

Research Group “Agile Knowledge Engineering and Semantic Web”

Founded 2006, initially hosted by the chair for Business Information Systems at Universität Leipzig

Now transition to Institute for Applied Informatics (InfAI)



- “An-institute” at Universität Leipzig
- Combines competences and resources of 8 University chairs from Computer Science and Economics faculties as well as industry and sponsors

AKSW aims:

- Contributing to the advancement of science in Semantic Web, Knowledge Engineering, Software Engineering
- Cost efficient, high-impact R&D, which proves usefulness at an early stage
- Bridge the gap between research results and applications

Main AKSW results: **OntoWiki** (Semantic Wiki), **DBpedia** (Wikipedia knowledge extraction), **DL-Learner** (Ontology Machine Learning), **Triplify** (light-weight semantic integration)

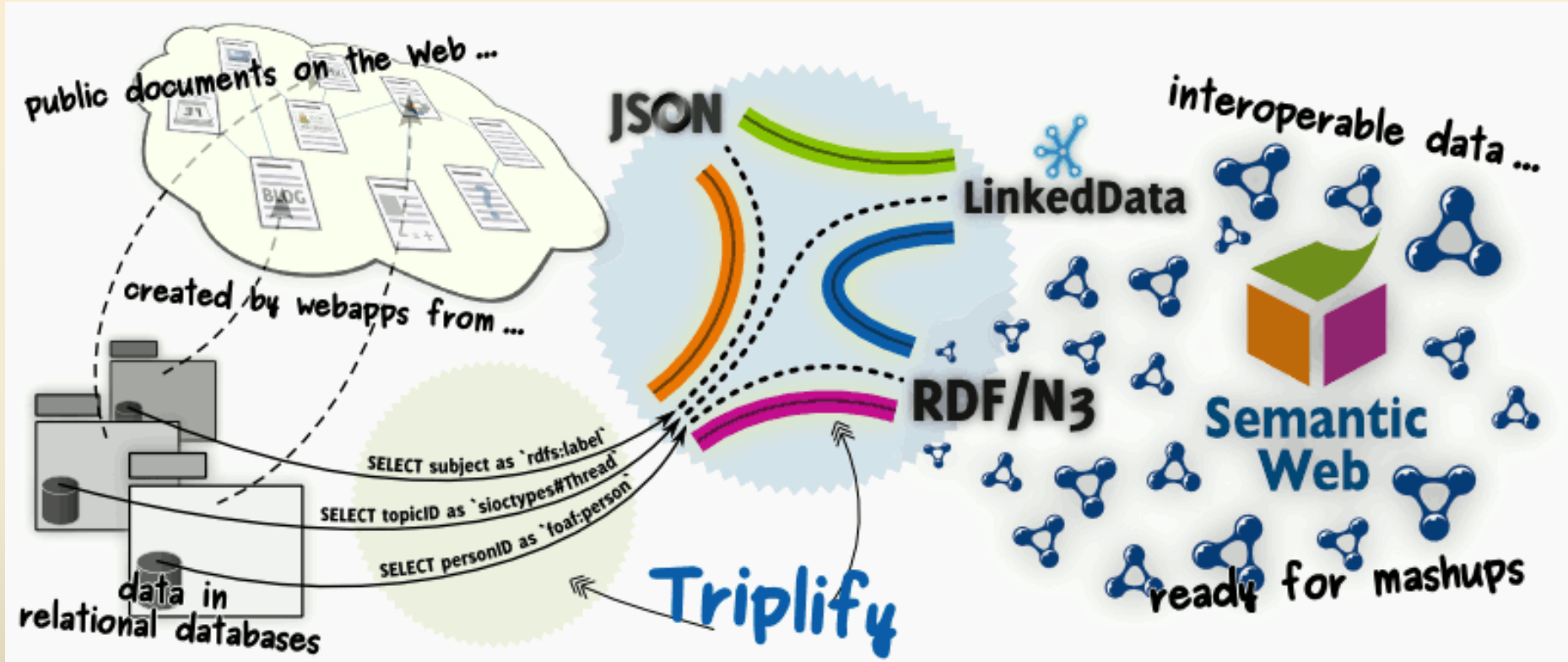
AKSW actively educates students in Semantic Technologies and serves the community by (co-) organizing events such as Conference on Social Semantic Web, I-Semantics, Scripting for The Semantic Web workshop series etc.

Triplify Motivation

- growth of semantic representations is probably still outpaced by the growth of traditional Web pages
- **overcome the chicken-and-egg dilemma of missing semantic representations and search facilities on the Web**
- Triplify tackles this dilemma by leveraging relational representations behind existing Web applications
- These are often open-source and deployed hundred thousand times
- structure and semantics encoded in relational database schemes (behind Web apps) is not accessible to Web search engines, mashups etc.

