

# Séances 6-8 - Introduction au Web sémantique

## 1. Introduction - Du classement des objets au Web des connaissances

Organiser l'information n'est jamais un geste neutre. Classer, décrire, nommer et relier des objets ou des documents revient toujours à produire une certaine vision du monde, à hiérarchiser des savoirs et à orienter les usages possibles de l'information. Dans le domaine du patrimoine culturel - bibliothèques, archives, musées, collections archéologiques - ces opérations sont au cœur même des missions de conservation, de recherche et de transmission. L'histoire des bibliothèques et des systèmes documentaires montre que chaque transformation technique majeure s'est accompagnée d'une transformation profonde des modes d'organisation des connaissances.

Pendant longtemps, l'organisation des collections a reposé sur des contraintes matérielles fortes : les objets physiques occupaient un espace donné, ne pouvaient être rangés qu'à un seul endroit à la fois et n'étaient accessibles qu'à travers un nombre limité de critères. Le passage au catalogue, puis aux bases de données informatisées, a progressivement déplacé le problème : il ne s'agissait plus seulement de ranger des objets, mais de décrire ces objets à l'aide d'informations structurées, appelées métadonnées et que nous avons déjà analysées dans les séances précédentes. Ce déplacement a permis une démultiplication des points d'accès, tout en introduisant de nouvelles rigidités liées aux modèles de données utilisés.

Cette tension entre richesse du réel et rigidité des systèmes de classement a été formulée de manière particulièrement éclairante par **David Weinberger**, dans son ouvrage *Everything Is Miscellaneous* (2007). Weinberger y propose une lecture historique des modes d'organisation de l'information en distinguant plusieurs « ordres » successifs : un premier ordre fondé sur le rangement physique des objets, un deuxième ordre fondé sur le catalogue et la séparation entre objets et descriptions, puis un troisième ordre fondé sur les bases de données numériques. Cette analyse met en évidence un point essentiel : plus les systèmes d'organisation gagnent en puissance, plus ils révèlent leurs limites face à la complexité, à l'ambiguïté et à l'incomplétude du réel.

Dans les bases de données relationnelles classiques, les objets patrimoniaux sont souvent contraints d'entrer dans des schémas prédéfinis : une œuvre doit avoir un auteur, une date, un lieu ; un objet archéologique doit être rattaché à une typologie stable ; une notice doit être complète pour être exploitable. Or, le patrimoine est précisément constitué d'objets lacunaires, de chronologies incertaines, d'attributions discutées et d'interprétations multiples. Cette

inadéquation entre la complexité des objets culturels et la rigidité des modèles de données constitue l'un des problèmes centraux des humanités numériques et de la documentation patrimoniale contemporaine.

C'est dans ce contexte qu'émerge, à la fin des années 1990 et au début des années 2000, une nouvelle proposition conceptuelle et technique : le **Web sémantique**. Pensé comme une extension du Web des documents, le Web sémantique ne vise pas seulement à relier des pages entre elles, mais à relier des **données** entre elles, de manière explicite, structurée et interprétable par des machines. Il ne s'agit plus uniquement de publier des contenus en ligne, mais de rendre leurs relations intelligibles, réutilisables et combinables dans des environnements distribués.

Le Web sémantique introduit ainsi un changement de paradigme profond : les informations ne sont plus enfermées dans des bases de données isolées ou des catalogues clos, mais deviennent des éléments d'un vaste réseau de connaissances interconnectées. Les identifiants globaux (URI), les relations explicites (triplets RDF) et les principes de données liées (Linked Data) permettent de dépasser les limites des systèmes documentaires traditionnels, en offrant la possibilité de relier des ressources issues de contextes institutionnels, disciplinaires et géographiques différents.

Cependant, ce nouveau cadre soulève une question fondamentale : **comment structurer le sens des données ?** Le Web sémantique fournit des mécanismes pour relier des informations, mais il ne définit pas, en lui-même, les concepts à utiliser ni les relations pertinentes entre eux. Autrement dit, relier des données ne suffit pas ; encore faut-il savoir *quoi* relier et *comment* le faire de manière cohérente. C'est à ce niveau qu'interviennent les **ontologies**, entendues comme des modèles conceptuels formels permettant de représenter un domaine de connaissances, ses entités et les relations qui les unissent.

