du Web de données*, INRIA, 2012, p.229-252.

complémentaires à la ressources initiale. Mais, en se libérant des applications web, l'information perd sa contextualisation et il est alors indispensable de la documenter par l'ajout de métadonnées qui l'accompagneront entre les systèmes via le protocole HTTP.

Les standards jouent un rôle essentiel dans ce contexte d'interopérabilité car ils garantissent l'unicité des syntaxes et vocabulaires dans les échanges entre systèmes. Qu'ils soient HTML, XML ou encore RDF, selon les technologies utilisées, ils unifient la manière d'exprimer la donnée.

Le protocole OAI-PMH, fondé par l'Open Archives Initiative en 2001, est la première étape vers l'ouverture des données sur le web. Il autorise l'échange de métadonnées par la construction d'entrepôts que des fournisseurs de service moissonnent afin de réemployer ces métadonnées dans d'autres programmes. Ce protocole est par exemple utilisé pour des plateformes d'agrégation de contenus scientifiques comme *Isidore*⁴ ou celles recensant les collections culturelles comme *Europeana*⁵. L'OAI-PMH repose sur l'utilisation du schéma de métadonnées Dublin Core⁶ dans la perspective d'une interopérabilité limitée à quelques éléments communs⁷. En utilisant le protocole OAI-PMH couplé avec des métadonnées exprimées en DC, il est alors possible d'« exposer » ses données sur le web pour permettre à des moissonneurs d'indexer leurs métadonnées, voire de les enrichir.

Avec un exemple concret issu du patrimoine culturel, nous verrons que le Dublin Core reste insuffisant pour décrire la complexité d'un objet, car il ne vise pas à proposer une structure exhaustive de description mais plutôt un schéma pivot général d'interopérabilité. Pour couvrir la richesse sémantique d'un tel objet et la diversité de l'information qui lui est attachée, il existe des ontologies informatiques destinées à donner plus de sens aux balises et à contextualiser davantage les informations relatives à un domaine.

Le terme d'ontologie informatique a été emprunté au courant philosophique de l'Ontologie qui aborde l'étude des propriétés générales de l'être. D'abord utilisées en Intelligence Artificielle avant de s'étendre à d'autres champs de l'informatique, les ontologies définissent un vocabulaire commun à un domaine en structurant l'information par des ensembles de concepts⁸. Elles servent ainsi à expliciter tout ce qui est implicite dans un domaine en décrivant les concepts eux-mêmes, leurs caractéristiques, leurs spécificités, leurs relations, et les conditions ou restrictions qui peuvent s'appliquer.

L'ontologie informatique est une aide précieuse non seulement pour exprimer et analyser des données hétérogènes dans des applications logicielles, mais aussi pour les partager et les documenter, car tout ce qui est « évident » dans la structuration de l'information d'une ressource devient explicite par les définitions mêmes des classes et propriétés utilisées. Mais l'interopérabilité des données n'est pas la seule raison motivant l'utilisation d'ontologies ; elles peuvent également s'insérer à différents stades d'un projet de création de systèmes d'informations, que ce soit pour établir une nomenclature commune entre différentes disciplines ou pour servir de socle lors de comparaison de jeux de données pour un même domaine.

⁴ <http://rechercheisidore.fr>

⁵ www.europeana.eu

⁶ Norme ISO 15836:2009 (révision d'ISO 15836:2003) : *Information et documentation – L'ensemble des éléments de métadonnées Dublin Core*.

⁷ Le Dublin Core (DC) possède plusieurs niveaux de raffinement : soit le Dublin Core Simple exprimant quinze éléments (titre, auteur, éditeur, sujet, droit, etc.), soit le Dublin Core qualifié nuançant certaines balises (par exemple, *dcterms:isPartOf* pour signifier une relation hiérarchique entre deux données).

⁸ LEBOEUF, P., « Le modèle CRM pour la documentation muséographique », in *Journée d'étude de l'ADBS – La modélisation : pourquoi l'intégrer dans les systèmes d'information documentaire ?*, 2003. www.cidoc-crm.org/docs/adbs_crm.doc.

L'implémentation des ontologies dans des systèmes d'informations en vue d'une ouverture des données sur le web va bien au-delà de l'utilisation de l'OAI-PMH et du Dublin Core. En effet, l'expression des données en RDF (Resource Description Framework) associée à l'emploi de chaînes de traitement informatique visant à faire correspondre, voire enrichir, les données par l'utilisation du vocabulaire et des dépendances liés à l'ontologie⁹, permet d'envisager le Web de données.

Les données du patrimoine culturel et le Dublin Core : le cas d'une statue composite.

Le Dublin Core, même dans sa version qualifiée, ne permet de traiter qu'une partie des informations liées à un objet du patrimoine archéologique ou muséal, et cela en appauvrissant certaines des notions sélectionnées. À titre d'exemple, une notice sur un objet de musée a été rédigée de façon à faire ressortir les limites de l'emploi des quinze éléments de description *Dublin Core Simple* pour rassembler, et surtout restituer, les informations riches et variées relatives à l'œuvre – une statue antique conservée au Musée du Louvre et représentant à l'origine Minerve¹⁰ puis, au XVIII^e s. et avant restauration, la personnification de Rome – mais aussi à l'enregistrement qui lui est consacré dans le corpus numérique LIMCicon (*LIMC-France*)¹¹, voire à sa photographie numérique (fig. 1).

<pre> <dc:title>Statue de Minerve</dc:title> <dc:title>Statue de Rome</dc:title> <dc:title>Paris, Musée du Louvre MR 341, Ma 1056</dc:title> <dc:creator>atelier romain?</dc:creator> <dc:creator>inconnu</dc:creator> <dc:creator>Francuccio Francucci</dc:creator> <dc:creator>LIMC-France</dc:creator> <dc:subject>Minerve ; Athéna ; Rome ; Antiquité ; sculpture ; archéologie...</dc:subject> <dc:description>statue en porphyre de Minerve</dc:description> <dc:description>statue en porphyre et bronze de Minerve</dc:description> <dc:description>statue en porphyre et bronze de Rome</dc:description> <dc:description>Personnage féminin assis, le bras droit levé</dc:description> <dc:publisher>LIMC </dc:publisher> <dc:contributor> LIMC-France</dc:contributor> <dc:date>Empire romain</dc:date> <dc:date>II^e s. apr. J.-C.</dc:date> <dc:date>XVII^e-XVIII^e s. apr. J.-C.</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://www.limc-france.fr/objet/14804</dc:identifiant> <dc:source>http://www.limc-france.fr</dc:source> <dc:language>fr</dc:language> <dc:relation> www.limc-france.fr/image/... OFR19638</dc:relation> <dc:coverage>Rome; Paris, Louvre; Antiquité; Empire romain ; II^e s. apr. J.-C.; XVII^e-XVIII^e s. apr. J.-C. ; 2011/08/20; 2012/09/27</dc:coverage> <dc:rights>domaine public</dc:rights> <dc:rights>LIMC-France</dc:rights> </pre>	
	 Cliché: A.-V. Szabados

Fig. 1 : notice descriptive en *Dublin Core Simple* d'une statue de Minerve du Louvre (MR 341, Ma 1056)

⁹ Par exemple, le portail international CLAROS (www.clarosnet.org), auquel participe le LIMC, est fondé sur l'utilisation du RDF et d'un schéma pivot construit sur l'ontologie CIDOC CRM. Cf. CLAROSwiki (<http://explore.clarosnet.org/wiki>) ; SZABADOS 2012, *supra* n.2, p.13-19.

