



Emmanuelle Bermès
Bibliothèque nationale de France

**Des identifiants pérennes pour les ressources
numériques**
L'expérience de la BnF

5 mai 2006

La création de ressources numériques en ligne, qu'il s'agisse de numérisation, d'une édition ou tout simplement d'un billet ou un commentaire de blog, soulève la question de l'identification fiable et durable de ces données sur le réseau. Les bibliothèques sont depuis longtemps déjà confrontées aux problèmes de numérotation, qu'il s'agisse des numéros attribués au moment de la publication, comme les ISBN et les ISSN, ou des cotes qui permettent de disposer et retrouver les livres dans la collection¹. La question des identifiants n'a donc rien d'une nouveauté, mais comme dans de nombreux autres cas, la transposition des pratiques dans le domaine du numérique et en particulier celui du Web pose un certain nombre de problèmes : l'impossibilité de faire des distinctions claires par supports ou par genres, la gestion de versions et de granularité...

Pour qui est confronté aux problématiques de gestion des collections numériques, ces difficultés sont courantes et les questions qu'elles soulèvent ont déjà été posées, et parfois résolues, souvent en faisant reposer ces solutions d'une part sur les métadonnées, d'autre part sur les identifiants. S'il est clair que les identifiants sont nécessaires à tous les niveaux d'organisation d'un système de gestion des objets numériques, la façon de construire et de gérer ces identifiants soulève encore bien des débats.

Les identifiants sur le Web : objectifs, fonctions, systèmes

La première question qui se pose concerne la finalité des identifiants. La terminologie reflète déjà deux axes possibles. Le terme de référence stable, ou lien permanent, évoque la capacité d'un tiers à nommer et retrouver la ressource : autrement dit, la citabilité. Quand on parle d'identifiant unique ou d'identifiant de ressource, on est plutôt dans une problématique d'unicité et de pérennité dans le cadre d'une utilisation pratique.

La citabilité peut être facilitée par le choix d'identifiants dits sémantiques ou signifiants, c'est-à-dire qui portent en eux-mêmes du sens. Ce sens repose en général sur les métadonnées de la ressource qu'ils décrivent, par exemple son titre. Ce type d'identifiants ont été assez largement adoptés sur le Web, notamment par les outils de gestion de plateformes de blogs. Cependant, ces identifiants signifiants, s'ils facilitent la compréhension et le référencement de la ressource, peuvent poser un certain nombre de problèmes. Tout d'abord, si la nature de la ressource change, le lien sémantique entre l'identifiant et la ressource peut être brisé. De plus, les identifiants signifiants sont profondément liés à la structure des documents qu'ils décrivent, ce qui peut rapidement dans le cas d'une collection de masse poser des problèmes de cohérence et de généricité. Enfin, d'une langue à l'autre, d'une époque à l'autre, la signification d'un mot ou d'un sigle peut changer, elle peut être explicite à une époque et incompréhensible un siècle plus tard, elle peut être anodine à un moment donné et devenir offensante en moins d'une décennie. Lorsqu'on travaille sur le très long terme et à l'échelle internationale, il peut donc être important de réfléchir à ces contingences et d'envisager le choix d'un système de nommage opaque.

Les identifiants opaques sont en principe générés par des machines à l'aide de logiciels. Parmi les normes existantes, nous pouvons citer UUID² (Universally Unique Identifier), un

¹ LUPOVICI, Catherine, « Le Digital Object Identifier : Le système du DOI », *BBF*, 1998, n° 3, p. 49-54
<http://bbf.enssib.fr>

² A Universally Unique Identifier (UUID) URN Namespace, RFC n° 4122 <http://www.ietf.org/rfc/rfc4122.txt>

identifiant construit par un algorithme normalisé sur la base d'informations techniques (quel ordinateur le génère) et temporels (à quel instant précis on le génère). De tels identifiants soulèvent une toute autre difficulté : la nécessité de préserver un lien entre l'identifiant et la ressource décrite, puisque l'identifiant ne porte pas, en lui-même, d'informations sur le contenu de cette ressource. Par ailleurs, l'automatisation des UUID a permis à des logiciels de générer des identifiants très longs, ce qui les rend impossibles à exploiter pour des utilisateurs humains.