Dans le champ du patrimoine culturel, cette problématique est particulièrement aiguë. Les institutions patrimoniales produisent et manipulent des données extrêmement hétérogènes : objets, personnes, lieux, événements, documents, interprétations scientifiques. Pour rendre ces données interopérables à l'échelle du Web, tout en respectant leur complexité historique et conceptuelle, il est nécessaire de disposer de modèles capables de représenter des processus, des temporalités multiples et des relations événementielles. C'est dans cette perspective qu'a été développé le **CIDOC CRM**, une ontologie de référence conçue spécifiquement pour la modélisation des connaissances patrimoniales.

Cette séance propose donc d'aborder conjointement deux dimensions indissociables : d'une part, le **Web sémantique** comme changement de paradigme dans la manière de publier et de relier des données ; d'autre part, les **ontologies**, et en particulier le CIDOC CRM, comme réponse conceptuelle aux limites des catalogues et des bases de données traditionnelles dans le domaine du patrimoine (pensez notamment aux bdd TEChSSON et AGIAs dont il a été question lors des séances 2-3). Il ne s'agira pas d'entrer dans des considérations techniques complexes, mais de comprendre les principes fondamentaux qui sous-tendent ces approches,

afin de saisir pourquoi elles sont devenues centrales dans les projets patrimoniaux contemporains.

## Partie 1. Le Web sémantique : du Web des documents au Web des données

### 1.1 Du Web comme vitrine au Web comme infrastructure de connaissance

Lorsque nous consultons aujourd'hui une bibliothèque numérique, un site de musée ou une collection patrimoniale en ligne, nous avons souvent l'impression d'accéder à un vaste réservoir de connaissances. Des milliers, parfois des millions d'objets sont visibles, décrits, illustrés et accessibles à distance. Pourtant, derrière cette apparente richesse se cache une limite structurelle : le Web tel qu'il s'est développé depuis les années 1990 est avant tout un **Web de documents**, et non un Web de connaissances.

Le Web, tel qu'il a été conçu à l'origine par **Tim Berners-Lee**, repose sur un principe simple : relier des documents entre eux grâce à des liens hypertextes. Une page renvoie à une autre page, qui renvoie à une autre encore. Ce modèle a été extraordinairement efficace pour diffuser l'information à grande échelle, notamment dans les domaines scientifiques, culturels et éducatifs. Il a permis aux institutions patrimoniales de publier leurs catalogues, leurs expositions virtuelles, leurs inventaires et leurs bases iconographiques.

Cependant, dans ce modèle, **le sens reste enfermé dans les pages**. Un site peut présenter une œuvre, une fouille archéologique ou un manuscrit, mais pour un ordinateur, il s'agit essentiellement d'un document HTML parmi d'autres. Les relations entre les informations - par exemple entre une œuvre et son auteur, entre un objet et son lieu de découverte, ou entre plusieurs états successifs d'un même artefact - ne sont pas explicitement compréhensibles par les machines. Elles sont destinées à être lues et interprétées par des humains.

Or, dans un contexte patrimonial, cette limite a des conséquences directes sur la **médiation** et la **valorisation**. Les interfaces proposent souvent des listes de résultats, des filtres simples, des parcours thématiques préconstruits, mais peinent à offrir des navigations réellement contextuelles et dynamiques. Il est difficile, par exemple, de faire émerger automatiquement des liens entre des collections conservées dans des institutions différentes, ou de proposer des parcours fondés sur des relations complexes (événements, usages, transformations, circulations).

### 1.2. Un changement de perspective : penser en données plutôt qu'en documents

Le Web sémantique propose précisément de changer cette perspective. Il ne s'agit pas de remplacer le Web existant, mais de le compléter en ajoutant une couche de structuration des informations. Là où le Web traditionnel relie des pages, le Web sémantique vise à relier des données.

Dans un contexte patrimonial, cela signifie que l'on ne se contente plus de publier une notice descriptive, mais que l'on cherche à rendre explicites les éléments qui la composent : un objet, une personne, un lieu, une date, un événement, une technique, une typologie. Chacun de ces éléments peut devenir une donnée identifiable, réutilisable et reliée à d'autres données.

**Ce changement est fondamental pour la valorisation du patrimoine. Une œuvre n'est plus seulement un point fixe dans une base de données ; elle devient un nœud dans un réseau de relations.** Elle peut être reliée à d'autres œuvres, à des événements historiques, à des lieux de production ou de conservation, à des acteurs multiples. Ce réseau peut ensuite être exploité pour produire des visualisations, des parcours narratifs, des croisements inédits entre collections.

