

Le modèle CRM pour la documentation muséographique :

s'attacher au sens pour n'être pas piégé par la forme

Journée d'étude organisée par l'ADBS : « La modélisation : pourquoi l'intégrer dans les systèmes d'information documentaire ? », Paris-La Défense, 20 mai 2003

Patrick Le Bœuf,
Bibliothèque nationale de France,
membre du groupe ISO/TC46/SC4/WG9
et du CRM Special Interest Group (CRM-SIG)

Le CRM : qui est-il ? d'où vient-il ? où va-t-il ?

Qui est-il ?

Le CRM (*Conceptual Reference Model*) est un modèle qui vise à expliciter sans ambiguïté le sens des informations relatives à un « objet de musée » dans l'acception la plus large du terme (œuvre d'art, vestige archéologique tel qu'un clou ou un four à pain, corps momifié, fleur séchée ou animal empaillé, mais aussi quelque chose qu'il est matériellement impossible de conserver à l'intérieur d'un musée mais qu'une institution du patrimoine culturel peut gérer, comme un monument ou un site tel qu'une grotte préhistorique), et ce *quelle que soit la structure des données* qui véhiculent ces informations.

Le CRM peut en outre se définir comme :

- un modèle orienté objet,
- un modèle sémantique,
- une ontologie de domaine.

– Un modèle orienté objet :

Cette première caractérisation ne fait référence qu'au formalisme du modèle. Le formalisme orienté objet a été préféré au formalisme entité-relation parce que, entre autres raisons :

- il permet de donner une vue d'ensemble sans s'encombrer de détails superflus, par le biais du mécanisme des « héritages » (c'est-à-dire : puisque tout ce qui peut être dit du cas général peut aussi être dit du cas particulier, il peut suffire, dans un contexte pédagogique, de ne parler que du cas général) ;
- il est plus aisément gérable et extensible que le formalisme entité-relation ;
- il ne préjuge d'aucun mode d'implémentation concrète : il est tout à fait possible de construire une base de données relationnelle à partir d'un modèle orienté objet.

– Un modèle sémantique :

Cette deuxième caractérisation insiste sur le fait que le CRM s'attache à faire émerger la signification réelle de tout ce qui est considéré comme « implicite » et « évident » dans une structuration d'informations. Si un individu A ne prend pas la peine d'explicitier le sens d'une information parce que cela lui semble évident (par exemple : « date = 1830 », en tenant pour acquis qu'il s'agit de la date à laquelle un tableau a été peint), il risque d'arriver que cette « évidence » ne saute pas aux yeux d'un individu B, qui va interpréter cette même information de façon erronée (par exemple : « ce tableau représente un événement qui s'est produit en 1830 »). En forçant un peu le trait jusqu'à la caricature, on pourrait dire que le but primordial du CRM est d'enfoncer les portes ouvertes...

– Une ontologie de domaine :

Le terme d'ontologie a été emprunté par les informaticiens à la philosophie, où il s'agit de déterminer quels sont les éléments qui sont réputés « être » et quelles relations ils entretiennent entre eux (c'est exactement ce à quoi correspond l'expression « entité relation »). Le CRM est une ontologie en ce sens qu'il vise à définir formellement et à expliciter un ensemble de concepts sous-jacents, non nécessairement exprimés.

La mention « de domaine » indique une restriction, une limitation à cette entreprise ontologique : il s'agit en fait tout simplement d'un effort visant à définir, non pas *tout* ce qui existe au monde (tâche virtuellement impossible et infinie), mais « seulement » tout ce qui présente un caractère de pertinence par rapport à la documentation muséographique.

D'où vient-il ?

Le CRM a été élaboré par le *Documentation Standards Working Group* du CIDOC (Comité international pour la documentation), lequel est lui-même une émanation du Conseil international des Musées (ICOM).

Il succède en fait à un modèle de données de type entité relation qui, au fil du temps, était devenu ingérable : la dernière version de ce modèle comportait quelque 430 entités et ses concepteurs eux-mêmes finissaient par ne plus s'y retrouver ! Ce modèle reposait en outre sur un présupposé, courant dans les années 80 mais qui commençait à être battu en brèche au milieu des années 90, à savoir que, pour pouvoir *échanger* de l'information, il fallait absolument que tout le monde fasse exactement la même chose de la même façon, selon les mêmes schémas de données. L'idée commençait à se faire jour que des systèmes de médiation pourraient être en mesure de gérer des données provenant de sources hétérogènes ; partant, l'objectif n'était plus de concevoir un modèle de données unique et énorme, destiné à répondre *a priori* à tous les types de besoins et à tous les cas de figure, mais de concevoir un modèle de *concepts* pouvant servir de point de *référence* (d'où le nom même du CRM, *Conceptual Reference Model*).

