

Introducing “BlogSpread”

First Ideas

Based on discussions and work of:

Ig Ibert Bittencourt

Amadeu Campos

Kai Eckert

Fred Freitas

Mathias Niepert

Heiner Stuckenschmidt

What is the goal?

- A *{framework|platform|application}* to **analyze** and **visualize** *{blogs|forums|websites|the social web}*.
- Yet another framework...
- Interoperable with existing data (by means of semantic web technologies)
- ...

Motivation, Why?

- Do **we** need it? Does **someone else** need it?
- Maybe...
 - ... many smaller research projects: BA/MA theses, student research projects → **no reinvention of wheel.**
 - ... not only smaller research projects → **Reduce overhead.**
- Is it part of active research?

Mannheim Use-Cases

- Development of research software
 - Opinion mining, sentiment analysis on blogs
 - Natural language processing
 - Learning ontologies/thesauri from blogs

Architecture

- The core of the project will be a desktop application.
 - Targeted user is a single researcher who uses BlogSpread for experiments.
- The APIs can also be implemented as webservices, so that analyses can be provided as webservice.
- Agent technologies are used where applicable to improve scalability and encourage modularization.

Environment

- Rich Client Desktop Application
 - That means: Use a framework!!!
- Programming language: Java
 - Two serious RCP candidates:
 - Eclipse RCP
 - Netbeans Platform
- We use Netbeans Platform
 - Main reason: expertize and existing code
 - There is no better or worse, it is more a matter of taste.

The Netbeans IDE

The screenshot displays the NetBeans IDE 6.5.1 interface. The main window is titled "Database - NetBeans IDE 6.5.1". The menu bar includes File, Edit, View, Navigate, Source, Refactor, Run, Debug, Profile, Versioning, Tools, Window, and Help. The toolbar contains icons for file operations, navigation, and execution. The left sidebar shows a project tree with a "Database" package containing several Java files, including "Concept.java". The central editor displays the following Java code:

```
public interface Concept {  
  
    List<Label> getPrefLabels ();  
    Label getPrefLabel (Language language);  
  
    List<Label> getAltLabels ();  
    List<Label> getAltLabels (Language language);  
  
    List<Concept> getBroader ();  
  
    ConceptScheme getConceptScheme ();  
  
    List<Label> getHiddenLabels ();  
    List<Label> getHiddenLabels (Language language);  
  
    Long getId ();  
  
    List<Label> getLabels ();  
    List<Label> getLabels (Language language);  
  
    List<Concept> getNarrower ();  
}
```

The right sidebar shows the "Concept.java - Properties" panel with the following table:

| Concept.java - Properties | |
|---------------------------|----------------------------|
| Properties | |
| Name | Concept |
| All Files | X:\daten\svn\core\semt... |
| File Size | 2011 |
| Modification Time | 28.10.2009 14:51:52 |
| Classpaths | |
| Compile Classpath | X:\daten\svn\core\semt... |
| Runtime Classpath | X:\daten\svn\core\semt... |
| Boot Classpath | C:\Program Files\Java\j... |

The bottom of the IDE features a status bar with "1:1 INS" and a row of panels: Usages, Analyzer, Output, Subversion - Property..., and Tasks.

Semtnel: Based on Netbeans Platform

The screenshot displays the Semtnel Platform interface, which is based on the Netbeans platform. The interface is divided into several panes:

- Tree View:** Located on the left, it shows a hierarchical structure of concept schemes. The "Immune System" is highlighted, and a blue arrow points from this pane to the "Thesauri" label.
- Explanation Graph Viewer:** Below the Tree View, it shows a list of concept schemes and record sources. A blue arrow points from this pane to the "Indexed Documents" label.
- Treemap Visualization:** The central pane displays a treemap visualization of the "MeSH (Tue Mar 31 09:40:32 CEST 2009)" concept scheme. A red arrow points from this pane to the "Analyses" label.
- Publication Search:** A red arrow points from this pane to the "Visualisations" label.
- Palettes:** On the right, there are palettes for "Analyses" and "Dataprovider". A red arrow points from the "Analyses" palette to the "Analyses" label.
- Immune System - Proper...:** A detailed view of the "Immune System" concept scheme is shown on the right. It includes a table of labels and a "Usage count" table. A red arrow points from this pane to the "Analyses" label.

Red arrows also point from the "Analyses" label to the "Tree View" and "Publication Search" panes, and from the "Visualisations" label to the "Treemap Visualization" pane.

