

## Prep for Meeting #2

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

**Objective of meeting:** To have the client rank each of the backlog stories.

### Components and their backlog stories (lettered bullets)

I. (Website) Svpply has a web UI to manage the web crawler and see results.

- a Svpply can stop and start Web Crawler daemons.

II. (Scheduler) Svpply is capable of scheduling Downloader and Analyzer jobs.

- a Svpply is capable of limiting the resource impact of downloading web pages for updates.
- b Svpply is capable of limiting the resource impact of creating profiles for stores.
- c Svpply is capable of limiting the resource impact of analyzing profiles for stores.

III. (Downloader/Spider) Svpply is capable of calling a procedure that downloads the website and stores it for further analysis.

- a Svpply is capable of storing the product page (HTML) a new product that was just registered including: URL, Product\_ID, JS\_price, JS\_category, JS\_status, UI\_price, UI\_category, UI\_status.

IV. (Storage) Store product page and analytics for analysis.

- a Svpply has an onsite store of all product pages in its database.

V. (Profile Builder) Creates profile filters for each store domain.

- a Svpply is capable of generating a profile of tags that hold **PRICE** information for a specific store.

Tasks:

- i Svpply provides ITP team top 10 non-API stores with 10 products(URL, Product\_ID, JS\_price, JS\_category, JS\_status, UI\_price, UI\_category, UI\_status).
- ii ITP team analyzes currently used JavaScript code.
- iii ITP team gets acquainted on Scrapy, XPATH, HTML.
- iv ITP team transform XPATH results into a Scrapy filter.
- d Svpply is capable of generating a profile of tags that hold **PRODUCT STATUS** information for a specific store.
  - i ITP team creates Scrapy filter for "Page Not Found".
  - ii ITP team creates Scrapy filter for "Sold-out"
- e Svpply is capable of generating a profile of tags that hold **CATEGORY** information for a specific store.

VI. (Parser/Crawler)

- a. Svpply compares Svpply's current **PRICE** value against the results of running the store's profile against the latest product page in the web crawler's database.
- b. Svpply compares Svpply's current **CATEGORY** value against the results of running the store's profile against the latest product page in the web crawler's database.
- c. Svpply compares Svpply's current **PRODUCT STATUS** value against the results of running the store's profile against the latest product page in the web crawler's database.