

XPath en ligne de commande avec xmllint

classé dans *GNU/Linux*, *xml*



Vous souhaitez analyser un gros document XML ? Exécuter des requêtes d’information structurée ? Vous avez peut-être l’habitude d’utiliser un visualiseur XML ou d’effectuer des recherches de chaîne pour ce genre d’opération. Mais si vous bénéficiez d’une bonne compréhension de la structure de votre document source, alors vous pouvez envisager de le parcourir encore plus rapidement avec XPath grâce à xmllint.

XPath ? toi-même !

xmllint dispose de la diabolique option `--xpath`, laquelle permet, comme son nom l’indique, de formuler une expression **XPath**. Rappelons à ce sujet que **libxml2**, dont xmllint fait partie, reconnaît XPath 1.0 et pas (encore) **XPath 2.0**. Vos requêtes ne s’inscriront donc, au mieux, que dans un sous-ensemble de **XQuery**, mais vous conviendrez, d’une part, que ça n’est déjà pas si mal et, d’autre part, que la flexibilité de xmllint constitue un atout certain. Pour bénéficier de cette comande, vérifiez ou procédez à l'**installation de libxml2**.

Commande d’exécution de requêtes XPath

Pour formuler une requête XPath "expr" sur un document XML *document*, exécutez la commande suivante dans un terminal :

```
xmllint --xpath "expr" document
```

Et c’est tout. Pour effectuer encore plus spontanément une requête XPath, un petit alias s’impose :

```
alias xq="xmllint --xpath"
```

(*xq* pour XQuery, bien sûr ;-)

Exemples de requêtes XPath

Essayons tout de suite notre nouveau jouet sur le fichier **books.xml** proposé sur le **tutoriel XPath de W3schools**. Vous pouvez bien sûr lancer les requêtes sur le document en ligne si vous le souhaitez.

Afficher les prix de tous les livres

```
xq "/bookstore/book/price" books.xml
```

Ce qui donne :

```
<price>30.00</price><price>29.99</price><price>49.99</price><price>39.95</price>
```

Afficher les titres dont le prix est supérieur à 35

```
xq "/bookstore/book[price>35]/title" books.xml
```

Ce qui donne :

```
<title lang="en">XQuery Kick Start</title><title lang="en">Learning XML</title>
```

Afficher le titre du dernier livre

Si les balises vous piquent les yeux et que vous êtes sûr que la requête ne renverra qu’un seul élément, par exemple lors d’une requête d’identifiant avec **id()**, terminez votre expression par **text()**.

```
xq "/bookstore/book[position() = last()]/title/text()" books.xml
```

Ce qui donne :

```
Learning XML
```

Notes finales

Bien que développée depuis de nombreuses années, l’excellente bibliothèque libxml2 ne semble disposer de l’option `--xpath` que depuis une version récente (?). Cette option m’avait tellement fait défaut que j’avais cessé de l’espérer. Heureusement, le billet suivant m’est tombé dessus : **Your XML friend XPATH command line xmllint** par *Linux by Examples* << merci !

Enfin, Les expressions XPATH en ligne de commande, c'est bien, mais si vous souhaitez aller plus loin et gérer automatiquement vos traitements avec XPath (« programmatiquement »), peut-être devriez-vous essayer de formuler vos **requêtes XPath avec Python**.