Thesauri

Indexed Documents

Analyses

Visualisations

Preferred: Immune System
Hidden: Immune Systems, System, Immune, Systems, Immune

Scope Note: The body's defense mechanism against foreign organisms or substances and deviant native cells. It includes the humoral immune response and the cell-mediated response and consists of a complex of interrelated cellular, molecular, and genetic components.

| Label | Value |
|---------------------------|-------|
| Pubmed 1986 Pubmed Manual | 2 |
| Pubmed 2006 Pubmed Manual | 14 |
| Pubmed 1996 Pubmed Manual | 5 |
| Pubmed 1991 Pubmed Manual | 6 |
| Pubmed 2001 Pubmed Manual | 17 |

Immune System

Talking about analyses

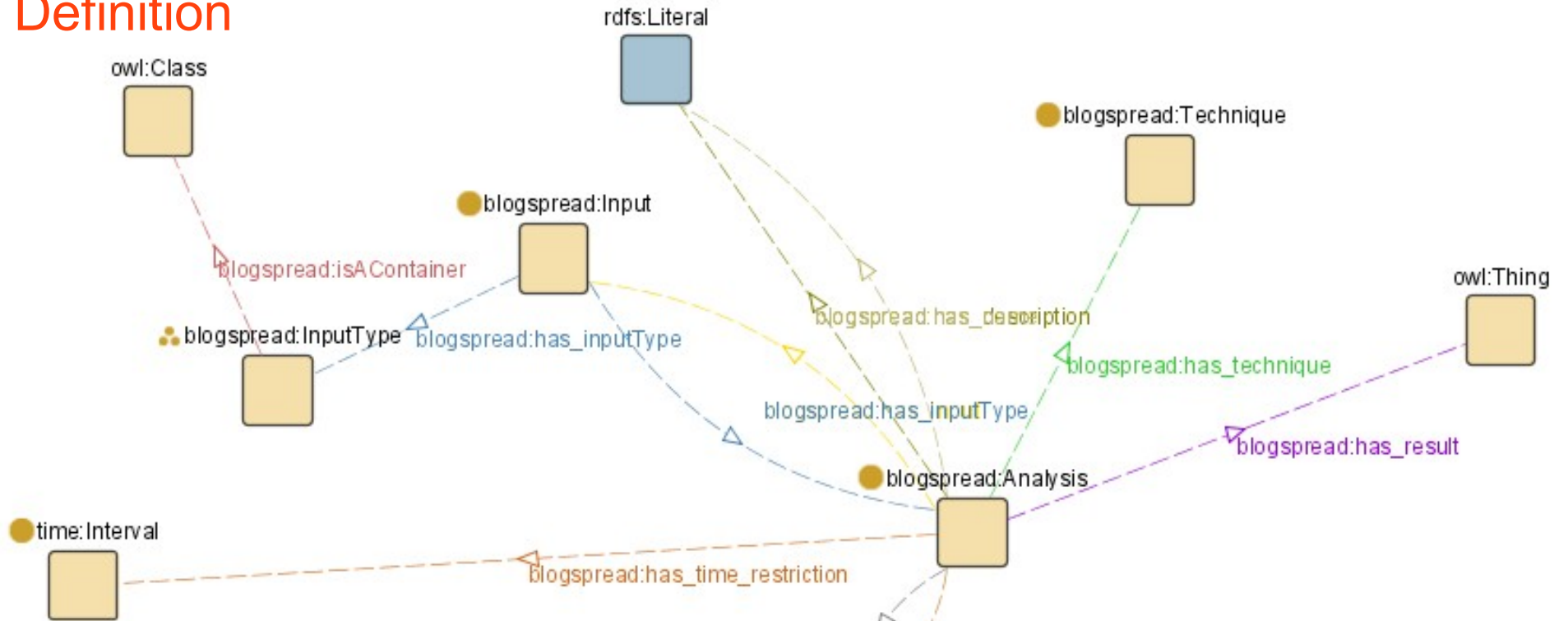
- Analyses have...
 - ... different numbers of inputs
 - ... different types of input
 - ... different numbers of results
 - ... different types of results
- An analysis can **built upon other analyses**
- An analysis is not necessarily pure math, but a **black box** (maybe it uses arbitrary background-knowledge internally)

Talking about visualizations

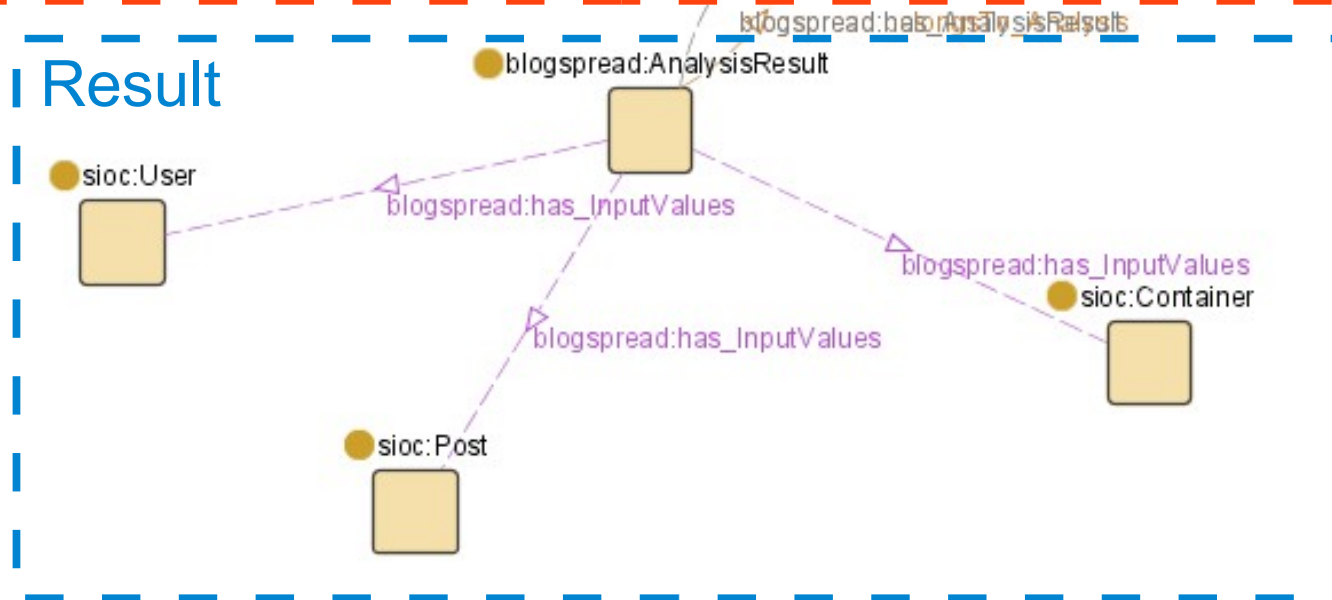
- Visualizations work for different kinds of data.
- Visualizations can combine several analysis results.
- Example: Treemap
 - Needs **hierarchical data** (so this property should be part of the data description)
 - Can simultaneously visualize two results by means of **color** and **size**, both usually are provided by a metric.
 - The color metric usually needs a **limited range**.