Project	Area	Downloads
phpBB	discussion forum	235480
Gallery	photo gallery	166005
XOOPS	CMS	115807
Coppermine	photo gallery	113854
Typo3	CMS	63641
Liferay Portal	Portal	39615
eGroupWare	group ware	33865
Alfresco	CMS	31914
e107	CMS	19996
Lifetype	Blogging	16730
Plone	CMS	13993
Compiere	ERP + CRM	13718
WebCalendar	Calendar	12832
Nucleus	Blogging	12739
Tikiwiki	Wiki	6368

Big Picture



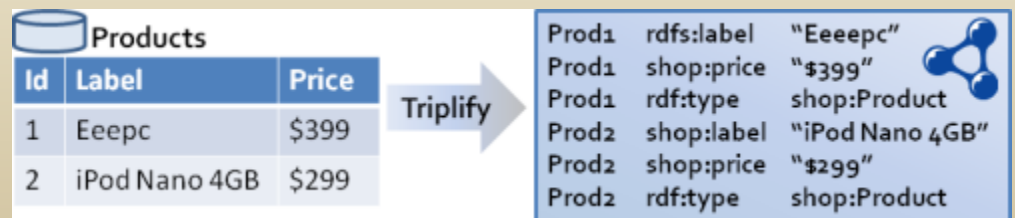
Triplify Approach

- Expose semantics as simple as possible
 - No (new) mapping languages
 - Few lines of code – easy to plug-in
 - Simple, reusable configurations
- Available for most popular Web app languages
 - PHP (ready), Ruby/Python under development
- Works with most popular Web app DBs
 - MySQL (extensively tested), PHP-PDO DBs (SQLite, Oracle, DB2, MS SQL, PostgreSQL etc.) should work, not needed for Virtuoso 😊
- Triplify exposes RDF/Ntriples, LinkedData and RDF/JSON

Triplify Solution: SQL-SELECT queries map relational data to RDF

Triplify Configuration basically consists of a definition of a number of SQL queries selecting information, which should be made publicly available. In order for Triplify to be able to convert the results of your SQL queries into RDF the query results are required to have a certain structure:

- the **first column must contain identifiers** which can be used to generate instance URIs (i.e. the primary key of your database table)
- **column names will be used to generate property URIs**, by renaming the columns of your database table (e.g. SELECT id,name AS '[foaf:name](#)' FROM users) you can reuse properties from existing vocabularies such as Dublin Core, FOAF, SIOC.
- the **individual cells of the query result contain data values or references to other instances** and will eventually constitute the objects of resulting triples



```
<?php
```

```
include('../wp-config.php');
```

```
$triplify['namespaces']=array(  
    'vocabulary'=>'http://triplify.org/vocabulary/Wordpress/',  
    'foaf'=>'http://xmlns.com/foaf/0.1/',  
    ... );
```

```
$triplify['queries']=array(  
    'post'=>array(  
        "SELECT id,post_author 'sioc:has_creator',post_date 'dcterms:created',post_title 'dc:title', post_content 'sioc:content',  
        post_modified 'dcterms:modified' FROM {$table_prefix}posts WHERE post_status='publish'",  
        "SELECT post_id id,tag_id 'tag:taggedWithTag' FROM {$table_prefix}post2tag",  
        "SELECT post_id id,category_id 'belongsToCategory' FROM {$table_prefix}post2cat",  
    ),  
    'tag'=>"SELECT tag_ID id,tag 'tag:tagName' FROM {$table_prefix}tags",  
    'category'=>"SELECT cat_ID id,cat_name 'skos:prefLabel',category_parent 'skos:narrower' FROM {$table_prefix}categories",  
    'user'=>array(  
        "SELECT id,user_login 'foaf:accountName',SHA(CONCAT('mailto:',user_email)) 'foaf:mbox_sha1sum',  
        user_url 'foaf:homepage',display_name 'foaf:name' FROM {$table_prefix}users",  
        "SELECT user_id id,meta_value 'foaf:firstName' FROM {$table_prefix}usermeta WHERE meta_key='first_name'",  
        "SELECT user_id id,meta_value 'foaf:family_name' FROM {$table_prefix}usermeta WHERE meta_key='last_name'",  
    ),  
    'comment'=>"SELECT comment_ID id,comment_post_id 'sioc:reply_of',comment_author AS 'foaf:name',  
        SHA(CONCAT('mailto:',comment_author_email)) 'foaf:mbox_sha1sum', comment_author_url 'foaf:homepage',  
        comment_date AS 'dcterms:created', comment_content 'sioc:content',comment_karma,comment_type  
        FROM {$table_prefix}comments WHERE comment_approved='1'",  
);
```

```
$triplify['objectProperties']=array(  
    'sioc:has_creator'=>'user', 'tag:taggedWithTag'=>'tag', 'belongsToCategory'=>'category','skos:narrower'=>'category','sioc:reply_of'=>'post');
```

```
$triplify['classMap']=array('user'=>'foaf:person', 'post'=>'sioc:Post', 'tag'=>'tag:Tag', 'category'=>'skos:Concept');
```

```
$triplify['TTL']=0; // Caching
```

```
$triplify['db']=new PDO('mysql:host='.DB_HOST.'.dbname='.DB_NAME,DB_USER,DB_PASSWORD);
```

```
?>
```

Example Config

Current state

- Configurations exist for
 - Joomla!, Wordpress, WackoWiki, OpenConf, OpenJournalSystems
- Aim is for Triplify to become integral part of Web app distributions (GSoC Joomla! project)
- Much work to convince/explain SemWeb and Triplify to Web app developer communities
- Triplification Challenge awards best “semantifications” (technologically not limited to Triplify)
- Very simple registry for Triplify endpoints
 - currently around 20 installations registered
 - Tip of the iceberg: registration optional and requires special PHP functionality (allow_url_fopen)

Future Work

- More configurations / make Triplify direct part of Web apps
- Showcase the benefits
 - Semantic Search
 - Semantic Mashups
 - Linked Data Browsing
- Vocabulary/views on updates
- Each Triplify installation exposes recent updates in a standardized way as a “LinkedData feed” or RSS/Atom feed
- (SPARQL support)