

Multiple pages

1. Here is a list of urls. Try running the script for each url in the list.

```
[ "http://www.dol.gov/olms/regs/compliance/cba/Cba_CaCn.htm",  
  "http://www.dol.gov/olms/regs/compliance/cba/Cbau_mamh.htm",  
  "http://www.dol.gov/olms/regs/compliance/cba/Cba_NfOz.htm" ]
```

2. Where did that list come from?
3. Scraping the main page
 1. Downloading the page
 2. Load into lxml
 3. Select the links (a tags)
 4. Sort the links by public/private
 5. Save public/private to the datastore
4. Loop over this list.

The final scraper is [here](#).