Courseware (/courses/MITx/6.00.1_4x/3T2014/courseware)   Updates & News (/courses/MITx/6.00.1_4x/3T2014/info)

Calendar (/courses/MITx/6.00.1_4x/3T2014/89309559b0414f6d8cbef9e48ca19f4b/)   Wiki (/courses/MITx/6.00.1_4x/3T2014/course_wiki)

iscussion (/courses/MITx/6.00.1_4x/3T2014/discussion/forum)   Progress (/courses/MITx/6.00.1_4x/3T2014/progress)

Help

GETTING STARTED

Download and save

code_ProblemSet7.zip (/c4x/MITx/6.00.1_4x/asset/code_ProblemSet7.zip): A zip file of all the files you need, including:

- `ps7.py`, a skeleton of the solution

- `ps7_test.py`, a test suite that will help you check your answers

- `triggers.txt`, a sample trigger configuration file. You may modify this file to try other trigger configurations

- `feedparser.py`, a module that will retrieve and parse feeds for you

- `project_util.py`, a module that includes a function to convert simple HTML fragments to plain text

The two modules (`feedparser.py and project_util.py`) are necessary for this lab to work, but you will *not* need to modify them. Feel free to read through them if you'd like to understand what's going on.

Canopy specific instructions: Before beginning, go to:
Edit -> Preferences -> Python tab and change the PyLab backend to Inline (SVG)

More Canopy specific instructions: Every time you modify code in `ps7.py` go to
Run -> Restart Kernel (or hit the CTRL with the dot on your keyboard)
before running `ps7_test.py`. **You have to do this every time you modify the file** `ps7.py` **and want to run the file** `ps7_test.py`, otherwise changes to the former will not be incorporated in the latter.

RSS OVERVIEW

Many websites have content that is updated on an unpredictable schedule. News sites, such as Google News (http://news.google.com/), are a good example of this. One tedious way to keep track of this changing content is to load the website up in your browser, and periodically hit the refresh button. Fortunately, this process can be streamlined and automated by connecting to the website's RSS feed, using an *RSS feed reader* instead of a web browser (e.g. Sage (https://addons.mozilla.org/en-US/firefox/addon/sage/)). An RSS reader will periodically collect and draw your attention to updated content.

RSS stands for *"Really Simple Syndication"*. An RSS feed consists of (periodically changing) data stored in an XML-format file residing on a web-server. For this project the details are unimportant. You don't need to know what XML is, nor do you need to know how to access these files over the network.

We will use a special Python module to deal with these low-level details. The higher-level details of the structure of the Google News RSS feed as described on the next page should be enough for our purposes.

Show Discussion

New Post

EdX offers interactive online classes and MOOCs from the world's best universities. Online courses from MITx, HarvardX, BerkeleyX, UTx and many other universities. Topics include biology, business, chemistry, computer science, economics, finance, electronics, engineering, food and nutrition, history, humanities, law, literature, math, medicine, music, philosophy, physics, science, statistics and more. EdX is a non-profit online initiative created by founding partners Harvard and MIT.

© 2014 edX, some rights reserved.

Terms of Service and Honor Code (https://www.edx.org/edx-terms-service)

Privacy Policy (Revised 4/16/2014) (https://www.edx.org/edx-privacy-policy)

About & Company Info

About
(https://www.edx.org/about-us)

News
(https://www.edx.org/news)

Contact
(https://www.edx.org/contact)

FAQ
(https://www.edx.org/student-faq)

edX Blog
(https://www.edx.org/edx-blog)

Donate to edX
(https://www.edx.org/donate)

Jobs at edX
(https://www.edx.org/jobs)

Follow Us

Twitter
(https://twitter.com/edXOnline)

Facebook
(http://www.facebook.com/EdxOnl

Meetup
(http://www.meetup.com/edX-Global-Community)

LinkedIn
(http://www.linkedin.com/company

Google+
(https://plus.google.com/+edXOnli