

Séance 4 - Publier, citer et relier des ressources patrimoniales sur le Web

1. Pourquoi parler d'URL et d'identifiants dans un cours de patrimoine ?

La médiation et la valorisation numérique du patrimoine sont aujourd'hui au cœur des missions des institutions culturelles, des universités et des projets de recherche. Numériser, publier en ligne, rendre accessibles des collections, des archives ou des travaux universitaires est devenu une pratique courante, presque évidente. Pourtant, cette apparente évidence masque une transformation profonde de la manière dont le patrimoine est produit, transmis et interprété dans l'espace numérique.

Comme nous l'avons déjà répété lors des trois premières séances, mettre une ressource patrimoniale en ligne ne consiste pas seulement à la montrer. C'est déjà **faire un choix de médiation**. Choisir ce qui est rendu visible, la manière dont cela est présenté, le contexte dans lequel l'objet est inscrit, les relations qu'il entretient avec d'autres ressources, tout cela participe d'un discours scientifique et culturel. Le numérique n'est donc pas un simple support de diffusion, mais un espace de construction du sens.

Dans ce cadre, la valorisation numérique ne peut plus être pensée comme une simple vitrine. Un site patrimonial, qu'il s'agisse d'une collection universitaire, d'un corpus d'archives ou d'un portail de mémoires de master, engage une responsabilité scientifique comparable à celle d'une publication imprimée. Il doit permettre non seulement la consultation, mais aussi la compréhension, la citation, la réutilisation et la mise en relation des ressources qu'il expose.

C'est ici que se joue une distinction fondamentale, souvent implicite mais essentielle pour comprendre les enjeux de ce cours : la différence entre **mettre en ligne** et **publier**. Mettre en ligne consiste à rendre accessible un contenu à un instant donné, dans un contexte technique précis. Publier, en revanche, suppose d'inscrire ce contenu dans un cadre plus large, capable de dépasser les évolutions des outils, des interfaces et des sites web. Publier, c'est assumer le temps long du patrimoine et/ou de la recherche.

Or, la médiation numérique du patrimoine se heurte rapidement à une question centrale : comment garantir que ce qui est publié aujourd'hui restera identifiable, compréhensible et citable demain ? Comment faire en sorte qu'un objet, une image, un modèle 3D ou un mémoire de master ne soit pas seulement visible, mais véritablement intégré dans un espace scientifique durable ?

C'est à ce moment précis que surgissent des notions qui peuvent sembler techniques - URL, URI, identifiants pérennes - mais qui relèvent en réalité d'un **choix de médiation et de valorisation**. Identifier une ressource de manière stable, c'est lui donner une existence au-delà de la page qui l'affiche. C'est permettre qu'elle circule, qu'elle soit référencée, qu'elle soit reliée à d'autres savoirs. Autrement dit, c'est transformer un contenu numérique en ressource patrimoniale et scientifique à part entière.

La Séance 4 s'inscrit donc pleinement dans une réflexion sur la médiation numérique. Elle vise à vous donner les outils conceptuels nécessaires pour comprendre pourquoi l'identification des ressources est une condition essentielle de leur valorisation. Les URL, les URI et les identifiants pérennes ne sont pas des détails techniques : ils constituent l'infrastructure invisible qui permet à un projet de médiation patrimoniale d'exister, de durer et de dialoguer avec d'autres projets.

C'est à partir de cette perspective — celle de la médiation, de la publication et du temps long — que l'on peut désormais entrer dans l'étude des mécanismes d'identification sur le Web, et comprendre pourquoi ils sont devenus incontournables dans tout projet fondé sur Omeka, sur le Web sémantique et, plus largement, sur la science ouverte.

2. URL : l'adresse d'une ressource sur le Web

Une **URL** (*Uniform Resource Locator*) est une **adresse** qui permet d'accéder à une ressource sur le Web.

Exemples :

- une page web,
- une image,
- une notice Omeka,
- un fichier PDF,
- un modèle 3D.

Une URL indique :

- où se trouve la ressource,
- et comment y accéder.

Le point fondamental à comprendre est le suivant : une URL localise une ressource, elle ne la définit pas. Si l'architecture d'un site change, si un serveur est déplacé, si une plateforme est refondue, l'URL peut :

- changer,
- disparaître,
- ou rediriger vers autre chose.

Or, en contexte patrimonial et scientifique, une citation ne peut pas dépendre d'une architecture technique instable.

3. URI : identifier plutôt que localiser

Une **URI** (*Uniform Resource Identifier*) est un **identifiant** au sens fort :

- elle désigne une ressource de manière univoque,
- indépendamment de son emplacement technique.

Une URI peut être :

- résolue sur le Web (elle “ressemble” alors à une URL)
- ou non directement cliquable.