¹⁰ Paris, Musée du Louvre inv. MR 341, usuel Ma 1056. DELBRÜCK, R., *Antike Porphyrwerke*, 1932, p.70-71, pl.18 ; SZABADOS, A.-V., LIMCicon ID 14804, in *LIMC-France*. www.limc-france.fr/objet/14804 (Consultation 29/10/2012).

¹¹ LIMCicon est diffusé sur *LIMC-France*, site web de l'équipe française de la Fondation Internationale pour le LIMC : <http://www.limc-france.fr>

Dans le cas de cette statue pourraient être privilégiées, soit les informations relatives au document numérique – l’enregistrement – traitant de l’objet (en caractères gras dans l’exemple), soit celles concernant l’œuvre elle-même, ou encore le thème représenté¹². Ici, les balises *dc:date* reçoivent des notions différentes sans les différencier, en l’occurrence les périodes de fabrication et de modification de la statue (*II^e s. apr. J.-C. ; Empire romain ; XVII^e-XVIII^e s. apr. J.-C.*), la date de création (2011/08/20) et la date de modification de l’enregistrement (2012/09/27)¹³. On aurait également pu ajouter la date de la découverte de l’objet lors de la fouille archéologique, ou encore celle de son entrée au Louvre. *Francuccio Francucci*, qui l’a rénovée et complétée, est proposé comme *dc:creator*, ainsi que le terme *inconnu*, qui pourrait être automatiquement extrait du champ « créateur de l’objet » (artiste antique) de la ressource numérique d’origine. L’auteur de l’enregistrement numérique – l’équipe *LIMC-France* – pourrait également être le *dc:creator*. Quant à la balise *dc:coverage* (couverture), elle mêle plusieurs informations temporelles et géographiques, la notion de couverture géographique pouvant concerner les lieux de création et de trouvaille de l’objet aussi bien que ses divers lieux de conservation, collections passées et actuelles¹⁴. Si l’on se limite à la statue antique, il est également impossible, avec *dc:title* et *dc:description*, de distinguer les divers états de la statue et les deux identifications : *Minerve* et *Rome*.

Cette notice et les pratiques témoignent d’une grande latitude dans l’application du DC¹⁵. Le choix des données à insérer dépendant fortement du contexte de production et d’utilisation ou de réutilisation des données, la nature et la granularité de l’information sur le document peuvent être bien différentes dans les métadonnées d’une photographie d’une banque d’images généraliste et dans celles, par exemple, d’une ressource numérique élaborée par des spécialistes de la statuaire antique. L’éventail des éléments couramment envisagé par les spécialistes de l’étude des objets culturels est ici réduit ; les notions retenues sont mises au même niveau, aplanies, dans le schéma de métadonnées DC. Les requêtes sur cette notice sélectionneront bien l’enregistrement demandé mais la restitution de l’information sera-t-elle détaillée assez précisément pour être utilisée par l’internaute ou par un autre système informatique ?

... et comment formuler ceci :

l’objet a été créé au II^e s. apr. J.-C.
... était en porphyre et représentait Minerve
... provient du jardin de San Martino de' Monti à Rome
... fut acquise en 1646 par le cardinal Mazarin ... transférée à Paris, Collection Mazarin,
puis dans la Collection de Louis XIV (en 1665) ... placée dans les jardins du Trianon ...
saisie à la Révolution
... complétée par du bronze et rénovée par Francuccio Francucci au XVII^e s.,
... fut transformée en Rome au XVIII^e s., par le changement de sa tête et de son bras
gauche
... a été étudiée par X qui l’a associée à la gravure réalisée en 1681 par Étienne
Baudet¹⁶, ... a été comparée à une autre statue par Y ...

¹² Dans un schéma destiné à un entrepôt OAI-PMH fondé sur l’utilisation du DC non qualifié, plusieurs champs DC pourraient être réservés aux informations sur le document numérique ou nécessaires au moissonnage des données et à leur gestion dans l’entrepôt.

¹³ Même en utilisant *dcterms:created* pour la date de création, celle de l’objet ne serait pas différenciée de celle de l’enregistrement.

¹⁴ Ces ambiguïtés ne seraient pas davantage levées si on avait isolé les données géographiques de <coverage> en utilisant le raffinement du DC qualifié pour la « couverture spatiale ».

¹⁵ Cf. le résumé de l’ISO 15836:2009.

www.iso.org/iso/fr/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=52142.

¹⁶ BAUDET, E., *Tableau du Cabinet du Roi, statues et bustes antiques des Maisons Royales*. t.1, 1681, pl.19 (estampe conservée au Musée national des châteaux de Versailles et de Trianon, inv. A15-133).

Si le référentiel de métadonnées Dublin Core est adapté aux informations sur la ressource numérique elle-même et peut rassembler une sélection d'éléments pertinents pour des besoins particuliers, par exemple pour un entrepôt OAI ou pour des requêtes posées lors de recherches généralistes, il faudrait recourir à des standards développés pour des domaines spécifiques aux GLAM¹⁷ afin d'élaborer des ressources numériques ou des schémas de métadonnées plus riches sémantiquement et mieux adaptés à ce type de notions. L'ontologie CIDOC CRM permettrait de modéliser ces connaissances et de restituer des données en tenant compte de l'évolution de l'œuvre, de son parcours, de son contexte historique et de son environnement culturel.

Le CIDOC CRM, une ontologie pour l'information relative au patrimoine culturel.

L'ontologie CIDOC CRM, qui émane du monde des musées¹⁸, pourrait couvrir les informations attachées à un document culturel, ou portées par celui-ci, en les « contextualisant » et en évitant les ambiguïtés précédemment présentées. Elle fournit en effet des définitions et un modèle pour décrire les notions et les relations permettant de documenter le patrimoine culturel.