Alors que les identifiants signifiants facilitent la citabilité et donc l'accessibilité des ressources, les identifiants opaques sont plus pérennes et préférables dans un contexte où l'on vise la préservation à très long terme de ces ressources. On retrouve bien dans cette dualité un problème qui a été éternellement celui des bibliothèques : le dilemme entre communication et conservation. Nous voyons donc que ces deux contextes d'usage possibles, préservation/archivage et consultation/indexation, constituent deux problématiques différentes pour les identifiants, et que le choix d'un système permettant de résoudre de manière idéale les deux problèmes à la fois risque de se révéler utopique. Il faut donc réussir à concilier ces deux aspects en prenant en compte les fonctionnalités des identifiants à des fins d'archivage et à des fins de consultation.

Fonctionnalités des identifiants

La réflexion sur les fonctionnalités des identifiants n'est pas nouvelle et on la trouve déjà, par exemple, dans la RFC 1737 intitulée *Functional Requirements for Uniform Resource Names*, datée de 1994³. Ce document définit une série de fonctionnalités qui sont encore celles qui sont questionnées aujourd'hui⁴.

La première de ces fonctionnalités est l'unicité, un identifiant étant supposé caractériser une ressource et une seule. En retour, la même ressource, même située à différents endroits, devrait avoir le même identifiant. On parle alors d'identifiants « globalement uniques », ce qui peut supposer une organisation plus ou moins centralisée à l'échelle internationale. Dans un second temps, on recherche pour ces identifiants globalement unique la pérennité, c'est-à-dire une forme de garantie qu'ils ne changeront pas et continueront à identifier la même ressource. La pérennité est la clef de la stabilité de la référence et la principale problématique de l'utilisation des identifiants. Cependant, il est désormais reconnu que la pérennité n'est en rien un problème technique, et qu'elle n'est garantie par aucun système connu d'identification. C'est la gouvernance qui assure la pérennité, c'est-à-dire qu'elle est garantie par la pérennité de l'institution qui donne les identifiants. Les institutions ou acteurs appelés à durer peuvent ainsi devenir des "autorités nommantes", à l'échelle d'une organisation locale, d'un pays, ou du monde. Ces autorités nommantes sont dès lors détentrices d'un pouvoir, celui de nommer, mais aussi d'une lourde responsabilité, celle de pérenniser leurs identifiants.

³ <http://www.w3.org/Addressing/rfc1737.txt>

⁴ Voir notamment *Persistent Identifiers*. University College Cork, Ireland, June 17-18, 2004 <http://www.erpanet.org/events/2004/cork/Cork%20Report.pdf>, *DCC Workshop on Persistent Identifiers*. University of Glasgow, 30 June – 1 July 2005 <http://www.dcc.ac.uk/events/pi-2005/>, ou encore *NISO Identifier Roundtable*, National Library of Medicine, March 13-14, 2006 http://www.niso.org/news/events_workshops/ID-docs/ID-06-report.pdf.

L'indépendance de l'autorité nommante doit dès lors être discutée : pour une institution, il peut être raisonnable de déclarer que le système le plus pérenne est celui qui lui impose le moins de contraintes, ou celui dont les contraintes correspondent le mieux aux besoins de l'institution. Ainsi un établissement comme la Bibliothèque nationale de France accorde une grande valeur à l'indépendance à la fois technique et budgétaire de son système d'identifiants, car cette indépendance lui garantit une liberté de mise en oeuvre qui va être favorable, dans le contexte d'un tel établissement, à la pérennité. En revanche, un petit éditeur pourra préférer un système très contraint, car cette contrainte va lui apporter un confort technique (en lui fournissant des outils) et une sécurité globale (en proposant par exemple un système de continuité si l'éditeur disparaît) qui sont indispensables à une véritable pérennité.