Ce nouveau modèle ne partait pas de zéro mais s'appuyait sur un ensemble d'éléments d'information prédéfinis dans un document normatif, les *International Guidelines for Museum Object Information : the CIDOC Information Categories*, publiés en 1995. Cet ensemble comporte 74 éléments informatifs, répartis en 22 grands groupes tels que « Acquisition », « Localisation », « Marques et inscriptions », « Matériau et technique », « Sujet représenté », etc.

Le travail d'élaboration du CRM a commencé en 1996. Une première version en a été rendue publique en 1998. On en est aujourd'hui à la version 3.4, dont il existe une traduction en français¹. Un groupe de travail spécifique, organisé par le CIDOC mais ouvert également aux personnes extérieures au CIDOC, s'est mis en place pour gérer les mises à jour du modèle : le *CRM Special Interest Group* (CRM-SIG). Depuis 2002, les grandes lignes du modèle sont réputées définitivement stabilisées, et il ne s'agit plus désormais que d'assurer des modifications de détail, notamment en ce qui concerne les définitions et les exemples.

Où va-t-il ?

Le CRM a été soumis à l'ISO en 2000 pour devenir une norme internationale. Les procédures de l'ISO sont toujours longues et compliquées : deux votes ont déjà eu lieu en 2000 et 2002, un troisième vote aura lieu dans le courant de l'année 2003. Un groupe s'est mis en place au sein de l'ISO pour travailler à ce projet de norme : le groupe ISO TC46/SC4/WG9. Le CRM est donc à l'étude parallèlement auprès du CRM-SIG du CIDOC et auprès du groupe TC46/SC4/WG9 de l'ISO. Cette situation qui pourrait paraître potentiellement conflictuelle ne l'est en fait pas du tout... car tous les membres du groupe ISO sont *de facto* également membres du CRM-SIG. Aux yeux de l'ISO, le CRM est actuellement connu comme projet de norme CD 21127.

Le fait que le CRM devienne une norme ISO ne signifie nullement qu'il deviendra « obligatoire ». Cette validation contribuera simplement à asseoir la position du CRM, à lui donner une plus grande « légitimité » et à le rendre disponible bien au delà du seul cercle de la documentation muséographique.

Le CRM : à quoi ressemble-t-il ?

Au cœur même du CRM, on trouve la notion d'événement, de phénomène temporel. En effet, si intuitivement on pourrait penser que ce qui est primordial c'est la description d'un objet matériel tel qu'il se présente aujourd'hui, il ne faut pas oublier qu'il s'agit avant tout de resituer cet objet dans un contexte historique. Ce qui importe, c'est à la fois l'environnement de cet objet au cours du temps (au moment de sa création, et lors de chacune de ses éventuelles modifications) et l'énoncé de tout ce qui a pu lui arriver. Du reste, dans la problématique d'un *échange* de données, ce n'est pas tant la description d'un objet *a priori* réputé unique et conservé à un seul

¹ http://cidoc.ics.forth.gr/docs/crm_french_version_3.4.zip

endroit qui importe², mais la description de ce contexte historique, des éléments qui peuvent être communs avec d'autres objets, ce qui à la fois rend cet objet « moins unique » et lui donne un sens culturel — même dans le domaine des sciences naturelles, où ce que l'objet documente est l'histoire de la discipline elle-même et de la construction d'une vision du monde qui nous entoure. Une grenouille dans un muséum n'est pas simplement une grenouille : une grenouille dans un muséum est un témoignage de la construction du concept de grenouille.

