GR4PHP: A Programming API for Consuming E-Commerce Data from the Semantic Web

First Workshop on Programming the Semantic Web, ISWC 2012
November 11-15, 2012, Boston, USA

Alex Stolz, Mouzhi Ge and Martin Hepp
{firstname.lastname}@unibw.de

E-Business and Web Science Research Group
Universität der Bundeswehr Munich, Germany

Structured E-Commerce Data (GoodRelations)

Simple, focused crawl:

- 2625 shops with GoodRelations data
- 1382 product offers per shop on average
- > 3.6 Mio. offers in total

Source: Focused crawl, May 2012

Web developers would love to take advantage of this rich data corpus!

Web development based on Semantic Web data lags behind opportunities

1. Domain knowledge (e.g. vocabulary patterns) and SPARQL skills required to formulate useful queries on Semantic Web data
2. Semantic Web programming frameworks mainly help experts only

<u>But</u>: Web technology standards are widely used and understood even by ordinary Web developers

Web Technology Trends

Web APIs (REST)

```
GET /shop/item123
```

Programming Libraries

```
include "shop.php";
$item123 =
 $shop->get("item123");
```

Similar Approaches

- NLP and Question Answering for SPARQL
  - Suffering inaccuracy
- Query Builders and Query Assistance
  - Integration problem for Web development
- Programming Frameworks
  - Little support for high-level data (conceptual level) Semantic Web applications

Our Approach

- Tap into existing Web technology standards
  - PHP programming library to alleviate consumption of e-commerce data on the Semantic Web
- Take advantage of Semantic Web benefits
  - E.g. decentralized architecture, flexibility of integrating diverse data sources, etc.

Benefits of betting on Web standards

- Eliminate / reduce learning efforts
  - Documentation, community support and reference implementations are already in place
- Hide complexity of Semantic Web
  - No RDF, SPARQL, GoodRelations skills needed

Implementation

- Powerful PHP library
  - Developer-friendly, straightforward programming API with uniform method signatures
  - Function calls translated to proper SPARQL queries
  - Queries cover most recurring data patterns

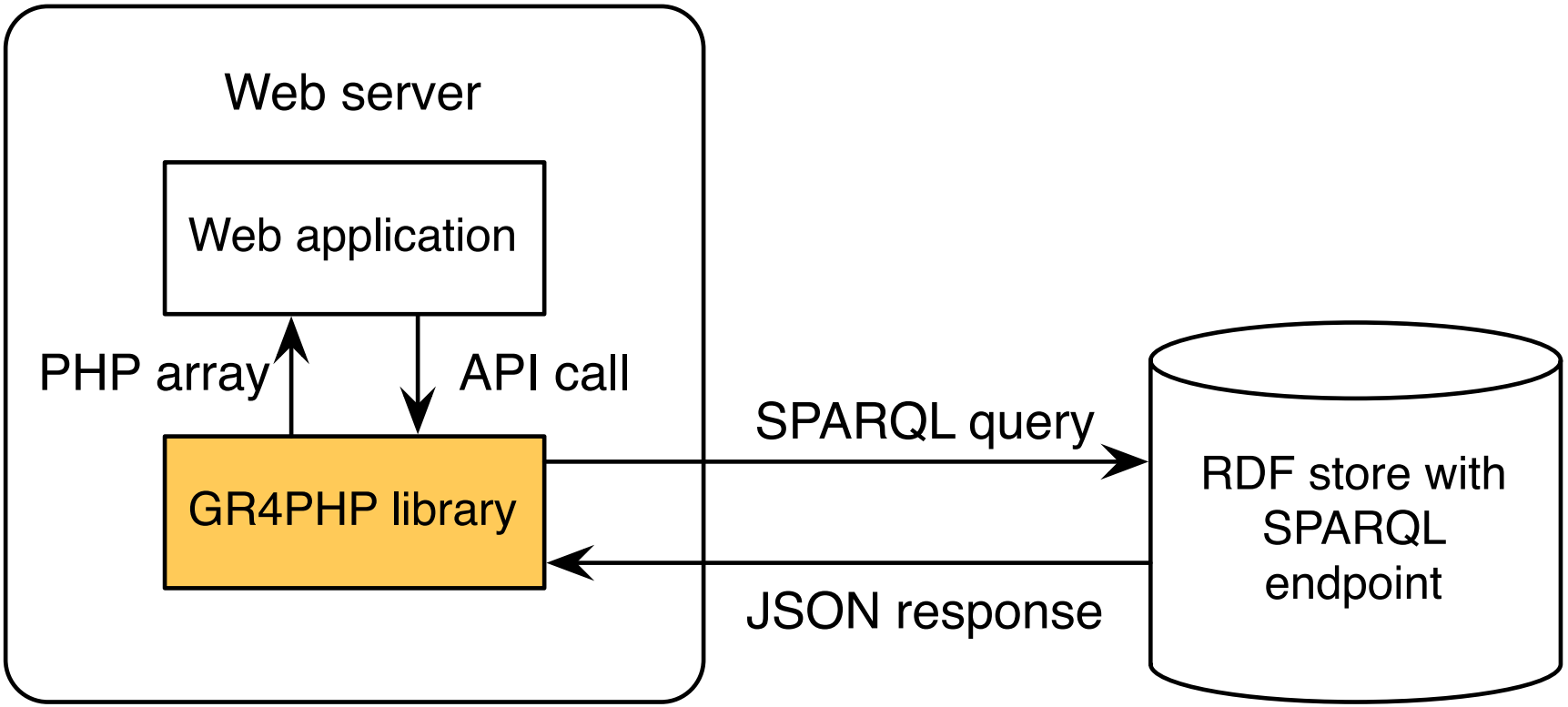
- Precise query mechanism under the hood