Toute URL est une URI, mais toute URI n'est pas une URL. Avec une URI :

- on ne pointe plus seulement vers une page,
- on désigne un **objet intellectuel** :
 - un objet patrimonial,
 - une œuvre,
 - une archive,
 - un mémoire de master,
 - un modèle 3D.

Cela permet :

- la citation scientifique stable,
- la mise en relation entre ressources,
- l'inscription dans le Web de données liées.

C'est indispensable dès que l'on quitte la logique du simple site vitrine.

4. URL et URI en pratique

Pour comprendre réellement ce que sont une URL et une URI, il faut quitter les définitions abstraites et regarder comment fonctionne un **site patrimonial concret**, par exemple un site construit avec Omeka pour publier des collections universitaires ou muséales.

Prenons le cas d'une notice d'objet patrimonial publiée dans Omeka. Lorsqu'un utilisateur consulte cette notice dans son navigateur, il accède à une **URL** du type :

<https://vergilius-s-test.univ-paris1.fr/s/vergilius/item/8153>

Cette URL permet d'afficher la notice dans une interface web donnée, avec un thème graphique, une langue, une mise en page particulière. Elle correspond à une page de consultation, destinée à la lecture humaine. Si le site change de structure, de thème, ou même de logiciel, cette URL peut être modifiée, redirigée ou disparaître. Elle est donc fonctionnelle, mais fragile du point de vue scientifique.

Or, ce que l'on cherche à identifier dans un projet de valorisation patrimoniale, ce n'est pas une page web, mais l'objet patrimonial lui-même : un moulage, une photographie, un document d'archive, un mémoire de master. C'est ici qu'intervient la logique d'**URI**.

Dans Omeka S, chaque ressource n'est pas seulement affichée : elle est identifiée. Derrière l'interface visible, la ressource possède une URI qui la désigne comme entité documentaire. Cette URI peut être utilisée dans différents contextes : affichage HTML, export en JSON-LD, interrogation via une API, ou alignement avec d'autres jeux de données. Autrement dit, la même ressource peut avoir plusieurs URL de consultation, mais une seule identité conceptuelle. Ce point est essentiel pour la médiation numérique du patrimoine. Un objet patrimonial peut être :

- montré dans une exposition virtuelle,
- intégré dans un parcours pédagogique,
- cité dans un mémoire ou un article,
- moissonné par un autre portail.

Dans tous ces cas, les URL peuvent varier, mais l'URI permet de maintenir l'unité de la ressource.

Imaginons maintenant un site patrimonial universitaire consacré aux mémoires de master. Un mémoire est mis en ligne, consultable via une URL dépendant du site, par exemple dans une rubrique « Promotions 2023 ». Deux ans plus tard, le site est réorganisé : les promotions sont restructurées, le thème change, les URL sont modifiées.

Si le mémoire n'est identifié que par son URL d'origine, toute citation devient problématique. Les liens cassent, les références deviennent ambiguës, et la ressource perd sa visibilité scientifique. Du point de vue de la médiation, cela revient à rompre la continuité du discours patrimonial.

À l'inverse, si ce même mémoire est associé à une URI stable, voire à un identifiant pérenne de type ARK, il peut être :

- déplacé sans perdre son identité,
- référencé dans d'autres sites,
- cité durablement,
- intégré dans des catalogues ou des portails nationaux.

Un dernier exemple permet de mesurer l'enjeu. Sur un site Omeka, un visiteur consulte une image d'objet archéologique via une URL conviviale, intégrée dans une exposition virtuelle. Un chercheur, de son côté, peut accéder à **la même ressource** sous forme de données

structurées, par exemple via un export JSON-LD ou une API. Ces deux accès sont différents, mais ils renvoient à **la même entité patrimoniale**.

C'est précisément cette dissociation entre l'accès (URL) et l'identité (URI) qui permet à un site patrimonial de remplir simultanément ses missions de **médiation**, de **valorisation** et de **recherche**.

À ce point de la séance, vous avez du avoir compris que si une URL sert à **montrer** une ressource, une URI sert en revanche à **désigner** une ressource, et que la valorisation numérique du patrimoine commence réellement lorsque l'on cesse de confondre les deux. C'est précisément cette distinction qui justifie l'introduction des **identifiants pérennes (ARK)** dans les projets patrimoniaux.

5. Identifiants pérennes : répondre au temps long

5.1 Le problème du temps

Le patrimoine et la recherche s'inscrivent dans le temps long. Le Web, lui, est fondamentalement instable :

- technologies évolutives,
- refontes de sites,
- changements institutionnels,
- migrations de serveurs.

Les **identifiants pérennes** sont une réponse directe à ce décalage.

