Prep for Meeting #3
Date: Feb 27, 2013
Location: 625 Avenue of the Americas - 3rd Floor

Needed technologies: HTML, CSS, JavaScript, Ajax, JQuery, MongoDB, XPATH, Python

Objectives of meeting: go over work done on sprint #1, plan for sprint #2

Sprint #1: write a script to determine the status of products and generate an xpath for a product's price.

- Pre-work for sprint #1:
    - Build the blueprint of crawler;
        - Logic process and decision tree.
        - Project view.
    - Analyze top 10 non-API website to find out data structure pattern.

- Work list of sprint #1:
    - Test if the URL is available;
    - Product Status: (based on Scrapy) using www.manufactum.com as an example, get the XPath of product's price, status and name;
    - Product Price: based on a correct user input, program would generate a profile(XPath) of that website. Using this profile update the prices of products within the same website;
    - Build profiles manually;

- Problems for Sprint #1:
    - Product status: still looking for an efficient way to generate XPath for different website.
    - Product price: only considered when URL is the product page.
    - Works in pieces, need a framework to combine all parts together.

Candidate contents for Sprint #2:
- Design and implement profile database structure;
- Build top10 non-API websites' profile database;
- Design and implement connection methods between profile database and crawler;
- Determine the structure of the project and create it on Github.