Project page:
http://code.google.com/p/gr4php/

PHP Functions

Function name	GoodRelations concept	Explanation
getCompany	BusinessEntity	*Information about legal entities*
getOffers	Offering	*List product or services offerings and relevant details*
getProductModel	ProductOrServiceModel	*Product model details, extensible to support custom product features*
getStore	Location	*Information about point of sales with option for filtering out stores closed at query time*
getLocation	Location	*List of nearby locations based on geo position, location name or GLN*
getOpeningHours	OpeningHoursSpecification	*List of opening hours for a given point of sale*

Conceptual Architecture



GR4PHP API

```
function get*(array $inputArray,                  // constraints array
  [array $wantedElements=FALSE],                  // selection array
  [string $mode=Configuration::MODE_LAX],         // search mode
  [int $limit=Configuration::LIMIT],              // result limit
  [array $searchProperties=FALSE])                // custom properties
```

GR4PHP API – Example

- Query n=10 offering descriptions with textual properties (gr:name, rdfs:label, etc.) that contain the search term ”Camcorder” and cost not more than 200 dollars.

GR4PHP API – Example (2)

```
require_once("gr4php.php"); // include library
$connection = new GR4PHP(Configuration::ENDPOINT_URIBURNER);

$result_array = $connection->getOffers(
  array("title" => "Camcorder",
    "maxPrice" => 200,
    "currency" => "USD"), // input
  array("uri", "title", "price", "currency"), // output
  Configuration::MODE_LAX, // rigidity of search
  10 // maximum result limit
);
```

GR4PHP API – Example (3)

	uri	title	price	currency
1.	...	Flip Video UltraHD Camcorder – Black/Chrome	199.99	USD
2.	...	...	...	...

GR4PHP – Example (4)

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX gr: <http://purl.org/goodrelations/v1#>
PREFIX dc: <http://purl.org/dc/elements/1.1/>

SELECT DISTINCT ?uri,?title,?price,?currency
WHERE {
 {?uri rdfs:label ?title. FILTER(contains(?title, 'Camcorder'))} UNION
 {?uri gr:name ?title. FILTER(contains(?title, 'Camcorder'))} UNION
 {?uri gr:description ?title. FILTER(contains(?title, 'Camcorder'))} UNION
 {?uri rdfs:comment ?title. FILTER(contains(?title, 'Camcorder'))} UNION
 {?uri dc:title ?title. FILTER(contains(?title, 'Camcorder'))}

 ?uri gr:hasPriceSpecification ?pricespec.
 ?pricespec gr:hasCurrencyValue ?price. FILTER(?price <= 200)
 ?pricespec gr:hasCurrency ?currency. FILTER(?currency = 'USD')
 ?uri a gr:Offering.
} LIMIT 10
```

Features and Restrictions

- SPARQL 1.1
- Validation of arguments passed through library functions
- Extensibility mechanism to meet for sophisticated use cases including external vocabularies

Results

Table 3. Times spent (mm:ss) for the completion of the assignments

	Group A							Group B		
	1	**2**	**3**	**4**	**5**	**6**	**ø**	**1**	**2**	**ø**
Task 1	3:15	3:30	8:00	2:30	6:00	16:00	6:32	11:30	15:00	13:15
Task 2	1:30	2:15	5:00	2:00	3:00	5:00	3:07	5:30	7:00	6:15

- Developer performance increased significantly by use of the library (group A) with respect to manually crafting SPARQL queries (group B)
- Independent-samples t-tests compared times of groups A and B → $p<0.05$ for task 1 and task 2

Applications

1. Implementation of and integration of semantic data into Web pages based on e-commerce data
2. Mobile application development taking into consideration contextual information (e.g. geo data)
3. Novel product search engines

Adoption by Mobile Web App Project

“Ravensburg App“

http://www.lieber-ravensburg.de/developer/

Acknowledgments

The work on GoodRelations and this paper have been supported by the German Federal Ministry of Research (BMBF) by a grant under the KMU Innovativ program as part of the Intelligent Match project (FKZ 01IS10022B).

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Thank you!

Martin Hepp
martin.hepp@unibw.de

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Project page:
http://code.google.com/p/gr4php/