Par ailleurs, la pérennité comme l'unicité posent un problème d'échelle. Il est désormais admis que des identifiants peuvent être réaffectés, voire détruits sur le réseau. La pérennité ne se définit donc pas par « éternellement », mais par « suffisamment longtemps » à l'échelle des besoins de l'institution qui gère la ressource. Pour prendre un exemple, un identifiant comme <http://www.lemonde.fr/> est tout à fait stable mais son contenu change quotidiennement. Peut-on alors parler d'identifiant pérenne ? Une distinction est à faire entre des ressources "abstraites", éventuellement mouvantes, et des ressources "concrètes", stables et uniques, toutes deux ayant besoin d'être identifiées. De même, il est nécessaire de définir l'échelle d'unicité recherchée : certains identifiants sont uniques au sein d'un système, et conviennent parfaitement pour un usage interne, bien circonscrit. Un tel usage n'impose pas les mêmes exigences que la diffusion en réseau, qui requiert quant à elle des identifiants globalement uniques plus complexes à gérer.

D'un point de vue technique, les systèmes d'identifiants doivent se montrer aptes à répondre aux besoins spécifiques qui ont été définis, comme nous venons de le voir, par chaque producteur en fonction de ses propres moyens.

Ainsi, les identifiants doivent être applicables à n'importe quel niveau de la ressource : la ressource elle-même mais aussi la collection dont elle fait partie, les articles qu'elle rassemble, et pourquoi pas, le paragraphe de l'article (ou le commentaire du billet), et également différentes versions d'une même ressource. Comme dans toute gestion de collection numérique, un choix initial est nécessaire entre granularité logique et physique. Il faut donc définir les différents niveaux de granularité de l'information qui doivent être identifiés, et comment ils vont se décliner dans le système d'identification : le choix peut aller de l'attribution d'identifiants complètement indépendants à chaque niveau (mais il faut alors gérer des liens entre ces niveaux d'identifiants, grâce à une carte de structure), jusqu'à un système hiérarchisé qui reflète l'organisation de la collection.

Les identifiants peuvent être capables d'intégrer des modèles préexistants pour le fournisseur qui les utilise : par exemple, les ISBN et ISSN, les cotes d'une bibliothèque, un système de nommage préexistant utilisé pour les URL ou les fichiers. La pérennité repose en outre sur la capacité à s'adapter aux changements de l'environnement, et il est nécessaire de pouvoir étendre les identifiants et les adapter au fur et à mesure de l'apparition de nouvelles ressources, des évolutions du réseau, des standards du Web, des capacités des navigateurs.

Enfin, les identifiants peuvent être actionnables, c'est à dire qu'ils fournissent une méthode pour accéder à la ressource. Aujourd'hui, seul le protocole 'http', le protocole du Web, est à même de répondre à cette exigence, car il est compréhensible pour un navigateur, l'outil que nous utilisons aujourd'hui pour parcourir le Web. Pour rendre des identifiants qui ne sont pas

basés sur le protocole 'http' actionnables, différentes méthodes ont été utilisées comme la mise en place de résolveurs, ou de greffons/plugs-ins que l'utilisateur ajoute lui-même à son navigateur. Ces résolveurs et plug-in utilisent eux-mêmes les technologies liées au protocole http pour rendre les identifiants actionnables.

Le contexte de l'identifiant pérenne doit permettre de savoir à quoi celui-ci correspond et d'accéder à la ressource elle-même. Soit on dispose d'autres informations sur la ressource, incluant éventuellement sa localisation, sous forme de métadonnées, soit on consulte une base de référence qui va donner l'adresse correspondant à l'identifiant, en passant par un service de résolution. Une solution n'exclut pas l'autre ; on peut avoir un identifiant associé à des métadonnées et en plus un résolveur qui va donner l'URL correspondante.