Nous nous intéressons donc avant tout à l'*historique* de tout ce qui est arrivé aux objets que nous décrivons, c'est pourquoi le cœur même du modèle est constitué des *Entités Temporelles*, qui expriment soit le fait qu'un objet est « dans tel état » (classe *État de Conservation*), soit le fait qu'il « s'est produit quelque chose » (un *Événement*, une *Activité*, un *Déplacement*, une *Production*, un *Début d'Existence* ou une *Fin d'Existence*, etc.) ; à partir de cette notion centrale, nous nous posons un certain nombre de questions fondamentales :

Quand est-ce arrivé ? nous répondons à cette question au moyen de la classe *Tranche Chronologique* ;

Où est-ce arrivé ? cette question implique l'existence de la classe *Lieu* (la notion de lieu étant comprise comme pouvant se définir par rapport à une constellation déterminée de matière, indépendamment de l'emplacement de cette constellation : comme un poinçon sur une bague, qui reste toujours au même endroit par rapport à la bague, alors que la bague elle-même se déplace à travers le monde au gré des déplacements de la personne qui la porte) ;

Qui a fait ça ? question qui met en jeu la notion d'*Agent* (laquelle implique une notion d'intentionnalité : un singe n'est pas considéré comme une instance d'*Agent*, même s'il détruit un tableau) ;

À quoi est-ce arrivé ? (donc, quel est l'objet décrit ?), question qui débouche à la fois sur la notion de contenant et sur celle de contenu immatériel, donc sur les notions de *Quelque Chose de Matériel* (qui peut être naturel ou avoir été fabriqué de main d'homme) et d'*Objet Conceptuel* (qui nécessairement ne peut qu'avoir été « fabriqué », c'est-à-dire être un produit de l'activité humaine) ;

À ce stade, on éprouve le besoin de *classifier* toutes ces informations : l'activité de classification met en jeu la notion de *Type* (les types peuvent s'appliquer absolument à tout élément du système)

Enfin, pour pouvoir parler aisément de toutes ces choses, il faut bien évidemment leur donner un nom, ce qui met en jeu la notion d'*Appellation*.

Le fait de placer systématiquement une *Entité Temporelle* (c'est-à-dire, en fait, le plus souvent, un *Événement*) au milieu d'une nébuleuse constituée d'un objet matériel ou conceptuel, d'un agent, d'un lieu et d'une tranche chronologique peut sembler très contraignant. En fait, le modèle CRM prévoit dans beaucoup de cas (mais pas tous) la possibilité de « court-circuiter » cet événement en reliant directement un objet au résultat de l'événement contourné. Ainsi, par exemple, l'événement constituant à mesurer un objet met en relation cet objet, l'agent qui mesure l'objet, l'endroit où a lieu cette activité, et la tranche chronologique au cours de laquelle elle a lieu ; il y a des cas très spécifiques où l'on peut effectivement avoir

² En cela, la notion d'*échange de données* recouvre quelque chose de sensiblement différent dans les bibliothèques et dans les musées.

besoin de tous ces éléments d'information et les avoir consignés dans une base de données, mais le plus souvent on a simplement besoin (ou on ne dispose que) du résultat de cette activité, à savoir : l'indication des dimensions de l'objet. Le modèle CRM prévoit donc à la fois un « chemin complet » de l'objet vers l'événement de mesurage et de l'événement de mesurage vers les dimensions de l'objet, et un « raccourci » qui dit simplement que l'objet a telles dimensions.

La notion d'événement est en outre précieuse en ce qu'elle permet de consigner avec précision des informations éventuellement contradictoires : par exemple, deux critiques d'art peuvent avoir attribué la même œuvre à deux artistes différents ; il importe alors d'indiquer qui a attribué cette œuvre à qui, et dans quelles circonstances.

Si le modèle CRM a essentiellement été conçu pour rendre compte de la sémantique des informations muséographiques, il peut relativement aisément servir également à expliciter la sémantique des informations bibliographiques contenues dans une notice en format MARC. Tout au plus faudrait-il peut-être simplement prévoir quelques extensions au modèle, entreprise facilitée par le choix du formalisme orienté objet. Voici par exemple une notice en format UNIMARC (ou plutôt un extrait de notice, car j'en omets volontairement un certain nombre de zones) :

```
010 .. $a 2-87811-089-7 $b br. $d 99 F
101 1. $a fre $c eng
200 1. $a Jérôme Bosch $b [Texte imprimé] $f Walter S. Gibson $g trad. de
l'anglais par Sophie Léchauguette
210 .. $a [Londres] $a [Paris] $c Thames & Hudson $d 1995
215 .. $a 177 p. $c ill. en noir et coul., couv. ill. en coul. $d 21 cm
300 .. Bibliogr. p. 174-175. Index
454 .1 $t Hieronymus Bosch
600 .1 $a Bosch $b Jérôme $f 1450?-1516
700 .1 $a Gibson $b Walter S.
702 .1 $a Léchauguette $b Sophie $4 730
960 .1 $a Hiëronimus Van Aeken
```