### 1.3. Les données liées (Linked Data) : relier plutôt que dupliquer

Le principe central du Web sémantique est celui des **données liées** (*Linked Data*). L'idée est simple : plutôt que de dupliquer sans cesse les mêmes informations dans des bases de données fermées, on utilise des identifiants partagés pour désigner les mêmes réalités.

Prenons un exemple patrimonial très simple. Imaginons un objet archéologique conservé dans un musée. Cet objet est associé à :

- un lieu de découverte,
- une période chronologique,
- un type,
- éventuellement à une personne ou à un groupe humain.

Dans un système classique, ces informations sont souvent stockées sous forme de texte : le nom du lieu, le nom de la période, le terme typologique. Dans un système de données liées, ces éléments peuvent être reliés à des **référentiels partagés**. Le lieu peut être relié à un identifiant géographique commun, le type à un thésaurus contrôlé, la période à une définition normalisée.

Ce mécanisme permet d'établir des liens au-delà d'une seule base de données. Un même lieu peut être partagé entre plusieurs institutions, un même type peut être utilisé dans des projets différents, une même période peut servir de point de convergence entre des corpus hétérogènes. Pour la médiation, cela ouvre la possibilité de relier des collections dispersées et de proposer des parcours transversaux à l'échelle du Web.

## 1.4. Le triplet RDF : une relation minimale mais expressive

Techniquement, le Web sémantique repose sur une structure très simple : le **triplet RDF**, composé de trois éléments :

- un **sujet**,
- un **prédicat**,
- un **objet**.

Cette structure permet d'exprimer une relation élémentaire, par exemple :

- « cet objet a été découvert à tel endroit »,
- « cette œuvre a été produite par tel artiste »,
- « ce site est localisé dans telle région ».

Ce qui est important ici, ce n'est pas la syntaxe, mais la **logique relationnelle**. Chaque information n'est plus une simple valeur dans une cellule de base de données, mais une relation explicite entre deux entités. Cette manière de structurer l'information est particulièrement adaptée au patrimoine, où les objets sont toujours pris dans des réseaux complexes de relations historiques, sociales et spatiales.

Pour la médiation, cela signifie que l'on peut passer d'une description statique à une exploration relationnelle. Un utilisateur peut suivre des liens : d'un objet vers un lieu, d'un lieu vers d'autres objets, d'un événement vers des acteurs, et ainsi construire son propre parcours de découverte.

Pour que ces relations fonctionnent à l'échelle du Web, il est nécessaire d'identifier les éléments de manière unique et stable. Comme nous l'avons déjà vu, c'est bien le rôle des **URI** (Uniform Resource Identifiers). Une URI n'est pas simplement une adresse Web ; c'est un identifiant qui permet de désigner sans ambiguïté une entité : un objet, un lieu, une personne, un concept.

Dans le domaine patrimonial, l'usage des URI est crucial pour la pérennité des données. Il permet d'éviter les ambiguïtés liées aux variations de vocabulaire, aux traductions, aux changements de nom. Un même lieu peut avoir plusieurs appellations, mais une URI unique permet de le reconnaître comme une seule et même entité.

Pour la valorisation, cela signifie que les données peuvent être réutilisées, croisées et enrichies dans le temps. Une exposition virtuelle peut s'appuyer sur des URI existantes, un projet pédagogique peut relier ses contenus à des référentiels externes, une bibliothèque numérique peut s'inscrire dans un écosystème plus large de données culturelles.

## 1.5. Pourquoi le Web sémantique est pertinent pour le patrimoine et la médiation

Le Web sémantique n'est pas une réponse purement technique. Il répond à des enjeux profondément culturels et patrimoniaux. Les objets du patrimoine ne prennent sens que par leurs relations : relations temporelles, spatiales, sociales, symboliques. Or, les systèmes documentaires traditionnels ont souvent du mal à rendre compte de cette richesse relationnelle.

En rendant explicites les relations entre les données, le Web sémantique permet :

- de dépasser la logique de la simple notice descriptive ;
- de relier des collections dispersées ;
- de proposer des parcours de médiation plus riches et plus personnalisables ;
- de favoriser la réutilisation des données dans des contextes scientifiques, pédagogiques et culturels.

