A prototype for faceted browsing of RDF data

Eyal Oren and Renaud Delbru
DERI Galway
`firstname.lastname@deri.org`

May 15, 2006

1 Faceted browsing

Faceted browsing is a superior exploration technique for large structured datasets [4], useful when users do not know the data that they are looking for. In faceted browsing, the information space is partitioned using orthogonal conceptual dimensions of the data (called facets). Each facet has multiple restriction values; users select a restriction value to constrain relevant items in the information space.

Existing approaches such as Flamenco [4] and Ontogator [1], cannot navigate arbitrary datasets but are limited to manually defined facets over predefined data structures. We have defined an approach for automatic construction of facets in semi-structured RDF data [3]. We extend the notion of faceted browsing to (graph-based) RDF data and define metrics for automatic ranking of facet quality. Our technique works for arbitrary RDF data, without the need to conform to any schema.

In this paper we present the accompanying prototype implementation.

2 Prototype

The prototype implementation is available at `http://browserdf.org`. A screenshot, showing an RDF dataset[1] representing the FBI most-wanted fugitives, is shown in Fig. 1. The interface is automatically generated for this data, and works similarly for arbitrary data.

On left-hand side of Fig. 1 we see the various facets in the dataset, e.g. alias, eye-color, hair-color, or nationality. The user can select any facet and restrict it, either by directly choosing restriction values (for literal values) or by indirectly (for resources which themselves have properties that are facets). At the top the currently applied constraints are shown, and the main part displays the resources that conform to the current constraints.

The user is prevented from following dead-ends: at each step, only those facets and restriction values that have non-empty results are available. Searching with keywords is also available, either within the current selected resources or in the whole set.

The screenshot shows a simple restriction-value over the weight of a person. The prototype offers several other types of constraints, such as arbitrary joins, arbitrary selection focus, property inversion, and intersection of these.
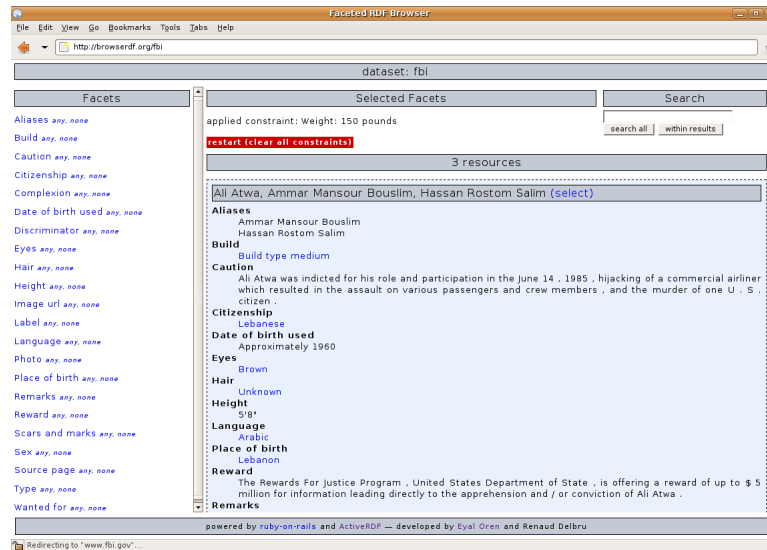
---

[1]`http://sp02.stanford.edu/kbs/fbi.zip`

Figure 1: Faceted browser showing FBI data

3   Implementation

The prototype is implemented using the web application framework Ruby-on-Rails[2]
and the object-oriented RDF API ActiveRDF[3] [2]. It works on arbitrary RDF data
sources through ActiveRDF datastore adapters: the YARS adapter is most suit-
able since it supports aggregation queries and keyword queries; arbitrary SPARQL
datasets can be displayed as well, but are relatively slow due to the lack of aggre-
gation queries.

We are currently performing an comprehensive evaluation, about which we will
report in future work. Preliminary evaluation results look promising: the metrics
give priority to the most important predicates, the automatically constructed in-
terface gives a usable overview of the dataset and users find the interface highly
usable.

References

[1] E. Hyvönen, S. Saarela, and K. Viljanen. Ontogator: Combining view- and
    ontology-based search with semantic browsing. In *Proceedings of XML Finland*.
    2003.

[2] E. Oren and R. Delbru. ActiveRDF: Object-oriented RDF in Ruby. In *Scripting
    for Semantic Web (ESWC)*. Jun. 2006.

[3] E. Oren, *et al.* Annotation and navigation in semantic wikis. In *SemWiki in
    ESWC*. 2006.

[4] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image
    search and browsing. In *CHI*. 2003.

---

[2]http://rubyonrails.org
[3]http://activerdf.org