La définition de cette norme ISO¹⁹ rappelle son usage pour l'interopérabilité des données, signale l'exigence de qualité de l'information traitée – notamment que le CRM est destiné « à supporter le niveau de détail et de précision exigés par des professionnels des musées et des chercheurs dans le domaine » – et précise le périmètre de son application : sont concernées les informations relatives au patrimoine culturel (*Cultural Heritage*), matériel et immatériel, à savoir « tout type de matériel rassemblé et exposé par des musées et des institutions apparentées », y compris « des collections, des sites et des monuments en rapport avec des domaines tels que l'histoire sociale, l'ethnographie, l'archéologie, les beaux-arts et les arts appliqués, l'histoire naturelle, l'histoire des sciences et de la technologie. » La définition souligne également la capacité du modèle à prendre en compte les « informations contextuelles (c'est-à-dire historiques, géographiques et théoriques qui donnent aux collections de musée leur signification culturelle et leur valeur). »

Ce dernier point et la large couverture du CRM permettent de l'envisager comme moyen de rassembler, de façon normalisée, des informations et des données sur des « objets » culturels aussi divers qu'une fouille archéologique, une pratique culturelle attachée à des objets ethnologiques (la coiffe de plume d'un amérindien et sa symbolique), un personnage historique et les événements qui lui sont liés (Napoléon...), des courants de pensée ou artistiques et leurs acteurs comme le groupe des artistes ayant participé en 1913 à l'*Armory Show* de New York et l'impact décisif de l'événement sur l'art américain du XX^e siècle, ou

¹⁷ Parmi les standards utilisés dans les domaines des GLAM (Gallery Library Archive Museum) : EAD (Encoded Archival Description) ; FRBR (Functional Requirements for Bibliographic Records) ; TEI (Text Encoding Initiative) et son émanation pour l'épigraphie, EpiDoc ; EDM (Europeana Data Model), pour Europeana (<http://pro.europeana.eu/edm-documentation>).

¹⁸ Le CRM (Conceptual Reference Model) a été élaboré par le Comité International pour la Documentation (CIDOC) du Conseil International des Musées (ICOM) : www.cidoc-crm.org. Norme ISO 21127:2006. L'ontologie continue d'évoluer : la norme et la version 5.0.4, de novembre 2011, du document de référence *Definition of the CIDOC Conceptual Reference Model*, peuvent être téléchargées en anglais sur le site du CIDOC : www.cidoc-crm.org/official_release_cidoc.html. La dernière mise à jour et sa nouvelle traduction française sont actuellement présentées à l'ISO. Pour une documentation abondante, cf. le site du CIDOC.

¹⁹ *Definition of the CIDOC Conceptual Reference Model*, v. 5.0.4, 2011, p.i-ii (www.cidoc-crm.org/docs/cidoc_crm_version_5.0.4.pdf). Cf. la définition résumée sur la page web de l'ISO 21127:2006.

même des phénomènes indicibles entre deux expressions artistiques comme l'influence du cabaret le *Bœuf sur le toit* sur l'introduction du jazz dans la musique française²⁰.

L'ontologie définit environ cent « Classes » (codées *E...*), qui correspondent à des notions, des concepts, et près de cent cinquante « Propriétés » (*P...*) qui sont des interactions entre les Classes (prédicats exprimés sous forme de verbe).

Par exemple, l'information « la statue est en porphyre » pourrait être reformulée :
 Classe **E24_Physical_ManMade_Thing** = *statue* | Propriété **P45_consists_of** |
 Classe **E57_Material** = *porphyre*

La granularité de l'information est assurée par une organisation hiérarchique des Classes, les définitions étant de plus en plus de précises au fil des « sous-classes » (fig. 2)²¹. La norme précise également les associations Classes/Propriétés (concepts/verbes) possibles, sachant qu'une sous-classe hérite des propriétés de ses classes parentes. Plusieurs de ces interactions et relations sont schématisées dans une série de diagrammes disponibles sur le site du CIDOC²².

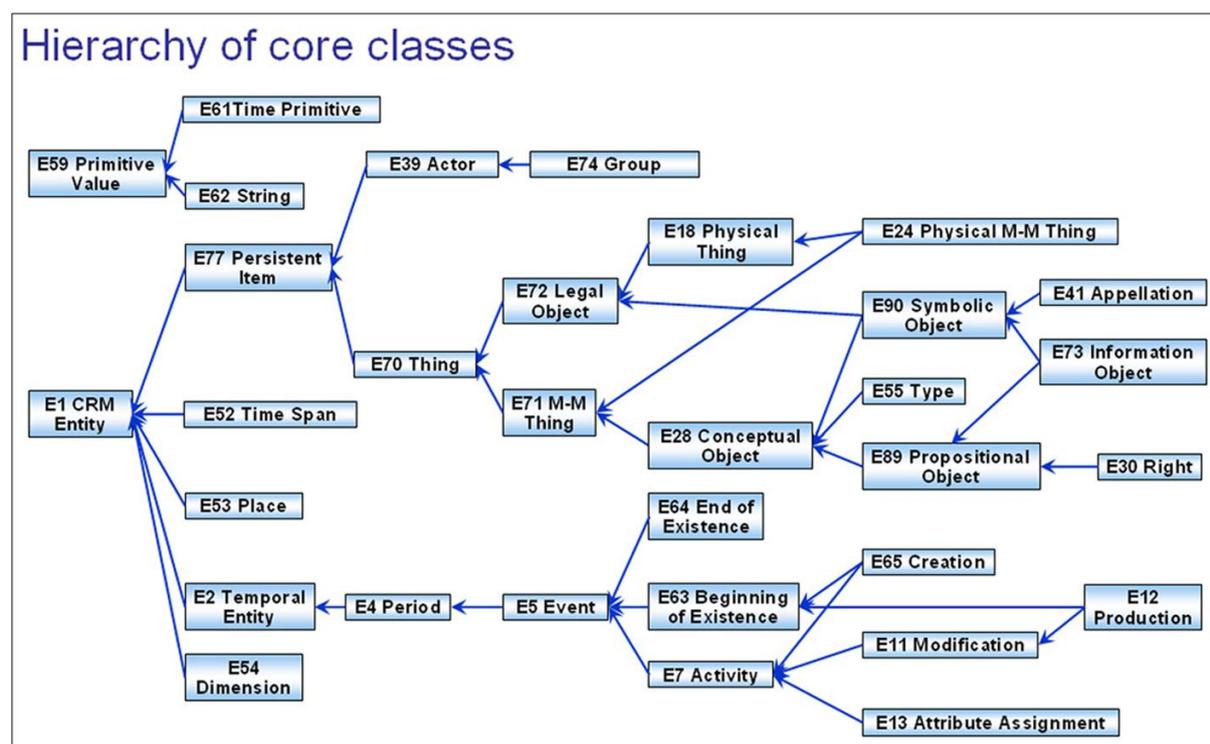


Fig. 2 : hiérarchie simplifiée des Classes du CIDOC CRM
 (http://www.cidoc-crm.org/cidoc_core_graphical_representation/hierarchy.html)

Le modèle s'articule autour de six axes forts²³ :

²⁰ Cf. les rencontres consacrées, à l'initiative d'A. THARAUD, à la Cité de la Musique (14-10-2012), à ce moment clé de la création musicale française des années 20 (http://citedelamusique.fr/pdf/note_programme/np_12269.pdf (Consultation : 13-10-2012)).