Il constitue ainsi un cadre particulièrement adapté aux objectifs contemporains de valorisation du patrimoine : ouverture des données, interopérabilité entre institutions, médiation numérique enrichie, et participation accrue des publics.

**Cependant, le Web sémantique ne suffit pas à lui seul. S'il fournit les mécanismes pour relier des données, il ne dit rien du sens de ces relations. Pour structurer les connaissances patrimoniales de manière cohérente, il est nécessaire de s'appuyer sur des modèles conceptuels explicites. C'est précisément le rôle des ontologies, et en particulier du CIDOC CRM, que nous allons aborder dans la seconde partie de cette séance.**

## Partie 2 - Ontologies et CIDOC CRM : structurer le sens des données patrimoniales

### 2.1. Pourquoi le Web sémantique a besoin d'“outils de sens”

Dans la Partie 1, nous avons vu que le Web sémantique permet de relier des données sous forme de relations explicites, par exemple grâce au modèle du triplet RDF (sujet – prédicat – objet). Mais une question reste ouverte : qu'est-ce qui garantit que ces relations “veulent dire quelque chose” de cohérent ?

Autrement dit : si deux institutions publient des données, comment s'assurer qu'elles parlent bien des mêmes types de choses, avec des relations comparables, sans produire un réseau de liens incompréhensible ? C'est ici qu'intervient l'idée d'**ontologie**.

## Vocabulaire, thésaurus, ontologie : trois niveaux différents

On confond souvent, y compris dans des projets patrimoniaux, trois objets différents : **vocabulaire**, **thésaurus** et **ontologie**. Pour vous, l'enjeu n'est pas de mémoriser des définitions, mais de comprendre à quoi ça sert et en quoi ça change la médiation.

**Un vocabulaire contrôlé** est une liste de termes normalisés (par exemple, "amphore", "cratère", "statuette") qu'on décide d'utiliser de manière stable pour éviter des variantes ("amphore grecque", "amphore antique", etc.).

- Intérêt médiation : cohérence d'indexation, filtres plus fiables.

**Un thésaurus** va plus loin : il organise les termes et leurs relations (synonymes, hiérarchies, relations associatives : "terme générique / terme spécifique", etc.).

- Intérêt médiation : navigation thématique, rebonds, élargissement / raffinement des recherches.

**Une ontologie** change d'échelle : elle ne se limite pas à organiser des mots ; elle modélise des **concepts** (des "types de choses") et surtout des **relations formelles** entre ces concepts. Elle décrit quels types d'entités existent dans un domaine (objets, personnes, lieux, événements...) et quelles relations il est pertinent d'établir entre elles (produire, découvrir, conserver, déplacer, documenter...).

- Intérêt médiation : raconter des trajectoires, relier des contextes, construire des parcours "par événements", croiser des collections hétérogènes sans tout aplatir.

Autrement dit : les vocabulaires contrôlés et les thésaurus aident à mieux *nommer* et *retrouver*. Une ontologie aide à *comprendre* et *relier* (dans une logique de connaissance).

## 2.2. Pourquoi le patrimoine a besoin d'une modélisation événementielle

Les objets patrimoniaux ne sont pas seulement des "choses" avec des attributs. Ils ont une histoire : ils sont fabriqués, utilisés, réparés, déposés, découverts, collectés, exposés, prêtés, restaurés, numérisés, interprétés. Une notice "plate" (même riche en métadonnées) a du mal à représenter ces trajectoires. C'est ici qu'une idée devient centrale : décrire le patrimoine par événements. Par exemple, si l'on veut médiatiser un objet archéologique, on a souvent besoin de distinguer :

- l'**événement de production** (où, quand, par qui, dans quel contexte technique)
- l'**événement de découverte / fouille** (où, quand, par qui, dans quel contexte scientifique)
- les **événements de conservation / déplacement** (où l'objet a été, quand, sous quelle responsabilité)
- l'**événement d'exposition** (quand et comment l'objet a été montré au public)
- l'**événement de numérisation** (quand et comment le fichier a été produit)

Ces distinctions ne sont pas un luxe “pour informaticiens” : elles changent la médiation. Elles permettent de raconter l’objet non plus comme une fiche figée, mais comme une **biographie**.