5.2 Principe général des identifiants pérennes

Un identifiant pérenne :

- ne dépend pas de l'architecture technique immédiate,
- est géré par une autorité identifiée,
- peut être redirigé si la ressource change d'emplacement,
- garantit la continuité de la référence.

Exemples bien connus :

- **DOI** pour les publications scientifiques,
- identifiants d'autorité (**VIAF**, **IdRef**),
- **ARK** pour les objets patrimoniaux et documentaires.

6. Les ARK : identifiants pour le patrimoine et les données

6.1 Principe des ARK

Un **ARK** (*Archival Resource Key*) est un identifiant pérenne conçu spécifiquement pour :

- les archives,
- les bibliothèques,
- les musées,
- les collections patrimoniales.

Un ARK :

- identifie une ressource,
- peut pointer vers différentes représentations (notice, image, métadonnées),
- reste stable même si le site change.

Ce n'est **pas** une simple URL améliorée, c'est une **clé d'accès durable** à une ressource.

Les identifiants **ARK** occupent une place particulière dans l'écosystème des identifiants pérennes, car ils ont été conçus dès l'origine pour répondre aux contraintes spécifiques des institutions patrimoniales et de recherche. Créé en 2001 par la California Digital Library, le schème ARK ne vise pas seulement à fournir une adresse stable, mais contrairement à une URL classique, repose sur une **responsabilité institutionnelle clairement assumée**. Une organisation qui attribue un ARK s'engage explicitement à maintenir l'accès à la ressource, ou à défaut à fournir un substitut explicite et documenté. Cette responsabilité est au cœur du dispositif : la pérennité ne repose pas sur la technologie seule, mais sur une politique documentaire et scientifique.

La structure même de l'ARK traduit cette philosophie. L'identifiant associe un schème explicite (« ark: »), un numéro d'autorité nommante (NAAN) attribué à une institution reconnue, et un nom propre à la ressource, volontairement opaque, afin d'éviter toute dépendance à des caractéristiques susceptibles d'évoluer (titre, localisation, statut). Ce choix d'opacité n'est pas un défaut : il protège précisément l'identifiant contre les changements de description ou d'interprétation scientifique.

L'exemple de la **Bibliothèque nationale de France** illustre concrètement ces principes. En tant qu'autorité nommante ARK, la BnF attribue des identifiants à des documents numériques, des notices bibliographiques ou des ressources pédagogiques, tout en assurant la résolution des identifiants et la maintenance de l'infrastructure, notamment par l'hébergement d'une réplique du registre des autorités nommantes. Ce choix institutionnel montre que l'ARK n'est pas un simple outil technique, mais un engagement sur la durée, pleinement intégré aux missions patrimoniales.

Dans le cadre des projets universitaires évoqués dans ce TD, l'ARK transforme une mise en ligne en acte de publication scientifique durable.

6.2 Pourquoi les ARK sont adaptés aux projets universitaires

Dans le contexte de l'UFR 03 et des projets que nous avons évoqués, les ARK permettent :

- de citer durablement un travail étudiant,
- d'inscrire ces productions dans un espace scientifique,
- de leur donner un statut de **ressource référencée**, sans dépendre d'un site particulier.

C'est un enjeu **pédagogique, scientifique et institutionnel**.

Dans **Omeka S**, chaque ressource :

- possède une URL de consultation,
- peut être exposée sous différents formats (HTML, JSON-LD),
- est pensée comme une **entité identifiée**.

Mais une URL Omeka **n'est pas automatiquement pérenne**. C'est pourquoi les projets institutionnels ajoutent des **identifiants externes** (ARK, DOI, etc.).

Par ailleurs, il est important de noter que dans Omeka S, l'important n'est pas la page affichée, mais la **ressource décrite** (Item, Media, Item Set). Une même ressource peut avoir plusieurs représentations, plusieurs usages, plusieurs points d'accès. L'identifiant pérenne sert de **pivot** entre ces usages.

7. Enjeux scientifiques directs pour les étudiants

Cette séance n'est pas abstraite. Elle concerne directement le travail que je vous ai demandé. Un mémoire de master publié sans identifiant pérenne est :

- difficilement citable,
- fragile dans le temps,
- isolé du reste de la recherche.

Un mémoire doté d'un identifiant stable, de métadonnées claires, devient :

- une ressource scientifique,
- intégrable dans des portails,
- réutilisable,
- visible au-delà de son contexte immédiat.

P.S. Les ARK ne sont ni une option cosmétique, ni une contrainte administrative, mais un outil structurant du raisonnement documentaire. Ils préparent directement la transition vers la séance suivante, où l'on ne se demandera plus seulement *comment identifier une ressource*, mais *comment la décrire de manière intelligible, partageable et interopérable*.

cours 2025-26 - V. Capozzoli