²¹ Schéma développé : www.cidoc-crm.org/cidoc_graphical_representation_v_5_1/class_hierarchy.html

²² www.cidoc-crm.org/comprehensive_intro.html (Consultation : 2012-10-25).

²³ TZOMPANAKI, K./ DOERR, M., *A New Framework for Querying Semantic Network*, 2011, partie « 3.1. Designing fundamental categories », p.6. www.ics.forth.gr/tech-reports/2011/2011.TR419_Querying_Semantic_Networks.pdf (Consultation : 2012-10-27).

- **Objet**, matériel ou immatériel (*Thing*)
- **Personne**, physique ou morale, individu ou groupe (*Actor*)
- **Événement** (*Event*)
- **Lieu**, notion géographique ou emplacement (*Place*)
- **Temps** (*Time*)
- **Concept**

À ces notions-clés s'ajoutent d'autres éléments permettant de traiter notamment les dimensions et l'état de l'objet, les droits, la documentation, la gestion de collections, etc. Le modèle est extensible et d'autres notions peuvent être librement ajoutés afin de couvrir des besoins particuliers, à condition de respecter les règles de l'ontologie, notamment l'héritage de la hiérarchie.

Dans la mesure où l'ontologie a acquis le statut de norme ISO, et grâce à l'ensemble de ses définitions, le document de référence devient un outil pour documenter les applications fondées sur son utilisation.

Les données sur la statue de Minerve/Rome : structurer l'information.

En appliquant les principes du CRM, on pourrait rendre compte du parcours et des états de la statue de Minerve/Rome, modéliser l'information en distinguant par exemple les matériaux d'origine et de rénovation, les diverses interprétations dues au changement du personnage représenté, ou encore les droits auxquels sont assujettis l'œuvre, ses reproductions et notices numériques (fig. 3). La sémantisation des données passe non seulement par le choix de la norme, de ses Classes et de ses Propriétés pour les métadonnées mais aussi par leur place dans le modèle. La proposition de la figure 3, exprimée dans une formulation simplifiée inspirée du XML²⁴, présente l'organisation de quelques unes de ces informations : les matériaux actuels sont signalés directement parmi les données sur la statue tandis que les matériaux d'origine sont « encapsulés » dans l'ensemble réservé à la fabrication de l'œuvre antique. Les droits de l'objet et ceux de la reproduction photographique sont eux aussi insérés dans leurs parties respectives et on aurait pu faire de même avec leurs dates.

Les données sur la statue en tant qu'objet conservé au Louvre pourraient être directement traitées avec le schéma LIDO, modèle fondé sur le CIDOC CRM et réalisé pour le traitement de l'information relative aux objets et collections de musées²⁵, mais l'information produite par le monde de la recherche diffère de la documentation muséal et LIDO pourrait n'en couvrir qu'une partie. Préserver la spécificité des approches scientifiques est important et le schéma-pivot CRM-IDA, en cours d'élaboration pour l'interopérabilité de

²⁴ Le principe de l'encapsulation des balises/données est restitué mais, pour simplifier, seul le cœur de la balise ouvrante est donné.

²⁵ LIDO (Lightweight Information Describing Objects) v1.0, novembre 2010. Pour la documentation, cf. les pages de l'ICOM sur LIDO : <http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/what-is-lido/> (Consultation : 2012-10-26). LIDO tient compte des schémas *SPECTRUM* (www.collectiontrust.org.uk/schema), *CDWA Lite* (Categories for the Description of Works of Art) [http://getty.art.museum/research/publications/electronic_publications/cdwa/cdwalite.html] et *museumdat* (www.museumdat.org/index.php). Pour des exemples d'utilisation de LIDO : <http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/lido-overview/examples/> . Pour les mappings d'un même objet avec LIDO et avec EDM : *Europeana Data Model. Mapping Guidelines v1.0.1* (24-02-2012) Annexe A, p.32-37 : <http://pro.europeana.eu/documents/900548/ea68f42d-32f6-4900-91e9-ef18006d652e> (Consultation : 2012-10-26).

corpus numériques d'objets produits par des équipes de recherche, pourra associer LIDO et les éléments du CRM²⁶.