La signification précise de ces mystérieux assemblages de chiffres, et de lettres combinées avec le signe « dollar », n'est guère compréhensible que des catalogueurs eux-mêmes (et du système sur lequel ils travaillent). N'importe quel catalogueur habitué à travailler avec le format UNIMARC peut « traduire » le code « 210 \$d 1995 » comme signifiant : « date d'édition = 1995 ». Mais ce signifié peut lui-même être ambigu pour quelqu'un qui n'est pas habitué à manipuler de l'information bibliographique — ou pour une machine à laquelle on n'a pas préalablement « expliqué » ce que c'est qu'une date d'édition. Le modèle CRM va donc expliciter tout le cheminement suivant — évident aux yeux des documentalistes et bibliothécaires, mais pas forcément pour tout humain ni pour toute machine :

— « Jérôme_Bosch_par_W._S._Gibson_traduit_par_S._Léchauguette_et_édité_par_T
hames_&_Hudson_ISBN_2-87811-089-7 » est un *Objet Informatif* de *Type*
publication ;

- Cet *Objet Informatif* a été créé par un *Événement de Création*, de *Type* édition commerciale ;
- Cet *Événement de Création* a forcément eu une certaine durée, une *Tranche Chronologique* ;
- Cette *Tranche Chronologique* est identifiée par une *Appellation de Temps*, et cette *Appellation de Temps* est même en l'occurrence, plus spécifiquement, une *Date* ;
- Cette *Date*, de *Type* « telle que transcrite de la publication elle-même »³, est : 1995.

Tout ce cheminement peut certes paraître bien laborieux, voire inutile ; mais il ne faut pas oublier qu'une machine, c'est très bête, et que plus on lui « explique » les choses, mieux cela vaut. En réalité, tout ce cheminement est *implicitement contenu* dans la simple expression : « date d'édition = 1995 », et dans l'encodage de cette expression en : « 210 \$d 1995 ». L'avantage, c'est que si la même expression est encodée différemment (dans un autre format MARC, par exemple, ou dans un format qui n'a rien à voir avec MARC), la machine pourra interpréter cet autre encodage comme ayant *la même signification*. Ainsi, comme l'annonçait le sous-titre de la présente intervention, on s'attache à une identité de *sens* sans se laisser piéger par des différences de *forme*.

L'extrait de notice bibliographique en format UNIMARC donné ci-dessus peut se « traduire » ainsi dans le modèle CRM :

```
<Entité_CRM> Objet informatif 2-87811-089-7 <dans_la_classe> Objet Informatif
</dans_la_classe>
  <a_pour_type> Publication, monographie <dans_la_classe> Type d'Objet
Informatif </dans_la_classe> </a_pour_type>
  <est_identifié_par> 2-87811-089-7 <a_pour_type> ISBN <dans_la_classe> Type
d'Appellation d'Objet Conceptuel </dans_la_classe> </a_pour_type>
</est_identifié_par>
  <a_pour_note> br. <a_pour_type> Type de reliure <dans_la_classe> Type de
Chaîne de Caractères <fait_partie_de> ISBD(M) <dans_la_classe> Document de
référence </dans_la_classe> </fait_partie_de> </dans_la_classe> <a_pour_type>
</a_pour_note>
  <a_pour_note> ill. en noir et en coul., couv. ill. en coul. <a_pour_type>
Mention d'illustration <dans_la_classe> Type de Chaîne de Caractères
<fait_partie_de> ISBD(M) <dans_la_classe> Document de référence </dans_la_classe>
</fait_partie_de> </dans_la_classe> <a_pour_type> </a_pour_note>
  <a_pour_note> Bibliogr. p. 174-175. Index <a_pour_type> Zone des notes
<dans_la_classe> Type de Chaîne de Caractères <fait_partie_de> ISBD(M)
<dans_la_classe> Document de référence </dans_la_classe> </fait_partie_de>
</dans_la_classe> </a_pour_type> </a_pour_note>
  <a_pour_dimension> Prix <dans_la_classe> Type de Dimension </dans_la_classe>
<a_pour_valeur> 99 </a_pour_valeur> <a_pour_unité> F </a_pour_unité>
</a_pour_dimension>
  <a_pour_dimension> Importance matérielle <dans_la_classe> Type de Dimension
</dans_la_classe> <a_pour_valeur> 177 </a_pour_valeur> <a_pour_unité> p.
</a_pour_unité> </a_pour_dimension>
  <a_pour_dimension> Format <dans_la_classe> Type de Dimension
</dans_la_classe> <a_pour_valeur> 21 </a_pour_valeur> <a_pour_unité> cm
</a_pour_unité> </a_pour_dimension>
  <a_pour_titre> Jérôme Bosch <a_pour_type> Titre propre <dans_la_classe> Type
de Titre <fait_partie_de> ISBD(M) <dans_la_classe> Document de référence
</dans_la_classe> </fait_partie_de> </dans_la_classe> </a_pour_type> <a_pour_type>
```