Certains systèmes d'identification pérenne demandent ou recommandent la saisie de métadonnées conjointement avec l'enregistrement de la ressource. C'est le cas par exemple de DOI et de ARK. D'autres systèmes sont dédiés à l'échange de métadonnées, mais incluent ou nécessitent un système d'identification pérenne pour accomplir leur rôle, qui est de donner accès à la ressource elle-même. Parmi ceux-ci on peut citer le système d'identification intégré dans le protocole OAI-PMH⁵.

Le rôle du résolveur d'identifiants est de faire correspondre au nom de la ressource son adresse réelle. Le résolveur peut être interne à l'institution qui donne les noms, ou externe et géré par une autorité indépendante⁶.

La combinaison du type de résolveur et du type de métadonnées associés à chaque système d'identification vont être un facteur de choix déterminant. Ils constituent le coeur du système, qu'il faut confronter avec les fonctionnalités attendues : par exemple, la possibilité de gérer plusieurs niveaux de granularité, la simplicité des mises à jour, la gestion des versions différentes d'une même ressource, etc.

Les systèmes d'identifiants pérennes aujourd'hui

Les systèmes d'identifiants pérennes actuels reposent sur l'existence d'une autorité nommante mondialement reconnue, qui dispose de la liberté et de l'indépendance nécessaires pour attribuer à des ressources des identifiants pérennes, uniques, et adaptables, que les navigateurs interprètent soit directement soit à l'aide d'un résolveur.

Comparer ou évaluer les différents systèmes est une tâche complexe, dès lors que, comme nous l'avons vu, aucun n'est supérieur aux autres en terme de pérennité ; ce sont plutôt les besoins des producteurs de ressources qui sont à prendre en compte dans le choix d'un système. La plupart des systèmes combinent les différentes fonctionnalités citées ci-dessus à différents niveaux. Il n'existe pas de critère de différenciation clair et net. Ils ont chacun leur façon de concevoir les choses, de les organiser.

⁵ <http://www.openarchives.org/OAI/2.0/guidelines-oai-identifier.htm>

⁶ Ex. : la Library of Congress dispose de son résolveur qui fonctionne pour les identifiants DOI et Handle. On dispose d'un identifiant comme : doi:10.1045/january2005-fox, et il suffit de le faire précéder par l'adresse du résolveur pour accéder à la ressource : <http://hdl.loc.gov/doi:10.1045/january2005-fox>. Il s'agit d'une simple fonction de résolution ; la ressource prise en exemple n'a aucun rapport avec la Library of Congress.

Les identifiants pérennes ont une syntaxe commune qui est basée sur la spécification des URI du W3C⁷. Cette syntaxe est composée de trois parties :

- un **préfixe** qui indique le contexte dans lequel l'identifiant est attribué (par ex. *http:*, *ftp:*, *urn:*, etc.)
- un élément qui permet de désigner l'**autorité nommante** qui a attribué l'identifiant au sein de ce système ; cette autorité peut être désignée par son nom, par un code attribué au sein du système, par un code ou un nom codé qu'elle détient par ailleurs (comme un nom de domaine), etc.
- enfin, le « **nom** » lui-même, c'est-à-dire une chaîne de caractères qui identifie la ressource de manière unique, au sein de ce système et pour cette autorité.