## 2.3. CIDOC CRM : une ontologie de référence pour le patrimoine culturel

Le **CIDOC CRM (Conceptual Reference Model)** est précisément une ontologie conçue pour le domaine du patrimoine culturel. Le site officiel rappelle qu’il s’agit d’un travail de longue durée, reconnu comme standard ISO depuis 2006. Ce point est important : CIDOC CRM n’est pas “un modèle parmi d’autres” bricolé localement ; c’est un cadre de référence international, précisément conçu pour favoriser l’interopérabilité et l’intégration de données culturelles.

En effet, il répond à un problème typique du patrimoine : comment faire dialoguer des données issues de musées, bibliothèques, archives, et projets de recherche, alors qu’ils n’utilisent pas les mêmes bases, pas les mêmes formats, et parfois pas les mêmes concepts ? CIDOC CRM propose une solution : décrire les connaissances patrimoniales avec un petit nombre de grands types d’entités (classes) et un ensemble de relations (propriétés) suffisamment génériques pour être partagées, mais suffisamment structurées pour éviter l’ambiguïté.

### Classes, propriétés, domaines : comment lire CIDOC CRM sans se noyer

Pour des novices, le risque est de voir le CIDOC CRM comme une liste intimidante de codes (E22, P108, etc.). Ce n’est pas simple mais on va essayer de rendre les choses le plus claires possible.

Dans CIDOC CRM, les classes (souvent notées **E...**) représentent des types d’entités :

- des **objets**,
- des **personnes / groupes**,
- des **lieux**,
- des **périodes / temps**,
- des **événements / activités**,
- des **documents**, etc.

L’idée n’est pas de tout apprendre, mais de saisir que CIDOC CRM force à poser une question saine :

“Ce que je décris ici, c’est un objet ? un événement ? un document ? une interprétation ?”

Rien que cette question améliore la qualité des données et donc la médiation.

Les propriétés (souvent notées **P...**) expriment une relation entre deux classes : par exemple “a produit”, “a eu lieu à”, “a un intervalle de temps”, “a été documenté par”, etc.

Dans le CIDOC CRM, une propriété ne doit pas être comprise comme un simple « champ » dans lequel on saisirait librement une information, comme on le ferait dans un formulaire classique. **Une propriété est avant tout un lien**, c'est-à-dire une relation précise entre deux types de choses bien définis. Lorsqu'on utilise une propriété, on ne fait donc pas qu'ajouter une donnée : on affirme qu'un certain type d'entité est relié à un autre type d'entité d'une manière déterminée.

C'est pour cette raison que **chaque propriété du CIDOC possède un domaine et un co-domaine**. Le domaine indique le type de chose depuis lequel la relation peut partir, et le co-domaine indique le type de chose vers lequel elle peut mener. Autrement dit, une propriété ne peut pas relier n'importe quoi à n'importe quoi : elle relie toujours un type d'entité donné à un autre type d'entité donné. Cette contrainte n'est pas un détail technique, mais un principe fondamental de rigueur du modèle.

Cette rigueur permet d'éviter des descriptions floues ou incohérentes. Par exemple, on ne peut pas, dans le CIDOC, relier directement une date à un objet en disant simplement « date : -450 ». Une telle information est ambiguë, car elle ne précise pas de quoi il s'agit : est-ce la date de fabrication, d'utilisation, de découverte ou de restauration ? Le modèle oblige donc à passer par un événement. On dira qu'un objet a été impliqué dans un événement de production, et que cet événement, lui, possède une durée ou une date. La date n'est ainsi jamais attachée directement à l'objet, mais à l'action ou au processus qui le concerne.

Ce fonctionnement peut sembler contraignant au premier abord, mais il constitue en réalité l'une des grandes forces du CIDOC CRM. Il oblige à expliciter ce dont on parle et à situer chaque information dans une chaîne de relations compréhensible : non seulement ce qu'est l'objet, mais aussi ce qui lui arrive, quand et comment. On ne se contente plus d'empiler des données ; on construit une description raisonnée du réel. C'est précisément ce passage d'une logique de formulaire à une logique de modélisation qui fait l'intérêt scientifique du CIDOC, en particulier pour l'archéologie et le patrimoine.

## La logique événementielle au cœur du modèle

Pour comprendre l'esprit du CIDOC CRM, on peut résumer ainsi : le patrimoine se décrit mieux quand on décrit les **événements** qui relient les choses. Un exemple classique est l'événement de **Production** : plutôt que de dire "cet objet a un créateur" (formule plate), CIDOC CRM modélise une production comme une activité qui implique :

- un acteur (le producteur),
- un lieu,
- un temps,
- et un produit (l'objet).