³ En toute rigueur, le véritable *Primitif Temporel* correspondant à la date d'édition est codé ailleurs en format UNIMARC, en zone 100 (zone omise dans l'exemple ci-dessus) ; si on voulait être parfaitement précis, l'encodage « 210 \$d 1995 » se traduirait dans le modèle CRM par une instance de la classe *Inscription*, de *Type* date d'édition. Mais dans le cadre du présent exposé, ce degré de rigueur et d'exactitude n'est pas nécessaire : il suffit ici de comprendre les grands principes du modèle et les mécanismes d'interprétation sémantique des données structurées.

```

Titre significatif < dans_la_classe > Type de titre < / dans_la_classe > < / a_pour_type >
< / a_pour_titre >
  < a_pour_titre > Walter S. Gibson < a_pour_type > Première mention de
responsabilité < dans_la_classe > Type de Titre < fait_partie_de > ISBD (M)
< dans_la_classe > Document de référence < / dans_la_classe > < / fait_partie_de >
< / dans_la_classe > < / a_pour_type > < / a_pour_titre >
  < a_pour_titre > trad. de l'anglais par Sophie Léchaugnette < a_pour_type >
Mention de responsabilité suivante < dans_la_classe > Type de Titre < fait_partie_de >
ISBD (M) < dans_la_classe > Document de référence < / dans_la_classe > < / fait_partie_de >
< / dans_la_classe > < / a_pour_type > < / a_pour_titre >
  < présente_des_caractéristiques_de > Texte < dans_la_classe > Objet Linguistique
< / dans_la_classe > < type_de_similitude > imprimé < / type_de_similitude >
  < a_pour_titre > Jérôme Bosch < a_pour_type > Titre traduit < / a_pour_type >
< / a_pour_titre > < a_pour_langue > fre < / a_pour_langue >
  < a_été_créé_par > Événement de Création de l'œuvre textuelle « Jérôme
Bosch »
    < réalisé_par > Personne < est_identifié_par > Gibson, Walter S.
< dans_la_classe > Appellation d'Agent < / dans_la_classe > < dans_le_rôle_de > auteur
principal du texte < / dans_le_rôle_de > < / est_identifié_par > < / réalisé_par >
    < réalisé_par > Personne < est_identifié_par > Léchaugnette, Sophie
< dans_la_classe > Appellation d'Agent < / dans_la_classe > < dans_le_rôle_de > traducteur
< / dans_le_rôle_de > < / est_identifié_par > < / réalisé_par >
    < / a_été_créé_par >
    < est_traduction_de > Texte original < dans_la_classe > Objet Linguistique
< / dans_la_classe > < a_pour_titre > Hieronymus Bosch < a_pour_type > Titre original
< / a_pour_type > < a_pour_type > Titre significatif < / a_pour_type > < a_pour_langue > eng
< / a_pour_langue > < / est_traduction_de >
    < / présente_des_caractéristiques_de >
    < a_été_créé_par > Événement de création de la publication « Jérôme Bosch »
< a_pour_type > Édition commerciale < / a_pour_type >
    < réalisé_par > Agent < est_identifié_par > Thames & Hudson
< dans_la_classe > Appellation d'Agent < / dans_la_classe > < dans_le_rôle_de > Éditeur
commercial < / dans_le_rôle_de > < / est_identifié_par > < / réalisé_par >
    < a_eu_lieu_dans > Lieu < est_identifié_par > [Londres] < dans_la_classe >
Toponyme < / dans_la_classe > < / est_identifié_par > < / a_eu_lieu_dans >
    < a_eu_lieu_dans > Lieu < est_identifié_par > [Paris] < dans_la_classe >
Toponyme < / dans_la_classe > < / est_identifié_par > < / a_eu_lieu_dans >
    < a_pour_tranche_chronologique > Date d'édition < est_identifiée_par >
1995 < dans_la_classe > Date < / dans_la_classe > < / est_identifiée_par >
< / a_pour_tranche_chronologique >
    < / a_été_créé_par >
    < est_au_sujet_de > Personne < est_identifiée_par > Bosch, Jérôme (1450?-1516)
< dans_la_classe > Appellation d'Agent < / dans_la_classe > < a_pour_autre_forme >
Hiérónimus Van Aeken < / a_pour_autre_forme > < / est_identifiée_par > < / est_au_sujet_de >
< / Entité_CRM >