Le degré de normalisation de chacun de ces trois éléments est un élément clé de l'unicité des identifiants et de leur pérennité. Les préfixes sont enregistrés auprès de l'IANA⁸, ce qui garantit leur unicité : un préfixe qui ne serait pas enregistré risquerait d'être utilisé par différentes communautés ou à différentes époques sans que l'on s'en rende compte, ce qui menacerait l'unicité des identifiants. La désignation des autorités nommantes est propre à chaque système, mais l'utilisation d'un code numérique ou alphanumérique semble aujourd'hui la solution la plus fiable, attendu les risques de changements de noms liés aux institutions. Enfin, le nom lui-même est souvent laissé à la discrétion de l'autorité nommante, ce qui laisse présager de grandes disparités si l'on n'adopte pas rapidement des règles de bonnes pratiques. On retombe sur les questions de citabilité, de sémantique et d'opacité des noms évoquées en début de cet article.

Un système d'identifiants pour la BnF

Le choix d'un système d'identifiants pérennes à la BnF a été réalisé en prenant en compte les différentes fonctionnalités évoquées ci-dessus, et en les priorisant par rapport aux besoins de l'établissement concernant d'une part la visibilité de ses ressources sur le Web et la construction de services associés, et d'autre part la mise en place du système de gestion cohérente des documents numériques au sein d'une archive compatible OAIS⁹, dite SI numérique, en vue de leur préservation sur le très long terme.

Le système d'identifiants pérennes devait donc répondre à la fois aux besoins du SI numérique et à ceux de la communication et de la consultation des documents, dans le respect de l'existant mais en préfigurant les futures extensions (intégration du Dépôt légal de la Toile, dépôts de masters électroniques par des partenaires, numérisation de masse...). Le système devait fonctionner dans l'état actuel des techniques du Web, et favoriser l'accès aux documents et la citabilité. Il convenait d'éviter aux usagers la manipulation d'outils techniques spécifiques et de s'intégrer dans leurs pratiques documentaires sur le Web. Un des enjeux était d'assurer une meilleure visibilité et compréhension des contenus numériques pour les usagers et pour les robots d'indexation des moteurs de recherche.

⁷ RFC 3986 : <http://www.ietf.org/rfc/rfc3986.txt>

⁸ Internet Assigned Numbers Authority. La liste des préfixes enregistrés est accessible en ligne : <http://www.iana.org/assignments/uri-schemes.html>

⁹ L'Open Archival Information System, reconnu comme norme ISO 14721, est un modèle conceptuel pour la mise en place d'un système ouvert de conservation à long terme des objets numériques.

La plus grande attention a été accordée à l'indépendance du système choisi, à la fois sur le plan budgétaire pour ne pas obliger l'établissement à s'engager dans un processus coûteux qu'il serait susceptible de vouloir abandonner un jour, et sur le plan technique afin que le système puisse être immédiatement intégré à l'architecture existante.

Avec ces contraintes, plusieurs solutions ont été explorées :

- **Le nommage OAI associé à un système de résolveur tel que POI** : dans la mesure où le protocole OAI est mis en œuvre à la BnF pour l'ensemble des documents numérisés, et prochainement pour tout le catalogue, la gestion d'identifiants OAI est un acquis. Il aurait donc pu être intéressant de s'appuyer sur ce système, en lui adjoignant un résolveur, pour gérer les identifiants pérennes de la BnF. Cependant, du point de vue des fonctionnalités, l'OAI ne répondait pas à l'ensemble des besoins. Les identifiants pointent vers des métadonnées et non vers les documents eux-mêmes. La granularité d'accès se situe donc obligatoirement au niveau de la notice bibliographique. De plus, l'utilisation d'OAI et du résolveur POI requiert un système complexe de redirections, ce qui n'est pas idéal du point de vue de l'accessibilité, de l'indexabilité et de la citabilité.
- **La mise en place d'un système local semblable à celui de la NLA** : l'expérience de la bibliothèque nationale d'Australie (NLA) a montré qu'un certain succès peut être rencontré dans l'élaboration d'un système local, garantissant l'indépendance vis à vis d'un domaine encore en manque de maturité, tout en prévoyant une compatibilité possible avec les systèmes internationaux dans l'optique d'une évolution future. C'est le travail de spécifications, d'organisation de la collection et de création des outils qui a semblé ici trop lourd et peu susceptible de relever le défi d'extensibilité posé par la prise en compte imminente de nouvelles ressources numériques¹⁰.
- **Le système ARK** : c'est le choix qui a été fait finalement.

