Aggregating Gift Registries

Elizabeth Schmutter

Faculty Advisor: Professor Zachary Ives

Abstract

Gift registries take the guesswork out of gift-giving by enabling registrants to specify the items they wish to receive. However, purchasing a gift for someone who is registered at one or more retailers requires knowledge of which of the person’s gift registries actually contain items within some desired price range. Existing “universal gift registries” address this situation by allowing registrants to create a single, replacement registry at the universal gift registry’s website. The Gift Registry Aggregator has the added benefit of allowing users to keep registries at multiple third-party retailers. By utilizing a user-provided link to each gift registry, the aggregator automatically updates as the items are bought off the third-party site’s registry. The summarizing page allows the items to be grouped by store, by price, by user-supplied category, and alphabetically by name. Each item has a link to enable its purchase from the registry on the third-party site. The registrant can also add and delete items that are not associated with a third-party registry.

System Overview

The Aggregator parses third-party registries (links supplied by registrant) and stores the information in a database.

Information is refreshed from source registries each time a user views the registry (except when resorting).

System Highlights

- Developed using Java Servlets, deployed on Apache Tomcat
- Supports registries from six retailers
- Users can add and delete additional items not on a third-party registry
- Potential gift-givers can search for gift recipient of interest
- Table-based parsing scheme can be extended to include other retailers

Parsing Methods

- HTML parsing performed using HTMLUnit (Java)
- Table structure is used to parse lists of items, one item per row
- Nested tags traversed to get content such as links, images, etc.
- Field validation (numeric price, non-empty fields) to separate data from headings
- Basic parsing strategies employed to extract data from cells containing multiple fields

The Aggregator parses third-party registries (links supplied by registrant) and stores the information in a database.

Information is refreshed from source registries each time a user views the registry (except when resorting).

Gift-givers viewing an Aggregated Registry can sort by item, price, store, or category. Links to the source registry allow simple and safe purchases.

Registrants can also edit item category, include comments on each item, and add and delete custom items not included on a third-party registry.

Senior Project Poster Day 2009, CIS Department, University of Pennsylvania