```

Comme on le voit, le fait d'explicitier la valeur sémantique *implicitement* contenue dans chacun des éléments de données d'une simple notice bibliographique (et encore, celle-ci était incomplète) gonfle démesurément le volume de cette notice. Mais que l'on se rassure : *en aucun cas* il n'est question de créer *directement* ces données dans le formalisme du CRM ; il s'agit d'établir une fois pour toutes un *mapping* entre un format donné et le CRM, pour qu'ensuite toutes les informations créées dans ce format puissent être interprétées par une machine en référence aux relations sémantiques définies par le CRM. Un des intérêts de l'entreprise, c'est de faire tomber les barrières entre les différents réservoirs de données, et notamment, comme dans l'exemple ci-dessus, entre un catalogue de bibliothèque se conformant aux ISBD et structuré en format MARC, et un inventaire de musée se conformant aux

CIDOC *Information Categories* et structuré selon un format local... C'est à bon droit qu'on a pu traiter le CRM de « ciment sémantique »⁴...

Le CRM : à quoi sert-il, à quoi peut-il servir ?

Outil de dialogue entre réservoirs de données hétérogènes

Comme l'exemple ci-dessus visait à l'illustrer, le modèle CRM peut être utilisé dans le cadre de l'élaboration d'outils de recherche intégrés et de systèmes de médiation. Lorsque, pour des raisons techniques, institutionnelles ou économiques, il n'est pas possible ou pas souhaitable d'intégrer matériellement des données dans un seul système, les systèmes de médiation permettent de *fédérer* plusieurs ressources d'informations tout en respectant la structure propre à chacune d'elles. Le but est de permettre à un utilisateur de lancer la même recherche sur toutes en même temps, par le biais d'une seule interface-utilisateur. L'une des conditions préalables à la conception d'un tel système de médiation est d'avoir procédé au *mapping* de chacune des ressources d'informations concernées vers le modèle de référence commun, en l'occurrence le CRM. C'est cette finalité qui a guidé tout le travail d'élaboration du CRM : permettre de combiner des données provenant de sources hétérogènes sans perte de sens et en préservant le degré de finesse respectif de chacune de ces sources.

C'est le type d'utilisation prévu dans le cadre du projet HEDD (*Historic Environment Data Directory*) de la commission *The English Heritage*⁵. Cette commission regroupe entre autres 22 musées, 3 bibliothèques, plusieurs dépôts d'archives ; chacun de ces établissements dispose de sa propre base de données, et ces données ne sont évidemment pas structurées de la même manière partout. Grâce au projet HEDD, le personnel de *The English Heritage* dispose d'un répertoire de toutes ces bases de données, avec une indication précise de la manière dont l'information y est structurée ; cet outil leur permet d'identifier les possibilités d'amélioration de l'accès à ces différents réservoirs de données, ainsi que les priorités à définir dans les projets de numérisation et de diffusion. Certains de ces réservoirs ont été sélectionnés pour être rendus disponibles au public via un portail unique, <http://www.britarch.ac.uk/HEIRNET> (*Historic Environment Information Resources NETWORK*).

C'est également le type d'utilisation que le RLG (*Research Libraries Group*) souhaite en faire pour sa base de données *RLG Cultural Materials* (<http://culturalmaterials.rlg.org> ; nécessite un mot de passe), réunion virtuelle des bases de données de chacune des institutions membres du Groupe, la cohérence de l'ensemble devant être assurée par le « ciment sémantique » du CRM.

⁴ En fait, l'expression fréquemment employée en anglais au sujet du CRM est : « *Semantic glue* », la « colle sémantique » ; je trouve que « ciment sémantique », outre le jeu de consonances qu'offre cette expression, passe mieux en français.

