External Queries Extension
Querying external sources using inline queries

Basil Ell
Overview

- Introduction
- Demarcation
- Example
- Main ideas
- Outlook

This work has been partially funded by EU IST FP7 project ACTIVE under grant 215040.
Yet another data integration extension?

- There is
  - SPARQLExtension
  - External Data
  - Wibit/Exhibit
  - SPARQLing
  - SMW+ Data Import Extension, Wiki Web Services
  - RDFIO
  - LinkedWiki

...
What is the difference?

- All extensions require the user to learn a certain syntax
- My approach: using inline queries and translating them into the remote query language
  - No new syntax: querying external sources using ASK
  - No import, no replication
  - Unified view on disparate data sources
  - Mapping information is separated from queries. Therefore it becomes reusable. Few users contribute, the community benefits.
  - Display query results using usual result printers.
Simple Example

{{#ask: [[Category:Person]]
 | format=table
}}

{{#ask: [[Category:Person]]
 | source=semanticweb.org
 | format=table
}}

{{#ask: [[Category:Person]]
 | source=freebase
 | format=table
}}
**Complex Example**

```
{{#ask: [[Karlsruhe]] [[Category:City]] [[Located in::Germany]]
 | ?Population
 | source=freebase
}}
```

MQL

```
{
  "/location/statistical_region/population" : [{
    "number" : null
  }],
  "/type/object/name" : "Karlsruhe",
  "/type/object/name" : "/location/citytown",
  "/location/location/containedby" : "Germany"
}
```

- Population → /location/statistical_region/population;number
- Category:City → /location/citytown
- Located in → /location/location/containedby

Page *Property:Population*

[[freebase pr mapping::/location/statistical_region/population;number]]

Page *Category:City*

[[freebase category mapping::/location/citytown]]

Page *Property:Located in*

[[freebase property mapping::/location/location/containedby]]
Configuring a new RemoteSMWStore

class SemanticWebOrgStore extends RemoteSMWStore {

    public function __construct() {

        $this->URLPart = "http://semanticweb.org/wiki/Special:Ask?title=Special%3AAsk";

        $this->MP = new MappingProvider();
        $this->MP->setGlobalProperties(
            "semanticweb.org page mapping",
            "semanticweb.org category mapping",
            "semanticweb.org property mapping",
            "semanticweb.org pr mapping",
            "semanticweb.org property type"
        );

        $this->MP->setLocalProperties(
            "Local semanticweb.org page mapping",
            "Local semanticweb.org category mapping",
            "Local semanticweb.org property mapping",
            "Local semanticweb.org pr mapping",
            "Local semanticweb.org property type"
        );

        $this->MP->setRemCategory("Category"); // Semanticweb.org uses an English SMW
        $this->MP->setLocCategory("Kategorie"); // Local SMW is a German SMW

    }
}
Conclusions

- Using external data as if it was internal data: querying, displaying using various result printers
- Import not necessary (but possible)
- Provenance of external data remains known, separation between external and internal not blurred
- No replication
- External data may also be imported/replicated -> stored in properties
- Live view on external data
- Reuse of mappings: some users add mappings, all users benefit
- Locality of each mapping property is well-defined
Outlook

- SPARQL Mediator?
- Performing aggregation on the results of mediators
- Global view on disparate data sources
- SMW knows where to retrieve which information
- Web of wikis: querying each other, sharing data, sharing mappings
- Integration with in-place editing?
Status

- Soon to be released
- Extension still buggy
- Demo wiki available:

www.basil-ell.de/mediawiki/ExternalQueries
Questions?