

Les Médecins Maîtres-Toile

[Accueil](#) ▶ [Boîte à Outils](#) ▶ [Initiation au xhtml](#) ▶ [Initiation au xhtml - 3 : Rappel sur les pointeurs internet \(URI/URL\)](#)

Publié le : 1er février 2004



Imprimer cet article

Auteur :
Frédéric Senis



Initiation au xhtml - 3 : Rappel sur les pointeurs internet (URI/URL)

Cette fois, l'article sera plus court mais aussi assez dense en informations nouvelles pour les uns, oubliées pour les autres.

URL/URI kesako ?

Vous connaissez tous la barre d'adresse de votre navigateur où vous taper en règle générale "www.le-nom-du-site.com".

Le navigateur rajoute alors "http://" devant.

De même si vous tapez "ftp.le-nom-du-site.com", le voici immédiatement affublé d'un ftp:// .

Ce petit bout de texte complète ce que vous avez tapé pour former un URI (Universal Resource Identifier = Identifiant de Ressource Universel). C'est à dire un pointeur qui identifie de façon certaine et unique une ressource disponible quelque part. L'une des activités de votre butineur web préféré est de comprendre cet URI et d'en tirer ce qu'il doit en faire, quel hôte internet contacter et que lui demander.

L'URI est l'élément le plus générique. Il en existe en fait 3 sous-versions :

- ▶ L'URN : Qui permet d'appeler une ressource par son nom
- ▶ L'URC : Qui décrit les caractéristiques d'une ressource
- ▶ L'URL : Qui donne sa localisation (ou plutôt le mode d'emploi pour l'atteindre).

A titre d'exemple, un URN pourrait être votre numéro INSEE personnel (y compris pour les enfants une fois qu'ils sont numérotés). Un tel URN vous identifie vous et vous seul.

Une autre possibilité serait le numéro ISBN d'une publication qui est unique pour chacune d'entre-elle. Une telle URN s'écrirait : urn:ISBN:2-841-77132-6 [1]

L'URC permettrait de retrouver un document dont on ne connaît pas le nom mais uniquement des bribes de contenu, un peu comme on recherche une publication dont on connaît l'auteur et l'année de parution seulement.

Malheureusement, ces 2 précédentes URx ne sont pas encore popularisées ni implémentées dans les logiciels que nous utilisons... Un jour peut-être.

Reste l'URL

L'URL (Uniform Resource Location, donc) fourni au butineur l'emplacement. Ainsi l'exemple du `http://...` et du `ftp://...` sont en fait des URL c'est à dire qu'ils ne décrivent pas l'objet en lui-même mais le moyen de l'atteindre.

Voyons comment est construite une telle URL : Il s'agit toujours du même prototype :
<scheme>:<partie spécifique>

- ▶ Le *scheme* correspond au qualificatif qui permet de comprendre la *partie spécifique*.
- ▶ Le sens de la *partie spécifique* dépend donc du *scheme*.

Les schemes se déclinent en 2 catégories :

- ▶ les schemes utilisant un protocole relatif à un hôte internet. Ceux-ci se codent toujours de la même façon : {scheme}://<nom d'utilisateur>:<mot de passe>@<hôte ou adresse IP>:<port de connexion>/<extension spécifique au protocole>, chaque partie peut être vide ou inexistante

Ainsi : `http://www.mmt-fr.org/article72.html` se comprend par :

- protocole *http*
- nom d'utilisateur : *aucun*
- mot de passe : *aucun*
- hôte : *www.mmt-fr.org* (c'est le nom de l'ordinateur à connecter)
- port : *standard pour le protocole, c'est à dire 80*
- partie spécifique : *article70.html*

mais nous aurions pu avoir :

`http://fred:xxxx@truc.example.org:676/fichier/plus/complexe/a/atteindre.html ?n+est+ce+pas` qui se comprendrait par :

- protocole *http*
- nom d'utilisateur : *fred*
- mot de passe : *xxxx*
- hôte : *truc.example.org* (c'est le nom de l'ordinateur à connecter)
- port : *676*
- partie spécifique : *fichier/plus/complexe/a/atteindre.html ?n+est+ce+pas*

- ▶ Les schemes non spécifiques à un hôte se décomposent différemment. Par exemple :

`mailto:mon.adresse@email.example.com`

On note en particulier qu'il n'y a pas les 2 slashes `"/"` après les `:"`, spécifiques au modèle ci-dessus.

Bien entendu, c'est surtout le prototype `http://www.example.com/repertoire/fichier.html` qui est le plus utilisé.

Et le xhtml dans tout ça ?

En fait tout ceci n'est pas du tout spécifique au xhtml. vous pouvez taper ce type d'adresse directement dans la ligne d'adresse de votre butineur web et si vous composez déjà des pages en HTML, vous utilisez déjà les URLs. Mais les changements entre HTML et xhtml étant tout entier contenus dans des petites finesses, n'y aurait-il pas un petit détail qui pourrait nous échapper ?

Et bien **OUI** (vous auriez été déçu, non ?)

lorsque vous utilisez par exemple un script pour gérer un formulaire avec 2 champs "nom" et "motdepasse", l'URL générée est du type :

```
http://www.example.com/formulaire.cgi?nom=fred&motdepasse=xxxx.
```

Or le "&" est un caractère réservé car il permet de construire les "entités", c'est à dire les "<", ">" via "<" et ">". Il doit donc être encodé en tant lui-même qu'entité : "&" (ce qui signifie "amperstand"). L'URL ci-dessus est donc à écrire :

```
http://www.example.com/formulaire.cgi?nom=fred&amp;motdepasse=xxxx.
```

Ceci est vrai dans le cadre d'un attribut contenant une URL mais aussi dans le texte de votre document. C'est une erreur classique alors méfiez-vous...

Conclusion C'est tout pour cette fois, le prochain article sera lui aussi plus court. Il expliquera une donnée fondamentale que j'ai déjà utilisé sans expliquer en profondeur : **les entités**.

[1] Il s'agit de "HTML & XHTML - La référence", éditions O'Reilly - 2001



 [Imprimer cet article](#)

Copyright Médecins Maîtres-Toile francophones
[Espace membres](#) - [Administration](#) - [Crédits](#)