ARK (pour Archival Resource Key)¹¹ est système d'identifiants pérennes créé et maintenu par la California Digital Library. Sa particularité réside dans deux éléments ajoutés aux trois principaux éléments de syntaxes que nous avons vu ci-dessus : les autorités d'adressage (Name Mapping Authority Hostports - NMAH) et les qualifieurs.

Le NMAH est l'adresse du serveur qui va résoudre l'identifiant *ARK*. Comparativement au rôle joué par l'autorité nommante, on pourrait parler d'autorité d'adressage. Cette portion ne fait pas vraiment partie de l'identifiant, elle est optionnelle et peut-être changée. De fait, il est possible d'avoir plusieurs autorités d'adressage pour un même identifiant, ce qui permet par exemple dans le cas de la BnF de respecter le contexte de visualisation des documents : un même document peut être ainsi indifféremment vu à l'adresse <http://catalogue.bnf.fr/ark:/12148/bpt6k101412s> et à l'adresse <http://gallica.bnf.fr/ark:/12148/bpt6k101412s>. Grâce à ce procédé, il devient possible de maintenir le même nom de domaine entre l'application d'accès (le catalogue ou Gallica) et le document lui-même. Pour Gallica, cette fonctionnalité est essentielle car elle garantit la possibilité pour les moteurs de recherche de ne pas être désorientés et d'indexer correctement les documents dans le domaine Gallica.

¹⁰ Consulter à ce sujet le rapport demandé par la NLA pour le choix de son système d'identifiants en mai 2001 : Diana Dack, *Persistent identifiers*. Rapport pour la National Library of Australia, mai 2001

<http://www.nla.gov.au/initiatives/persistence/PIcontents.html>.

¹¹ Consulter la spécification du format ARK : <http://www.cdlib.org/inside/diglib/ark/arkspec.pdf>

Le qualifieur est une chaîne de caractères qui permet de qualifier ce que l'on veut de l'objet. Il est optionnel. L'autorité nommante est libre de développer une hiérarchie et rendre visible des variantes : il devient ainsi possible de développer la granularité d'un document, de faire référence à des versions distinctes, ou d'appeler des services particuliers. Contrairement au nom lui-même, le qualifieur n'est pas soumis à une garantie de pérennité ; il peut donc être signifiant et il peut également être associé à des services susceptibles de changer d'apparence ou de disparaître. On retrouve la dualité, évoquée plus haut, entre un identifiant abstrait qui est parfaitement pérenne et pointe vers un objet logique (l'identifiant ARK), et un identifiant concret qui désigne ou peut désigner une partie physiquement toujours identique de l'objet logique. Ainsi, l'identifiant `ark:/12148/bpt6k85329c` désigne l'objet logique qui correspond à la numérisation de « Wheler, George. Voyage de Dalmatie, de Grèce et du Levant » tandis que l'identifiant <http://catalogue.bnf.fr/ark:/12148/bpt6k85329c/f4.pagination> désigne l'objet physique correspondant à la 4^e image de cet objet logique, dans le contexte du catalogue et avec le service d'affichage de la pagination. De ces deux identifiants, seul le premier est pérenne ; le second contient des informations difficilement pérennisables, y compris des informations sémantiques (« catalogue », « pagination ») qui ont seulement vocation à rendre un service.