⁵ Dont le véritable nom est : *Historic Buildings and Monuments Commission*.

Outil de dialogue entre spécialistes d'horizons différents

Mais ce n'est pas la seule utilisation potentielle du modèle CRM. Le CRM peut aussi, tout simplement, aider des humains à se comprendre entre eux. Le CRM est un document de *référence* qui peut contribuer à instaurer une « base conceptuelle commune » entre différents domaines. Dans l'élaboration de projets techniques touchant au patrimoine culturel, il est indispensable d'établir une communication claire et sans ambiguïté entre les experts du domaine (historiens, archéologues, biologistes...) et les informaticiens. Il faut que les techniciens comprennent toutes les subtilités de l'information relative au patrimoine culturel, mais parallèlement il faut que les experts du domaine soient à même d'exposer leurs besoins d'une manière accessible à des spécialistes de la technologie de l'information, et d'évaluer les solutions que ces derniers leur proposent. Par l'analyse riche et approfondie que le CRM donne de l'information sur le patrimoine culturel, il peut faciliter le dialogue entre experts et techniciens. Les documentalistes peuvent y trouver une représentation formalisée de concepts qui leur sont familiers, tandis que les informaticiens peuvent y voir les grandes lignes, à un niveau très général, d'un projet de système d'information.

Outil d'aide à la conception de systèmes et extensibilité du modèle

Bien que le but premier du CRM ne soit certainement pas de fournir un format de données, cette utilisation ne saurait être exclue *a priori*. Il faut toutefois souligner qu'il n'est pas nécessaire d'implémenter la totalité du modèle en l'état. Certains de ses aspects peuvent s'avérer superflus dans le cadre d'un projet donné, ou au contraire insuffisants dans le cadre d'un autre projet. Dans le premier cas, il suffit tout simplement de ne pas en tenir compte. Dans le second cas, le formalisme orienté objet permet de concevoir des *extensions* au modèle, notamment par le biais de la création de sous-classes, mais aussi par la connexion du modèle avec d'autres ontologies.

Le CRM a notamment été utilisé par la Galerie nationale de Finlande lorsque celle-ci a voulu changer de base de données : il a servi de point de référence dans la conception de la nouvelle base, et la structure de données de l'ancienne base a été « mappée » vers les sections pertinentes du CRM. Les règles de catalogage des documents vont en outre être révisées pour prendre en compte le CRM afin d'assurer une meilleure cohérence interne de ces règles.

Quant à l'extension du modèle par connexion avec une autre ontologie, c'est ce qui a été entrepris par une chercheuse de l'Université du Queensland en Australie, Jane Hunter. Partant de la constatation qu'il existait une ontologie pour les objets matériels conservés dans les musées, le CRM, et une ontologie pour la description du contenu de ressources multimédia, MPEG-7, mais *pas* d'ontologie pour les ressources multimédia propres aux musées, elle a réuni les deux en prenant le CRM pour base de départ et en l'étendant au moyen d'entités empruntées à MPEG-7. Pour ce faire, elle a simplement défini une nouvelle sous-classe de la classe *Objet Informatif*, appelée *Contenu Multimédia*, dont plusieurs classes issues de MPEG-7 sont à leur tour des

sous-classes. Elle a également rattaché certaines classes de MPEG-7 aux classes *Élément Visuel* et *Objet Linguistique* du CRM, en tant que sous-classes.

Il est évident que le travail de Jane Hunter peut également servir aux bibliothèques, voire aux archives, pour l'interprétation sémantique des données relatives aux documents audiovisuels.

Outil de comparaison, d'échange et de stockage de données

Le CRM peut en outre servir de *référence technique* :
pour *comparer* entre eux et *évaluer* des systèmes d'information et des schémas de données. Ce type de comparaison, que l'on en soit à un stade de projet ou déjà d'implémentation, permet de mettre en lumière d'éventuelles divergences, dont un examen approfondi pourra déterminer si elles sont justifiées ou non.
pour assurer dans de bonnes conditions le *transfert de données* entre systèmes incompatibles. Le CRM peut fournir l'infrastructure sémantique d'un format de données *commun* que pourront partager différents systèmes, afin que les données puissent être transférées de l'un vers l'autre.
pour *partager* des données entre des systèmes différents. Il est plus facile d'utiliser un seul format de référence auquel on va comparer chacun des schémas de données, que de concevoir des protocoles et des mappings entre tous les systèmes deux par deux.
pour assurer l'*archivage à long terme* d'informations relatives au patrimoine culturel, grâce à des formats de données indépendants de tout système, conçus à partir d'un modèle sémantique commun tel que le CRM.

