

# Les URL permanentes

## Un moyen de pérenniser les adresses web

**T**out le monde le sait, l'obsolescence des URL (Uniform Resource Locators) est une plaie pour qui veut répertorier des ressources publiées sur le Web. Entre les réorganisations de l'arborescence des sites – une spécialité des serveurs universitaires – et les noms de domaine dont le détenteur cesse de payer l'abonnement alors même que le site correspondant, hébergé chez un prestataire gratuit, continue son existence, il est très difficile de maintenir un répertoire d'URL à jour.

*Étienne Hustache*

Bibliothèque nationale de France  
[etienne.hustache@bnf.fr](mailto:etienne.hustache@bnf.fr)

Tant que les URN (Uniform Resource Names) ne deviendront pas une réalité, le problème ne connaîtra pas de solution pleinement efficace et on continuera à obtenir le message fatidique « 404-Fichier non trouvé ». Il existe toutefois un moyen intermédiaire de simplifier le travail de mise à jour des URL proposé depuis 1996 par le département recherche d'OCLC et qui présente l'énorme avantage d'être à la fois gratuit et simple à mettre en œuvre : les Persistent URLs ou PURL, qu'on peut traduire par « URL permanentes »<sup>1</sup>.

Le système mis en place par OCLC permet d'attribuer un nom à des ressources et de le convertir pour obtenir

1. L'essentiel de cet article doit beaucoup au document de Keith Shafer, Stuart Weibel, Erik Jul et Jon Fauser, « Introduction to Persistent Uniform Resource Locators », dont j'ai traduit de larges extraits en adaptant les exemples à partir de ma pratique personnelle. Il est consultable à l'adresse <http://purl.oclc.org/OCLC/PURL/INET96>. On trouvera aussi une présentation des PURL dans l'article de Catherine Lupovici, « Le Digital Object Identifier », *BBF*, 1998, t. 43, n° 3, p. 49-54. Si des collègues ont envie d'utiliser les PURL et souhaitent me poser des questions, ils peuvent le faire à l'adresse [TUISP@le-site.net](mailto:TUISP@le-site.net)



une étape intermédiaire. OCLC se déclare convaincu qu'il sera facile et peu coûteux d'en assurer la traduction en URN.

Par exemple, la PURL :

<http://purl.oclc.org/tuisp/contact/forum.shtml>

[protocole] [adresse du traducteur] [nom]

pourrait ressembler à ceci si l'on utilise la syntaxe des URN :

<URN:/org/oclc/purl/tuisp/contact/forum.shtml>

[autorité de nommage] [nom]

Notez que l'« autorité de nommage » dans cet exemple est simplement l'« adresse du traducteur » de l'exemple précédent écrite à l'envers avec des barres obliques à la place des points et que le « nom » reste le même. Il suffirait que les navigateurs soient rendus capables de reconnaître la syntaxe <URN:/org/oclc/purl/tuisp/contact/forum.shtml> de façon à se connecter au traducteur situé à [purl.oclc.org](http://purl.oclc.org) et le traducteur pourrait alors transformer la partie qui précède le nom [/tuisp/contact/forum.shtml](http://purl.oclc.org/tuisp/contact/forum.shtml). Les PURL ne sont pas une solution propriétaire, elles ne font pas appel à de nouveaux protocoles, seulement à des protocoles HTTP standards pour la connexion aux traducteurs et à des redirections HTTP tout aussi standard.

## Notion de domaine

Les domaines sont des subdivisions de l'espace de nommage sur un traducteur de PURL et ressemblent aux répertoires d'un système de gestion de fichiers. Le droit d'écriture à l'intérieur d'un domaine est contrôlé. Il en existe deux sortes, les « domaines supérieurs » et les « sous-domaines ».

- Comme leur nom l'indique, les domaines supérieurs occupent le niveau supérieur de l'espace de nommage d'un traducteur de PURL. Les détenteurs d'un domaine supérieur contrôlent l'accès à toute cette subdivision de l'espace de nommage. Par exemple, <http://purl.oclc.org/tuisp> est un domaine supérieur.
- Les sous-domaines sont placés sous un domaine supérieur ou bien un autre sous-domaine et peuvent être créés librement par les utilisateurs dans les domaines qu'ils contrôlent. <http://purl.oclc.org/tuisp/contact> est un sous-domaine.

## Redirections partielles

Créer une redirection partielle consiste à employer un domaine comme préfixe d'une hiérarchie locale d'URL.

Un traducteur de PURL cherchera à transformer tout ce qu'il peut dans une PURL en fonction de ce qu'il trouve dans sa base de données et y collera le reste pour le transformer en URL.

Par exemple, <http://purl.org/tuisp/corresp/> est associé à l'URL <http://tuisp.online.fr/corresp/>. Un essai de résolution de <http://purl.org/tuisp/corresp/ID.htm> effectuera la traduction en <http://tuisp.online.fr/corresp/ID.htm>

Une redirection partielle peut servir de préfixe permanent pour toutes les ressources d'un site web ou un sous-ensemble de celui-ci. Tout document stocké sur le serveur peut être consulté en ajoutant son chemin relatif (c'est-à-dire partiel) à la redirection partielle. Ainsi, un changement dans la hiérarchie de fichiers du site web ne nécessiterait que de modifier une seule redirection partielle et serait transparent pour les utilisateurs.