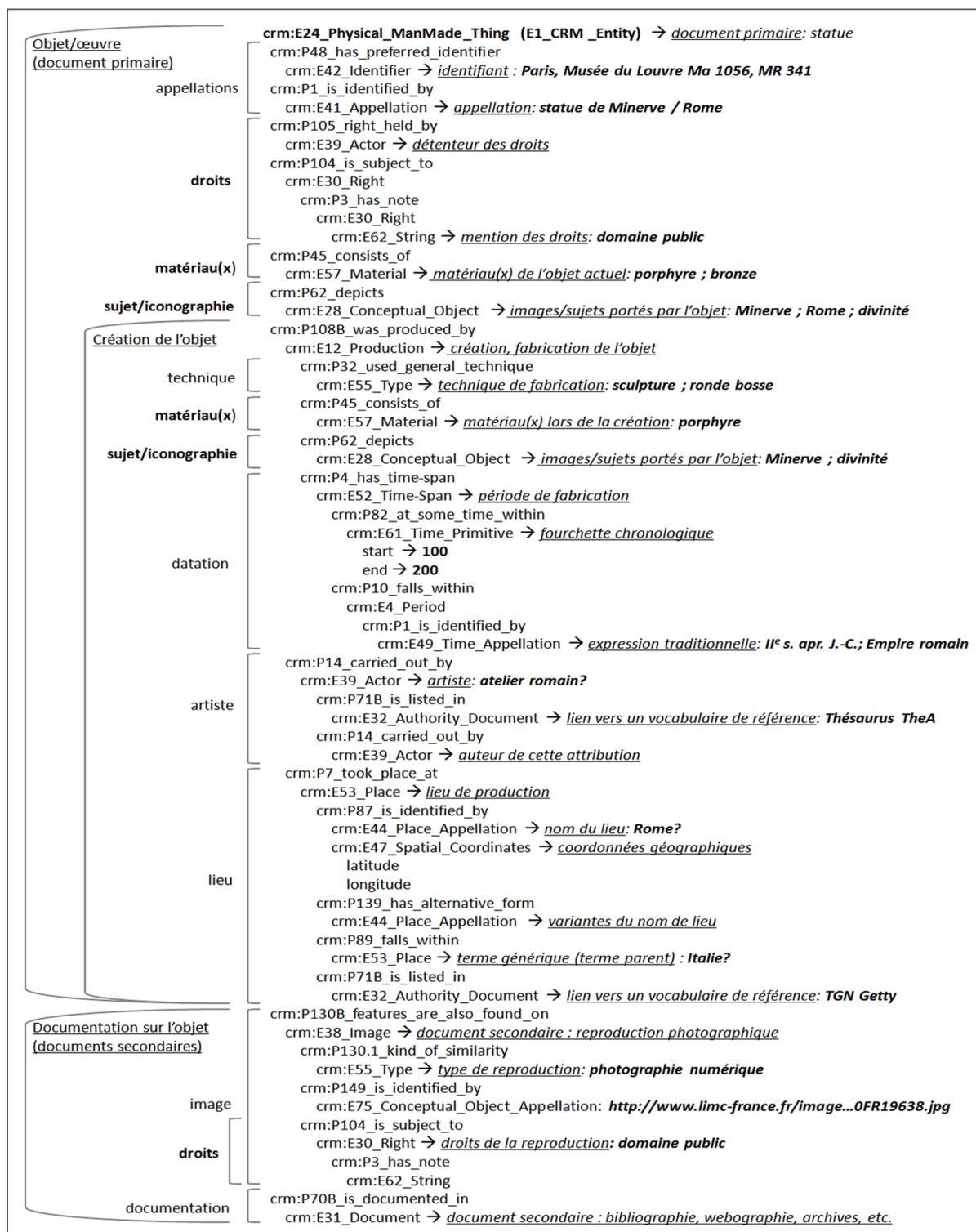


Fig. 3 : extrait d'une proposition de schéma simplifié CIDOC CRM/XML de la statue de Minerve (Louvre MR 341, Ma 1056)

²⁶ IDA (Interopérabilité des Documents Anciens), soutenu par le TGE Adonis. Cf. SZABADOS, A.-V., « Le site Web LIMC-FRANCE : iconographie de la mythologie et corpus d'œuvres antiques », § « L'optimisation de la diffusion et l'insertion dans le Web de données », *Mouseion* (UniLaSalle), à paraître.

Traiter les informations complexes d'un objet de fouilles

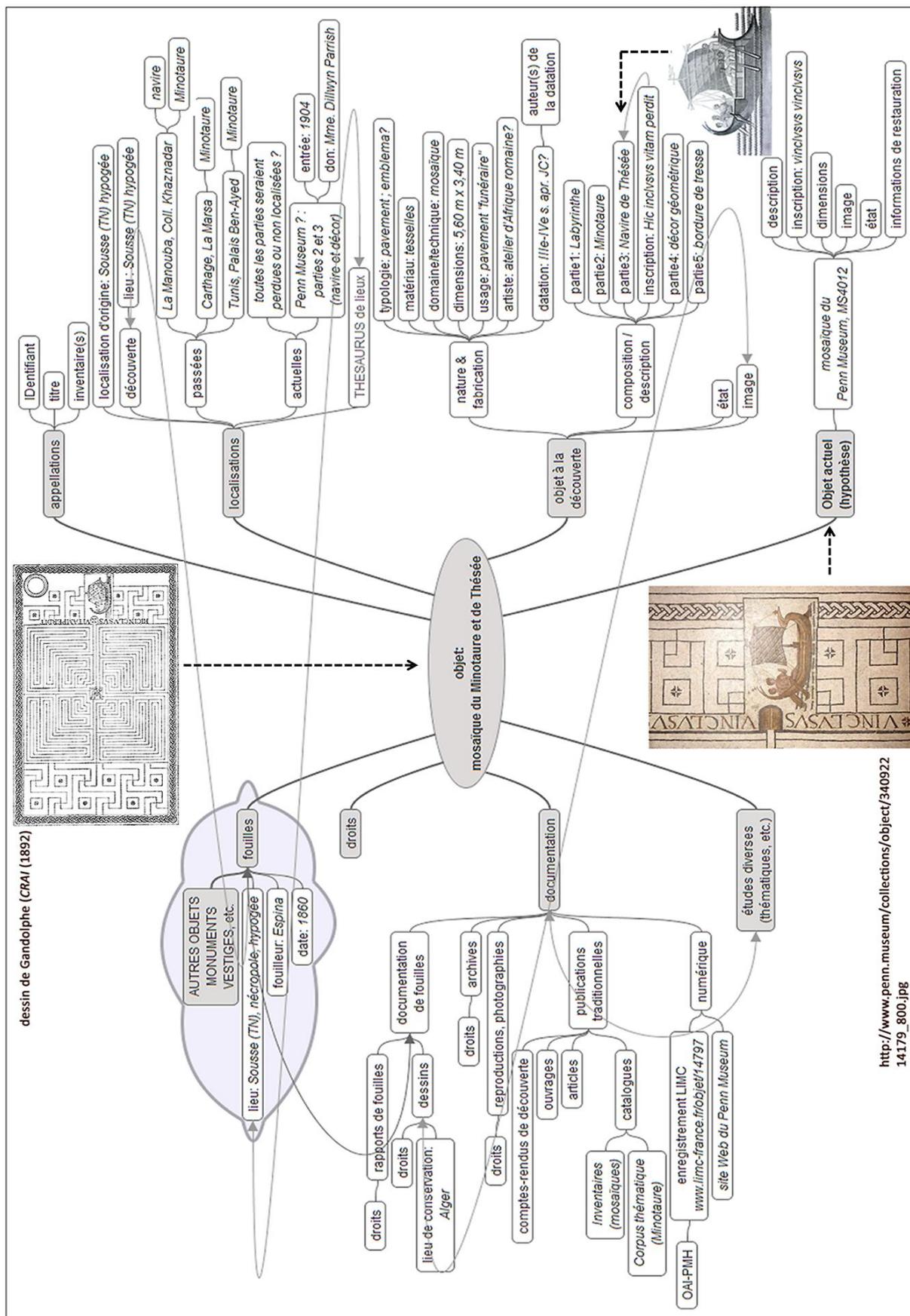


Fig. 4 : exemple de modélisation de l'information d'une mosaïque découverte à Soussse

L'étude d'une mosaïque romaine au destin tourmenté réunit des connaissances variées et lacunaires, et fait appel à de multiples documents épars (fig. 4) : découvert en 1860 par M. Espina lors de la fouille d'une nécropole de Sousse, l'antique Hadrumetum, le pavement fut démantelé et ses scènes figurées transférées dans le palais de M. Khaznadar, alors Grand Vizir de la Régence de Tunis, puis, après la chute de ce dernier, en partie déplacées dans d'autres palais. Bien que considéré comme actuellement perdu – comme le reste du pavement – le panneau du navire de Thésée est rapproché hypothétiquement d'un document presque similaire conservé au Penn Museum de Philadelphie²⁷ grâce aux documents et dessins réalisés lors de la découverte, aux comptes rendus et témoignages de l'époque, ainsi qu'aux recherches plus récentes comme celles effectuées par W. A. Daszewski pour établir un corpus des représentations musivales antiques de la légende de Thésée et du Minotaure²⁸.

