

25th September 2012 Simple and easy XML parsing with lxml

Let me show you couple of ways to parse simple xml file using the [lxml](http://lxml.de/) [http://lxml.de/] library in Python. First of all, you have to install lxml library. If you are an ubuntu user like me, you can get it using synaptic package manager. Or you can download and install it from the website.

Now, let's assume that we want to parse a xml file named 'file.xml' with the following content:

```
<stores >
  <store >
    <store_id >120 </store_id >
    <city >xyz </city >
  </store >
  <store >
    <store_id >140 </store_id >
    <city >jkl </city >
  </store >
  <store >
    <store_id >150 </store_id >
    <city >def </city >
  </store >
  <store >
    <store_id >160 </store_id >
    <city >abc </city >
  </store >
  <store >
    <store_id >170 </store_id >
    <city >pqr </city >
  </store >
</stores >
```

Now I want to get all the store_id for my purpose from the file. The following python code serves the purpose:

```
from lxml import etree

doc = etree.parse('file.xml')
store_list = doc.findall('store')

for store in store_list:
    store_id = store.findtext('store_id')
    print store_id
```

Here is another way, using iterators to loop over the elements:

```
from lxml import etree

doc = etree.parse('file.xml')

for store in doc.getiterator('store'):
    store_id = store.findtext('store_id')
    print store_id
```

Posted 25th September 2012 by [Tamim Shahriar](#)

Labels: [lxml](#), [xml](#), [xml parsing](#)

0 Add a comment

Comment as: Sarod (Google ▼) Sign out
Publish Preview Notify me