À titre d'exemple, j'ai dû passer un temps considérable il y a six mois à modifier dans les pages du site des TUISP toutes les URL des mémoires en texte intégral stockés sur les sites des IEP (instituts d'études politiques) de Lyon et de Rennes parce que leurs administrateurs avaient modifié la hiérarchie du site. Si j'avais pris la précaution de créer d'emblée les redirections partielles :

[http://purl.org/tuisp/iep\\_lyon/MFE2000/](http://purl.org/tuisp/iep_lyon/MFE2000/)  
 = <http://doc-iep.univ-lyon2.fr/Ressources/Documents/Étudiants/Memoires/MFE2000/>  
 et [http://purl.org/tuisp/iep\\_rennes/MFE2000/](http://purl.org/tuisp/iep_rennes/MFE2000/)  
 = <http://www.rennes.iep.fr/html/Fauvet/Mémoires/Memoires-00/Memoire-en-ligne/>  
 j'aurais pu simplifier considérablement la tâche de mise à jour...

## Enregistrement

Il est bien entendu nécessaire de s'enregistrer comme utilisateur à l'adresse <http://purl.org/maint/register.html> (ou l'adresse correspondante du serveur de PURL que vous utilisez) afin de créer un nom de login (ID) et un mot de passe, associé à un mnémonique que vous pourrez utiliser pour vous rappeler votre mot de passe si vous l'avez oublié. L'enregistrement est immédiat, vous pouvez créer votre première PURL !

## Création d'un domaine supérieur

Sur le traducteur de PURL d'OCLC, si vous vous contentez d'un simple enregistrement, toutes les PURL et toutes les redirections partielles que vous créerez le seront à l'intérieur du domaine supérieur /NET/ (<http://purl.oclc.org/NET/>)

Vous pouvez préférer disposer de votre propre domaine supérieur et en faire la demande à cette adresse : [http://purl.org/maint/request\\_domain.html](http://purl.org/maint/request_domain.html)

Contrairement à l'enregistrement d'un utilisateur, la procédure n'est pas automatique et nécessite une intervention humaine. Vous serez averti par courrier électronique de la suite donnée. Cela va très vite : j'ai demandé la création du domaine supérieur /tuisp/ l'année dernière et ai eu une réponse positive (et fort aimable) sous vingt-quatre heures. Il est toutefois fortement conseillé de rédiger toute correspondance associée en anglais...

### À quelles ressources affecter des PURL ?

Une PURL devrait correspondre à une unité logique indépendante pour laquelle il est jugé souhaitable de fournir un accès durable. De bons exemples sont une page d'accueil de site, une revue électronique, un article isolé ou une communication. Il n'est en revanche pas souhaitable d'affecter une PURL à un chapitre ou à un graphique au sein d'un document, ni à des documents de nature transitoire.

Les objets situés au sommet d'une hiérarchie qui pourrait être déplacée en bloc sont d'excellents candidats pour des redirections partielles. Il est toutefois délicat d'estimer précisément quelle est la partie d'une hiérarchie de répertoire qui n'est pas susceptible d'être modifiée. Dans les exemples que j'ai donnés plus haut pour les IEP de Lyon et de Rennes, il me paraît préférable de descendre le plus bas possible dans la hiérarchie des répertoires pour éviter de me trouver confronté à un changement de nom de l'un d'entre eux (ce qui s'est déjà produit), quitte à devoir créer une PURL pour chaque année de soutenance. Il faut calculer les risques...

### Maintenance des PURL

Il est évident qu'une PURL n'est pas mise à jour automatiquement lorsque l'URL associée est modifiée. Cela nécessite une intervention humaine et c'est pour cela que la création de PURL implique un engagement de long terme qui résiste en particulier aux changements de personnel. C'est tout sauf évident, mais on peut se conso-

ler en pensant que cela n'aggrave pas la situation existante où les URL sont citées directement...

On peut mettre à jour l'URL associée à une PURL en se connectant sur le formulaire approprié (dans le cas d'OCCLC, <http://purl.org/maint/modify.html>). Seuls les utilisateurs enregistrés peuvent modifier une PURL.

Il est également possible de mettre fin à la traduction d'une PURL en entrant une URL vide dans le formulaire de modification. Dans ce cas, un appel de cette PURL se traduira par l'affichage de l'historique des modifications.

### Risques de détournement

En matière de détournement, il en va des PURL comme des URL. On voit des URL rachetées lorsque leur détenteur cesse de payer sa cotisation et qui aboutissent dès lors à tout autre chose que ce qu'on pourrait imaginer. Ainsi, <http://www.dinant.net/> aboutit aujourd'hui sur le site d'une « Académie européenne d'herboristerie » alors qu'il correspondait bel et bien à un site consacré à la ville de Dinant en Belgique jusqu'en juillet 2001, comme le prouvent les pages d'accueil archivées sur la « Wayback Machine » :

[http://web.archive.org/web/\\*/http://www.dinant.net/](http://web.archive.org/web/*/http://www.dinant.net/)

Le procédé est régulièrement utilisé par des sites pornographiques ainsi que le montre une affaire récente où l'un d'entre eux a dû faire des excuses publiques pour avoir racheté une URL précédemment utilisée par un groupe de lutte contre la pédophilie.

C'est pourquoi il me paraît déconseillé d'utiliser des PURL qui n'ont pas été créées par vous-même ou par un membre d'un organisme que vous considérez comme fiable. Par les temps qui courent, il vaut mieux se méfier ! Mais si on fait attention, je suis persuadé, par une utilisation personnelle vieille maintenant d'un an, qu'un recours judicieux aux PURL peut considérablement simplifier la vie des concepteurs de sites et des bibliothèques qui ont entrepris d'entrer des notices de ressources Internet dans leurs catalogues.

Janvier 2002