Pour la médiation, c'est décisif : on peut raconter la fabrication, la commande, les versions, les matériaux, les contextes, et relier tout cela à d'autres objets produits dans des conditions comparables.

## Exemple patrimonial simple : ce que cela change pour une notice en ligne

Prenons un exemple volontairement simple (sans syntaxe informatique). **Dans une notice “plate”**, on aura typiquement :

- Titre : “Statuette en terre cuite”
- Date : “IVe s. av. J.-C.”
- Lieu : “Tarente”
- Auteur : “Inconnu”
- Matériau : “terre cuite”
- Provenance : “fouille 1974”
- Localisation actuelle : “Musée X”

C’est utile, mais ce n’est pas une histoire. Pour la médiation, cela donne souvent une fiche descriptive isolée.

**Avec une logique CIDOC CRM**, on distingue des événements :

- un événement de production (même si l’auteur est inconnu)
- un événement de découverte (fouille 1974)
- des événements de déplacement / conservation (provenances successives)
- un événement de numérisation (si on publie un modèle 3D, des photos, un RTI)

Résultat : on peut proposer à l’utilisateur :

- une exploration par lieux (tous les objets “découverts à...”)
- une exploration par campagnes de fouilles
- une exploration par chaînes de conservation
- une exploration par types d’événements (production, découverte, exposition)

Ce sont des parcours de médiation qui deviennent possibles parce que les relations ont été structurées.

## 2.4. Les référentiels : PACTOLS, GeoNames, etc.

Dernier point essentiel pour la séance en question : CIDOC CRM ne remplace pas les référentiels ; il les **articule**.

- Un **thésaurus** (par exemple PACTOLS) aide à stabiliser les concepts (types d’objets, techniques, périodes, thèmes).
- Un référentiel géographique (par exemple GeoNames) aide à stabiliser les lieux.

En revanche, les classes et les propriétés du CIDOC CRM fournissent la structure : elles disent comment relier un objet, un événement, un lieu, un acteur, un temps.

Autrement dit : les référentiels donnent des “valeurs fiables” ; l’ontologie donne une “grammaire” pour relier ces valeurs de manière intelligible.

## Conclusions : ontologie et outil de publication : le cas d’Omeka S et la difficulté de modéliser l’information patrimoniale

L’utilisation d’une ontologie comme le CIDOC CRM dans un environnement de publication tel que **Omeka S** met en évidence une tension structurelle entre la rigueur conceptuelle du modèle et les contraintes pratiques de l’outil. D’un point de vue théorique, le CIDOC CRM repose sur une distinction fondamentale entre les **entités** (ou classes), qui correspondent à des types de choses clairement identifiables — objets matériels, personnes, lieux, événements, périodes — et les **propriétés**, qui servent exclusivement à exprimer des relations formelles entre ces entités. Cette distinction est au cœur de la cohérence ontologique du modèle : une propriété n’a de sens que parce qu’elle relie deux classes précises, selon une logique de domaine et de co-domaine. Or, cette séparation conceptuelle essentielle n’est pas rendue visible dans l’interface d’Omeka S. Pour l’utilisateur, et en particulier pour les étudiants, les propriétés issues du CIDOC CRM apparaissent comme de simples champs applicables à une ressource, au même titre que des propriétés descriptives plus classiques. Les entités intermédiaires — notamment les événements, qui sont pourtant centrales dans la logique du CIDOC CRM — ne sont ni imposées, ni explicitement encouragées par l’outil. Il en résulte une pratique fréquente consistant à utiliser des propriétés ontologiques comme des attributs descriptifs directs, sans passer par une modélisation événementielle complète.

Cette situation ne relève pas d’une mauvaise compréhension de la part des étudiants, mais découle directement des choix de conception d’Omeka S. La plateforme n’a pas été pensée comme un éditeur d’ontologies ou comme un environnement de modélisation formelle, mais comme un **outil de publication, de valorisation et de médiation** des collections numériques. À ce titre, elle privilégie la flexibilité, la prise en main rapide et l’adaptabilité à des projets très variés. Imposer une distinction stricte entre entités et propriétés, ou contraindre systématiquement la création d’entités intermédiaires conformes au CIDOC CRM, rendrait l’outil excessivement complexe et inadapté à de nombreux usages patrimoniaux, notamment pédagogiques. Le flou fonctionnel observé — où propriétés et entités semblent, dans la pratique, de nature équivalente — n’est donc pas un défaut accidentel, mais le résultat d’un **arbitrage assumé** entre expressivité ontologique et utilisabilité.

