Social Web Design
社群網站設計
Darby Chang
張天豪
Bot
機器人
Bot, spider, crawler
Bot is useful

- Extract data from other sources
  - e.g. hot news from PIXNET
- Commit data to other destinations
  - e.g. vote cheating or Search Engine Optimization (搜尋引擎最佳化, SEO)
- Automation
  - e.g. backup (daily archive user data and ftp it out)

- It could be performed by-request or periodically
Tip

- Don’t worry about POST forms, usually they can also be executed via URL
- **WWW::Mechanize**
  - if the POST method is strictly required, you can always find a browser simulator
  - some services check the user agent to prevent crawlers
    - Mozilla/5.0 (Windows NT 6.1; rv:13.0) Gecko/20100101 Firefox/13.0
    - [UserAgentString.com](http://UserAgentString.com)
- A demo, the backup of zoro
- crontab
  - if you want to perform some works periodically
  - [例行性工作排程的建立](http://例行性工作排程的建立)
- **Net::Telnet**
  - if you want to deal with BBS rather than web platforms
  - [Re: 又出現一個PTT備份站了](http://Re: 又出現一個PTT備份站了)
Today’s assignment
今天的任務
Extract something from other sites

- Adopt any way mentioned in today’s class to enrich the content of your site. If you have no ideas, try providing hot news from news sites or hot keywords from search sites.

- Reference
  - Regular-Expressions.info

- Your web site (http://merry.ee.ncku.edu.tw/~xxx/ex12/ and http://merry.ee.ncku.edu.tw/~xxx/cur/) will be checked not before 23:59 6/18 (Mon). You may send a report (such as some important modifications) to me in case I did not notice your features.