Les Classes et les Propriétés du CIDOC CRM permettent de traiter l'une des modélisations possibles de ces connaissances et de leurs relations²⁹ (fig.4), qui restitue notamment le parcours géographique et historique de la mosaïque³⁰, son état au fil du temps, son identification, ou encore ses liens avec la fouille de la nécropole et les autres sépultures découvertes, avec les acteurs des événements, avec sa documentation (étude de l'hypogée, dessins, publications, photographies, etc.). L'information géographique, particulièrement importante dans ce cas, peut être détaillée (fig. 5-6), tant au niveau du lieu lui-même, *Sousse* (ses appellations : *E48_Place Name*, *E44_Place Appellation* ; ses coordonnées : *E47_Spatial Coordinates* ; etc.) que des différentes localisations de l'objet (*P53_has former or current location/ P55_has current location/ P54_has current permanent location | E53_Place*).

```
<crm:E53_Place>
  <crm:P87_is_identified_by>
    <crm:E48_Place_Name>Sousse</crm:E48_Place_Name>
    <crm:E47_Spatial_Coordinates>
      <latitude>35.8333</latitude>
      <longitude>10.6333</longitude>
    </crm:E47_Spatial_Coordinates>
  </crm:P87_is_identified_by>
  <crm:P139_has_alternative_form>
    <crm:E44_Place_Appellation>Hadrumetum</crm:E44_Place_Appellation>
    <crm:E44_Place_Appellation>Hadrumète</crm:E44_Place_Appellation>
  </crm:P139_has_alternative_form>
  <crm:P89_falls_within>
    <crm:E53_Place>Tunisie</crm:E53_Place>
  </crm:P89_falls_within>
  <crm:P71B_is_listed_in>
    <crm:E32_Authority_Document>TGN
    <crm:E32_Authority_Document>Getty</crm:E32_Authority_Document>
  </crm:P71B_is_listed_in>
</crm:E53_Place>
```

Fig. 5 : extrait de schéma CIDOC CRM/XML consacré à la ville de Sousse

²⁷ Inv. MS4012 : LUCE, S. B., « Five Roman Mosaics », *The Museum journal. University of Pennsylvania* 7, 1916, p.18-26 ; SZABADOS, A.-V., LIMCicon ID 14822, in *LIMC-France*. www.limc-france.fr/objet/14822 (Consultation : 22-10-2012).

²⁸ FOUCHER, L. *Inventaire des mosaïques. Sousse*, 1960, p.76 n°57.167 ; DASZEWSKI, W. A., *Nea Paphos II. La mosaïque de Thésée*, 1977, p.125-126 n°54 ; SZABADOS, A.-V., LIMCicon ID 14797, in *LIMC-France*. www.limc-france.fr/objet/14797 (Consultation : 22-10-2012).

²⁹ Pour des associations objet/ personne/ temps/ événement/ lieu..., cf. *Object Association Information* : http://www.cidoc-crm.org/cidoc_graphical_representation_v_5_1/object_association.html

³⁰ Diagramme espace/temps : www.cidoc-crm.org/cidoc_graphical_representation_v_5_1/spatial_temporal.html

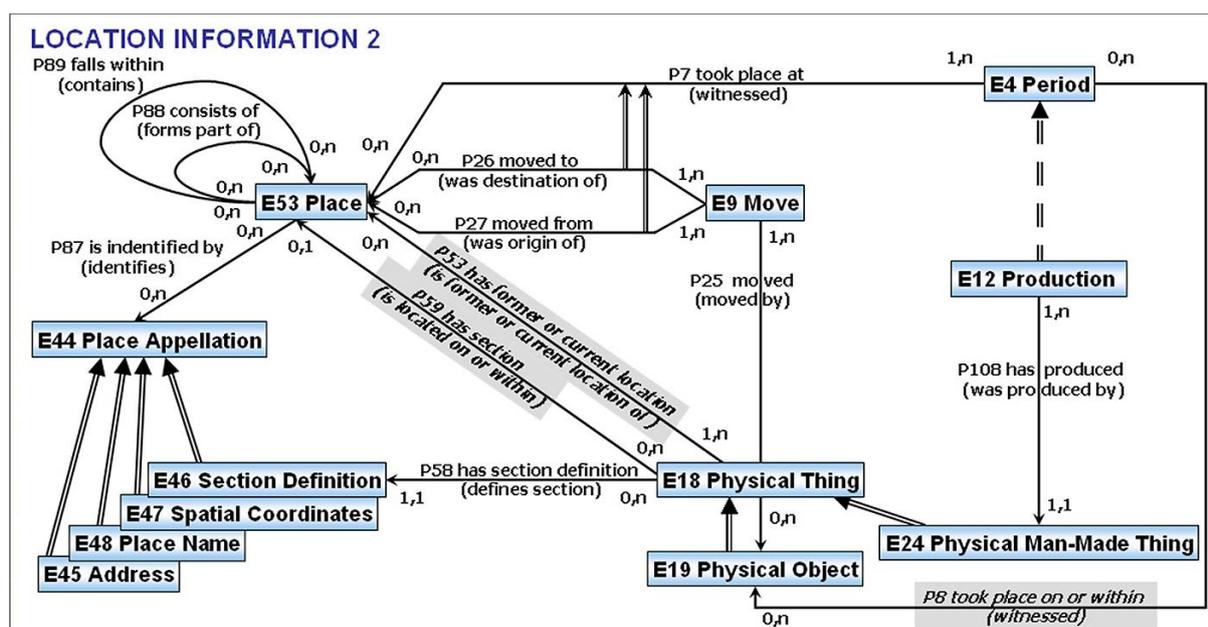


Fig. 6 : traitement CIDOC CRM des lieux et localisations
http://www.cidoc-crm.org/cidoc_graphical_representation_v_5_1/location.html

Par sa capacité à réunir des informations hétérogènes et à conserver leurs liens avec leur histoire, leur parcours ou les personnes impliquées, cela en particulier grâce à la préservation du contexte événementiel, le CRM pourrait couvrir une grande partie des données produites à l'occasion d'une fouille archéologique. Le modèle anglais CRM-EH, élaboré dans le cadre du projet STAR (Semantic Technologies for Archaeology Resources), pour un système d'information destiné à rendre plusieurs ressources archéologiques interopérables, est une extension de l'ontologie³¹.