Outil d'intégration au Web sémantique

Enfin, et ce n'est certainement pas là le moindre de ses emplois potentiels, le CRM, en tant qu'*ontologie*, est très probablement appelé à jouer à terme un rôle majeur dans tous les secteurs du *Web sémantique* concernés par l'information sur le patrimoine culturel. Mais c'est là encore une autre problématique, qui sera abordée cet après-midi au cours d'une autre présentation.

Reconnaissance de dette et orientation bibliographique :

Cette présentation, surtout dans sa troisième partie, doit énormément à : Crofts, Nicholas ; Doerr, Martin ; & Gill, Tony. The CIDOC Conceptual Reference Model: a standard for communicating cultural contents. Dans : *Cultivate Interactive*[en ligne]. Issue 9. February 2003. Disponible sur World Wide Web : <<http://www.cultivate-int.org/issue9/chios/>>.

Ont en outre abondamment été consultés, utilisés, cités, voire pillés :

Sur les utilisations potentielles du CRM en général :

Doerr, Martin ; & Crofts, Nicholas. Electronic Esperanto : the role of the oo CIDOC Reference Model. Dans : *Proceedings of the ICHIM '99, Washington DC, September 22-26, 1999*. Également disponible en ligne : <http://cidoc.ics.forth.gr/docs/doerr_crofts_ichim99_new.doc>.

Doerr, Martin. *The CIDOC CRM, a standard for the integration of cultural information* [en ligne]. Rethymnon (Crète) : ICS-FORTH, 2002. Disponible sur World Wide Web : <http://cidoc.ics.forth.gr/docs/crm_for_caa_crete-new.ppt>.

Sur le formalisme de la « conversion » de données vers le CRM :

Doerr, Martin ; & Karvasonas, Iraklis. *Converting object documentation into a CRM-compatible XML form using Data Junction 7.5 (DOC to XML conversion)* [en ligne]. Rethymnon (Crète) : ICS-FORTH, May 2001. Disponible sur Internet : <http://cidoc.ics.forth.gr/docs/data_index_mapping.doc>.

Sur le projet HEDD :

Stiff, Matthew. *Semantic glue, not sticky tape : the CRM and the National Heritage in England*. Dans : *Sharing the knowledge : international symposium, Washington DC, March 26-27, 2003*. Abstract disponible en ligne : <http://cidoc.ics.forth.gr/symposiun_abstracts/list_of_abstracts.htm#stiff>.

Sur RLG Cultural Materials :

Gill, Tony. *Touring the RLG information landscape : the CIDOC Conceptual Reference Model*. Dans : *RLG Focus* [en ligne]. Issue 45, August 2000. Disponible sur World Wide Web : <<http://www.rlg.org/r-focus/i45tour.html>>.

Gill, Tony. *Using the CIDOC CRM : RLG's Cultural Materials Initiative* [en ligne]. [2000]. Disponible sur World Wide Web : <<http://www.chin.gc.ca/Resources/Cidoc/English/Presentations/tgill.ppt>>.

Gill, Tony. *When the rubber hits the road : using the CIDOC CRM in the real world*. Dans : *Sharing the knowledge : international symposium, Washington DC, March 26-27, 2003*. Disponible sur Internet : <http://cidoc.ics.forth.gr/docs/symposium_presentations/gill_2003-when-rubber_hits_road.ppt>.

Sur l'utilisation du CRM à la Galerie nationale de Finlande :

Inkari, Juha ; & Oy, Aimari. *The Finnish National Gallery database implementation* [en ligne]. [S. d.] Disponible sur Internet : <http://cidoc.ics.forth.gr/docs/FNG_CRM_5thMeeting.ppt>.

Sur la connexion CRM-MPEG7 :

Hunter, Jane. *Combining the CIDOC CRM and MPEG-7 to describe multimedia in museums* [en ligne]. [2002 ?] Disponible sur World Wide Web : <http://www.metadata.net/harmony/MW2002_paper.pdf>.

Signalons enfin qu'il existe un site Web entièrement et exclusivement consacré au CRM : <<http://cidoc.ics.forth.gr>>.