Dans cette perspective, il est important de comprendre que l’usage “approximatif” du CIDOC CRM dans Omeka S n’est ni une erreur méthodologique, ni une trahison du modèle. Il s’agit d’un **compromis pragmatique**, qui permet de bénéficier partiellement des principes du Web sémantique — identifiants, relations explicites, interopérabilité minimale — sans entrer dans une modélisation ontologiquement exhaustive. Pour les étudiants, l’enjeu n’est donc pas de produire une modélisation parfaitement conforme au CIDOC CRM, mais de prendre conscience

des écarts entre le modèle théorique et son implémentation dans un outil de publication. Comprendre ces limites, être capable de les nommer et de les justifier, fait pleinement partie de l'apprentissage. Cela permet également de situer Omeka S à sa juste place dans l'écosystème des données patrimoniales : non comme un outil de modélisation conceptuelle complète, mais comme une interface de diffusion et de médiation, susceptible d'être complétée, en aval, par des opérations de mapping, d'alignement et de publication RDF plus rigoureuses lorsque les objectifs scientifiques ou d'interopérabilité l'exigent.

## De la théorie à la pratique dans Omeka S

Dans une ontologie comme le **CIDOC CRM**, la distinction est **fondamentale** :

- une **entité / classe** (E22 Objet matériel, E21 Personne, E53 Lieu...) → correspond à *un type de chose* ;
- une **propriété** (P108 a produit, P53 a pour localisation...) → correspond à *une relation entre deux entités*.

Or, dans **Omeka S**, cette distinction **n'est pas matérialisée dans l'interface** :

- tout apparaît sous la forme de **“propriétés”** que l'on ajoute à une **ressource** ;
- une ressource peut être :
  - un item
  - un media
  - ou une autre ressource liée, **sans que la logique ontologique soit visible graphiquement**.

Résultat : pour Omeka S, **une propriété CIDOC CRM et une entité CIDOC CRM “se ressemblent” fonctionnellement**, du point de vue de l'utilisateur. Cela a comme conséquence que les utilisateurs d'Omeka font recours aux propriétés et/ou entités CIDOC CRM **comme des quasi-synonymes**, et ce n'est ni une faute, ni une incompréhension.

Il faut être très clair (et c'est un point clé du cours) : **Omeka S n'est pas un éditeur d'ontologies, mais un outil de publication**. Il permet :

- d'utiliser des vocabulaires RDF,
- de produire des triplets,
- de relier des ressources entre elles

Toutefois, Omeka S ne force pas l'utilisateur à raisonner “ontologiquement” au sens strict. Pourquoi ? Parce qu'Omeka S vise l'usage, pas la modélisation formelle. C'est un choix assumé pour garantir une **souplesse maximale**, au prix d'un **affaiblissement de la rigueur ontologique visible**. Ainsi, l'outil **ne distingue pas visuellement** une relation entre deux

entités d'un attribut descriptif simple. Ce n'est pas une erreur conceptuelle, mais une **abstraction volontaire**, orientée médiation et publication.

Dans un monde idéal :

- on ferait une modélisation CIDOC CRM "pure",
- avec événements explicites,
- entités intermédiaires,
- validation stricte.

Dans le monde réel des projets patrimoniaux et pédagogiques :

- on fait des **compromis opératoires**.

Omeka S propose :

- une **approximation praticable** du modèle ontologique,
- suffisante pour :
  - l'interopérabilité minimale,
  - la publication sémantique,
  - la médiation numérique.

**Ce n'est pas une trahison du CIDOC CRM** mais une **traduction pragmatique**.

**En conclusion :**

- Omeka S permet l'usage des ontologies, mais **ne contraint pas la modélisation**.
- La distinction entité / propriété est **conceptuelle**, pas toujours visible dans l'outil.
- Les usages "impropres" des propriétés du CIDOC-CRM sont :
  - compréhensibles,
  - souvent inévitables,
  - parfois suffisants pour les objectifs de médiation.