Objets, modèles, répliques, typologies

Le CRM rassemble ces informations complexes non seulement en donnant les moyens de reformuler les connaissances mais aussi en facilitant les points de jonction, comme dans le cas d'une base de données relationnelle conçue pour un programme pluridisciplinaire et archéologique sur l'histoire des mines³². Les rapprochements avec d'autres standards métiers sont possibles, ainsi qu'avec des documents de référence. Par exemple, le lieu de découverte de la mosaïque, *Sousse*, peut être directement associé à un vocabulaire spécialisé, un thésaurus de référence comme le *TGNGetty*³³ grâce à *P71_lists (is listed in) | E32_Authority Document* (fig. 5, texte souligné)³⁴. L'ontologie donne les moyens de créer des typologies, des nomenclatures, des taxonomies dans des domaines relevant du concept, de l'iconique, du

³¹ BINDING, C./ MAY, K./ TUDHOPE, D., « Semantic Interoperability in Archaeological Datasets: Data Mapping and Extraction via the CIDOC CRM », *Proceedings of the 12th European Conference on Digital Libraries (ECDL)*, 2008, p.280-290. Diagramme de la modélisation : www.cidoc-crm.org/docs/AppendixA_Diagram_V9.pdf. Projet STAR : <http://hypermedia.research.glam.ac.uk/kos/star/>.

³² HIEBEL, G./ HANKE, K./ HAYEK, I., « A relational database structure and user interface for the CIDOC CRM with GIS integration », Nuremberg, 20-22/12/2010. www.cidoc-crm.org/docs/Hiebel_crm_sig_2010.ppt

³³ <http://www.getty.edu/research/tools/vocabularies/tgn/>

³⁴ Un thésaurus peut aussi être inséré grâce à la Classe *E55_Type* qui « type » des données ou des ensembles de données.

sonore, par exemple des modèles de référence visuels (type iconographique), des objets-types, des séquences sonores, etc.³⁵

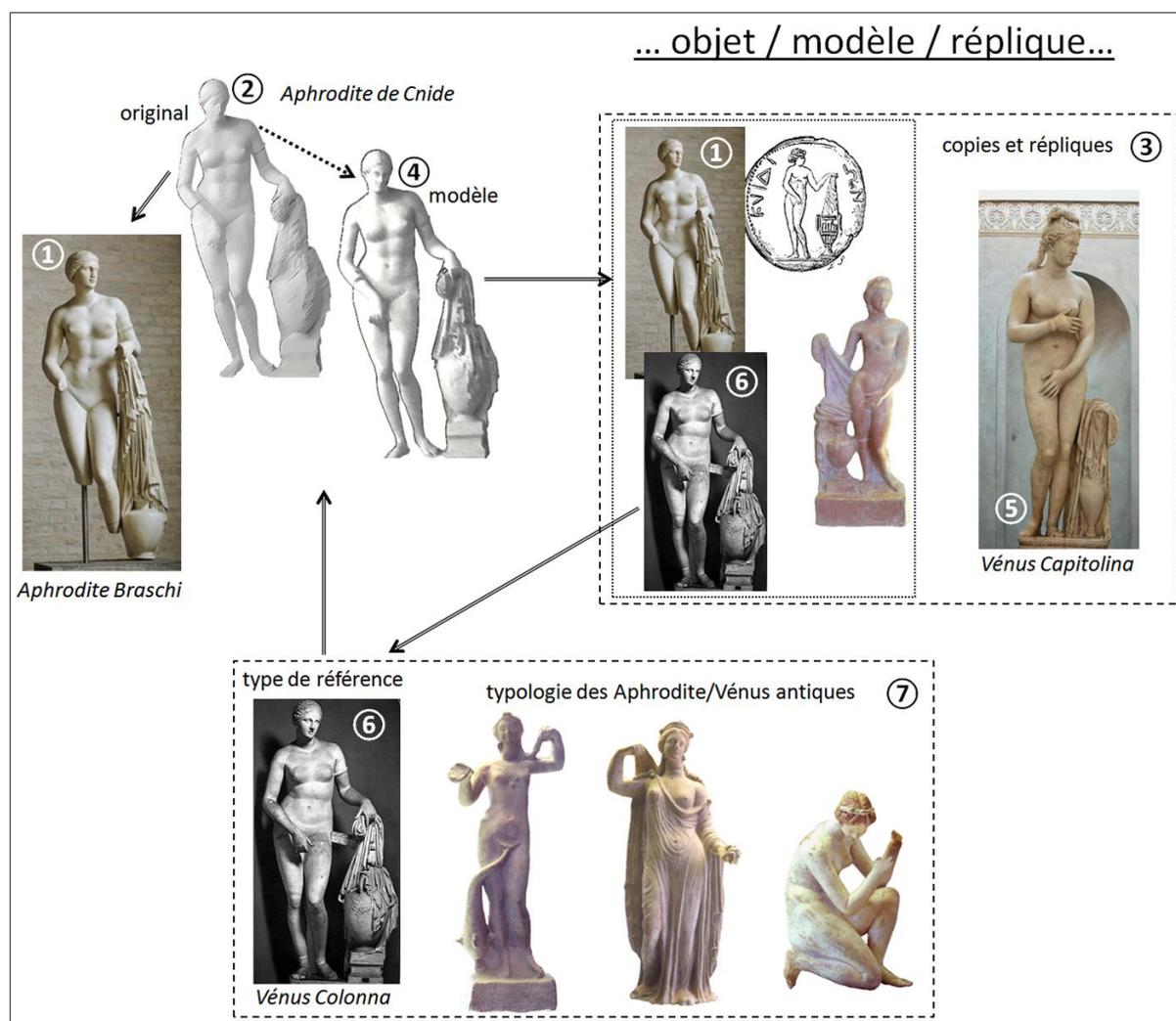


Fig.7 : L'Aphrodite de Cnide : original, modèle, copies et types de référence (Cnidienne et clichés des terres cuites du Louvre : A.-V. Szabados. 1, 3 et 6 : Wikimedia Commons)

Sur la figure 7, l'Aphrodite Braschi, une statue d'époque hellénistique aujourd'hui conservée à Munich³⁶, est placée au centre d'un réseau complexe d'informations :