Les objets numériques appelés à intégrer le SI numérique de la BnF sont d'une grande diversité et se présentent avec leurs identifiants, précédemment attribués dans les processus de création ou de publication, qui sont de différentes sortes : des ISSN, des ISBN, des cotes de bibliothèques, des URI, des DOI, des identifiants automatiques de type UID, etc. Les processus de collecte ou de numérisation eux-mêmes sont variés : il existe trois chaînes de numérisation (une pour les livres, une pour les images et une pour l'audiovisuel) auxquelles on ajoutera la dépôt légal de la Toile et la production documentaire interne (records management). Chacun de ces processus dispose nécessairement de ses propres identifiants de production, qui logiquement sont tous différents puisqu'ils répondent à des besoins différents. En particulier, ces différentes chaînes sont confrontées à des problèmes de granularité très variés et ont chacune leur façon de les résoudre.

Au moment où tous ces objets intégreront le SI numérique, il deviendra impossible de gérer ces différents identifiants ; ils seront donc enregistrés dans les métadonnées tandis que le système attribuera un nouvel identifiant ARK qui se trouvera dès lors au centre du système. Cet identifiant est chargé de faire le lien entre les métadonnées et les objets eux-mêmes, et grâce au système de qualifieurs d'ARK, il est également possible de gérer à ce niveau la granularité des objets. Rien n'empêche, lors de l'attribution du nouvel identifiant, de réutiliser dans la partie « nom » de ARK une partie de l'identifiant de production si cela semble pertinent. En l'occurrence, la BnF a fait le choix d'utiliser des identifiants opaques alphanumériques, plus pérennes et extensibles, qui intègrent en partie les anciens codes à barres qui servent à gérer la production de numérisation.

Du point de vue de l'accès, on dispose avec le système ARK d'un identifiant pérenne au niveau de l'objet et aux différents niveaux de granularité qui le composent, ce qui va servir aux applications et à la gestion du système lui-même, mais également aux usagers. Ceux-ci bénéficient désormais d'un système stable pour accéder aux objets et aux parties d'objets, ce qui permettra d'améliorer la citabilité. En outre, nos partenaires pourront construire sur le système leurs propres applications, par exemple des interfaces d'annotation ou d'enrichissement de contenu, sans avoir à faire une copie locale des fichiers source et ce même s'ils veulent atteindre un niveau de granularité très profond. Utilisé également pour le catalogue BN Opale Plus, ARK permettra de rendre les notices adressables : il deviendra

possible de faire un lien permanent sur une notice du catalogue, ce lien pouvant être utilisé dans des applications connexes comme l'entrepôt OAI, des catalogues collectifs ou des moteurs de recherche.

Glossaire et références

URL	Uniform Resource Locator	Chaîne de caractères permettant de localiser une ressource sur le Web. Cette chaîne est précédée du préfixe correspondant à la localisation des documents sur le Web : http http://www.ietf.org/rfc/rfc1738.txt
URN	Uniform Resource Name	Chaîne de caractères permettant d'identifier (par son nom) une ressource sur le Web. Cette chaîne est précédée du préfixe URN. http://www.fags.org/ftp/rfc/rfc2141.txt
URI	Uniform Resource Identifier	Chaîne de caractères permettant d'identifier une ressource sur le Web, par sa localisation ou par son nom. Cette chaîne est précédée d'un préfixe enregistré tel que : http, urn... http://www.ietf.org/rfc/rfc3986.txt
DOI	Digital Object Identifier	Système d'identification des objets dans un environnement numérique, qui repose sur la suite logicielle Handle et est géré par l'International DOI Foundation. http://www.doi.org/
ARK	Archival Resource Key	Système d'identification créé et maintenu par la California Digital Library. http://www.cdlib.org/inside/diglib/ark/
OAI	Open Archive Initiative	Initiative à l'origine du protocole OAI-PMH, qui permet d'échanger des métadonnées sur le Web et intègre son propre système d'identification. http://www.openarchives.org/
POI	Purl-based Object Identifier	Système d'identification qui combine les identifiants attribués dans le cadre de l'OAI-PMH et le système de redirection PURL développé par OCLC. http://www.ukoln.ac.uk/distributed-systems/poi/