The Analysis Ontology: A first draft

Definition



Result



Existing projects

- University of Mannheim
 - Sementinel (Kai)
 - Blogalyzer (Kai)
- University of Pernambuco, Recife
 - Master-Web, Agathe (Fred)
 - BWI + Pos Tagging (Rinaldo)
 - GATE Pipelines (Luciano)
 - Visualisation Tool (Amadeu)
 - Tag Suggestion (Icaro)

BlogSpread on Sourceforge

- <http://blogspread.sourceforge.net/>
- We will use it (at least) for:
 - Wiki
 - Mailinglists
 - Code-Repository
 - Downloads
- Everyone in this room can join this project!
- **Spread the word!**

SourceForge.net > Find Software > Blogsread

Blogsread Planning by kaiec

[Share](#) [More](#)

[Summary](#) | [Files](#) | [Support](#) | [Develop](#)

Blogsread is an open platform for the development of applications that analyze data from websites like blogs and forums. Blogsread is jointly developed by the University of Mannheim, Germany, and the Universidade Federal de Alagoas, Maceio, Brasil.

[View all files](#)

<http://blogsread.sourceforge.net>

TAGS

[EDIT](#)

[Show project details](#)

Rate and Review

Be the first person to add a text review.

Would you recommend this project?



or



Ratings and Reviews



Be the first to post a text review of Blogsread. Rate and review a project by clicking thumbs up or thumbs down in the right column.

Project Feed

Show:



 [Code committed](#)

Related Projects

SQUID2

SQUID2 software processes SHRIMP ion microprobe data. It derives ages for the U-Th-Pb system, ...

micro-read Fast Alignment Search Tool

Fast Alignment Search Tool suite. A fast sequence read mapper suite for short reads generated ...

[NBM Center](#)

[Blogspread](#): Project Web Hosting - Open Source Software


Blogspread

Users

[Download Blogspread files](#)

[Project detail and discuss](#)

[Get support](#)



About this project:

This is the **Blogspread** project ("blogspread")

Blogspread is an open platform for the development of applications that analyze data from websites like blogs and forums. Blogspread is jointly developed by the University of Mannheim, Germany, and the Universidade Federal de Alagoas, Maceio, Brasil.

Developers

Join this project:

To join this project, please contact the project administrators of this project, as shown on the [project summary page](#).

Get the source code:

Source code for this project may be available as [downloads](#) or through the CVS or Subversion SCM repository used by the project, as accessible from the [project summary page](#).

Project Web Hosted by SOURCEFORGE.NET*

©Copyright 1999-2008 - [SourceForge](#), Inc., All Rights Reserved

[About](#) - [Legal](#) - [Help](#)



Welcome, Guest! [Log In](#) | [Create Account](#)

[Visit project blogspread](#)



navigation

- [Main Page](#)
- [Community portal](#)
- [Current events](#)
- [Recent changes](#)
- [Random page](#)
- [Help](#)

search

toolbox

- [What links here](#)
- [Related changes](#)
- [Special pages](#)
- [Printable version](#)
- [Permanent link](#)

[page](#) [discussion](#) [view source](#) [history](#)

Existing Projects

University of Mannheim

- [Semtinel \(Kai\)](#)
- [Blogalyzer \(Kai \(not my project, but I have access to it\)\)](#)

University of Pernambuco, Recife

- [Master-Web, Agathe \(Fred\)](#)
- [BWI + Pos Tagging \(Rinaldo\)](#)
- [GATE Pipelines \(Luciano\)](#)
- [Visualisation Tool \(Amadeu\)](#)
- [Tag Suggestion \(Icaro\)](#)

University of Maceio

This page was last modified on 11 November 2009, at 14:28. This page has been accessed 2 times.