L'Aphrodite Braschi ① serait une copie de l'Aphrodite de Cnide de Praxitèle³⁷, réalisée au IV^e siècle av. J.-C et détruite depuis des siècles. Lors de sa fabrication, la statue de Praxitèle était un objet matériel ② (*E12_Production* → *E18_Physical Thing*) mais sa renommée, transmise dès l'Antiquité par ses évocations dans des sources écrites (*E31_Document*) et ses copies ③ (*E11_Modification/ E12_Production* → *E18_Physical Thing*) en fit un type iconographique, un modèle conceptuel ④ (*E65_Creation* → *E28_Conceptual Object*). L'Aphrodite Braschi n'en est pas la seule copie et d'autres, comme

³⁵ Cf. diagramme *Taxonomic Discourse* : www.cidoc-crm.org/cidoc_graphical_representation_v_5_1/taxonomic_discourse.html

³⁶ Munich, Glyptothek GI 258. VIERNEISEL-SCHLÖRB, B., *Glyptothek München, Katalog der Skulpturen II*, 1979, p.323-336 n°31, fig. 158-164 ; SZABADOS, A.-V., LIMCicon ID 14767, in *LIMC-France*. www.limc-france.fr/objet/14767 (Consultation : 28-09-2012).

³⁷ Pour l'Aphrodite de Cnide et ses copies : PASQUIER, A., in PASQUIER, A./ MARTINEZ, J.-L. (dir.), *Praxitèle. Paris, Louvre 2007*, 2007, p.139- 151.

la *Vénus Capitolina*³⁸ ⑤, présentent des différences dans la pose du personnage (position des mains). L'*Aphrodite Braschi* est considérée comme l'un des modèles de référence de sa série (*P137_exemplified*) ce qui en fait aussi un type iconographique – un modèle conceptuel – dont l'exemple-type complet pourrait être la *Vénus Colonna*³⁹ ⑥ (+ *P136.1* ou *P137.1_in the taxonomic role*).

Avec le CRM, ces informations sont organisées en distinguant clairement les objets matériels (*E18_Physical Thing*) des objets conceptuels (*E28_Conceptual Object*) et en conservant les relations entre ces œuvres à la fois semblables (« *statue antique d'Aphrodite nue, avec une draperie et un vase* ») et bien différentes. L'*Aphrodite de Cnide* peut être présentée comme le modèle de référence de l'*Aphrodite Braschi*, cette dernière comme un autre modèle de référence, la *Vénus Colonna* comme un taxon, au sein d'une typologie des schémas iconographiques antiques d'Aphrodite/Vénus ⑦ (-*anadyomène*, -*genitrix*, -*accroupie*, etc.). L'activité savante consistant à établir un type à partir de la comparaison de ces sources matérielles et écrites peut quant à elle être exprimée grâce à l'événement *E83_Type Creation* et ses articulations (*P135_created type*, *P136_was based on*, etc.).

S'appropriier l'ontologie.

La documentation produite autour de l'ontologie est de plus en plus abondante sur le site du CIDOC mais, malgré les tutoriels, schémas explicatifs et partages d'expérience, la large couverture de la norme et les nuances subtiles de ses définitions ne facilitent pas l'appropriation du CRM. L'ontologie est relativement facile à comprendre et à manipuler – en particulier dans la phase de modélisation de l'information durant laquelle elle fournit, pour la réflexion, un éventail de concepts liés au domaine culturel – mais on constate qu'elle permet de modéliser une même information de plusieurs façons différentes. Dans ce cas, même si les données sont bien enrichies sémantiquement, qu'en est-il effectivement de l'interopérabilité de ces données ?⁴⁰

La création et l'adoption par les utilisateurs de modèles et de schémas communs, comme LIDO, pourraient apporter des solutions et il a semblé nécessaire, lors des journées *THATCamp Paris 2012*, de réunir en France une communauté d'intérêt et de pratique (*ontologie-patrimoine*), largement ouverte et trans-disciplinaire, afin de faciliter les échanges et les partages d'expérience, ainsi que les retours vers les concepteurs de l'ontologie⁴¹.

La sémantisation des ressources numériques, destinée à préserver les diverses facettes du savoir et à favoriser l'insertion de la production scientifique dans le

³⁸ Rome, Musées Capitolins, Palazzo dei Conservatori, inv. 409. DELIVORRIAS, A. / BERGER-DOER, G. / KOSSATZ-DEISSMANN, A., « Aphrodite » p.52 n°409, in *LIMC* vol.2 (1984).

³⁹ Vatican, inv. 812. DELIVORRIAS/ BERGER-DOER/ KOSSATZ-DEISSMANN, *supra* n.38, p.50 n°391.

⁴⁰ Publication de l'atelier proposé par A.-V. SZABADOS, K. BRIATTE et R. LETRICOT : Collectif, « Utiliser l'ontologie CIDOC CRM pour l'information relative au patrimoine culturel », in *THATCamp Paris 2012*, 2012, § 34-35, 41 sq. <http://editionsmslh.revues.org/319> (Mise en ligne 28-09-2012. Consultation 04-10-2012). Cf. les retours d'expérience : HASLHOFER, B./ NUSSBAUMER, Ph., « CIDOC CRM in Practice », novembre 2009, slides 10 sq. (<http://fr.slideshare.net/bhaslhofer/cidoc-crm-in-practice>) ; OLDMAN, D., « The British Museum, CIDOC CRM and the Shaping of Knowledge », 04-09-2012 (www.oldman.me.uk/blog/the-british-museum-cidoc-crm-and-the-shaping-of-knowledge/) ; BINDING/ MAY/ TUDHOPE, 2008, *supra* n.31 : « The abstractness of the CRM and the lengthy relationship chains arising from the event-based model also raised issues for designing appropriate user interfaces » (p. 283), « As with BRICKS, it proved necessary to create technical extensions to the CIDOC CRM to deal with attributes required for practical implementation concerns. » (p.289).

⁴¹ *THATCamp Paris 2012*, *supra* n.40, § 17 sq. Liste de diffusion : *liste.ontologie-patrimoine*

Web sémantique, voire le Web de données, est optimisée par l'association de standards techniques et documentaires. Parmi ceux-ci, l'ontologie de domaine CIDOC CRM se démarque par sa capacité à rassembler et traiter des connaissances et des données hétérogènes, mais associées, en conservant leurs liens et leurs contextualisations. L'un des constats formulés dans les retours d'expérience sur l'application pratique de la norme porte sur l'importance de l'implication des spécialistes du domaine scientifique traité lors de la modélisation de leur discours. Les exemples présentés s'inscrivent dans cette démarche et montrent que le CRM permet d'exprimer ces connaissances culturelles avec une granularité fine.