



salzburgresearch



REDLINK

Dr. Sebastian Schaffert

Von Stanbol nach Marmotta

Content Enhancement und Smarte Annotationen für die Medienindustrie



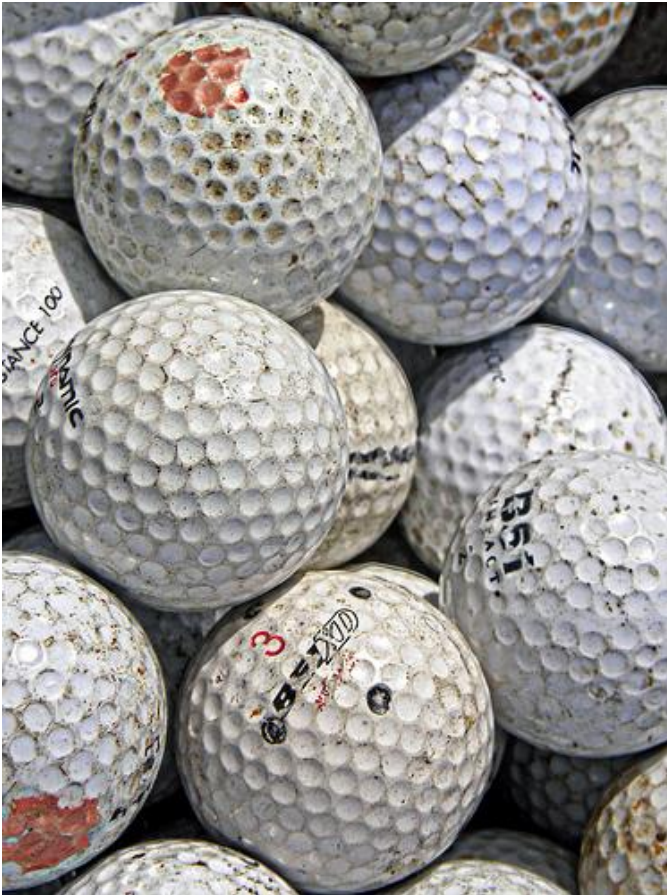
“When you can measure what you are speaking about and express it in numbers, you know something about it, but when you cannot measure it, your knowledge is of a meagre and unsatisfactory kind”

- Lord Kelvin





(Medien-)Inhalte sind wertvoll!



- Art: Golfball
- Farbe: weiß
- Kategorie: Sport - Golf
- Zustand: gebraucht
- Durchmesser: 43mm
- Preis: \$0,50

Quelle: School of Data, <http://schoolofdata.org/handbook/courses/what-is-data/>



(Medien-)Inhalte sind wertvoll!



- Art: Golfball
- Farbe: weiß

leider sind die wertvollen
Informationen vor Maschinen
versteckt!

olf

n

Quelle: School of Data, <http://schoolofdata.org/handbook/courses/what-is-data/>



Versteckte Werte ...

So sieht eine Webseite in natürlicher Sprache für eine Maschine aus ...

林克昌 根留台灣 可能增高

在愛戴者熱心奔走之下，華裔名指揮家林克昌根留台灣的可行性又提升了幾分。兩廳院主任李炎、國家音樂廳樂團副團長黃奕明日前親赴林克昌、石聖芳寓所拜會，並提出多場客席邀約。此外，台灣省立交響樂團團長陳澄雄也早早「下訂」，邀請林克昌赴台中霧峰，從八月十日起訓練省交，為期長達一個月。

在台灣諸多公家樂團中，陳澄雄是以實際行動表達對林克昌肯定的樂界人士之一，曾多次公開表示對林克昌指揮才華的欽佩，而且幾乎每個樂季都邀請林克昌客席演出。

此外，林克昌上個月赴俄羅斯與頂尖的「俄羅斯國家管絃樂團」灌錄了柴可夫斯基晚期三大交響曲以及「羅密歐與茱麗葉」、「斯拉夫進行曲」、「義大利隨想曲」，最後的DAT母帶也在前兩天寄回台灣。製作人楊忠衡與林克昌試聽之後，都對錄音效果—尤其音質表現感到相當滿意，楊忠衡估計呈現了七分林克昌指揮神韻。

俄羅斯國家管絃樂團首席布魯尼日前也讚譽林克昌的指揮藝術有三大特點：一是控制自如的彈性速度；二是強烈的動態對比；三是宛如呼吸歌唱的旋律處理。這些對錄音師而言都構成很大挑戰。俄國錄音師雖然採用多軌混音，但定位、場面都有可觀之處。。

Folie von James Hendler (Univ. Maryland)

Versteckte Werte ...



Mit XML (o.ä.) versuchen wir, dem Text durch Tags „Bedeutung“ hinzuzufügen

林克昌 根留台灣 可能增高

<name>

<education>

在愛戴者熱心奔走之下，華裔名指揮家林克昌根留台灣的可行性又提升了幾分。兩廳院主任李炎、國家音樂廳樂團副團長黃奕明日前親赴林克昌、石聖芳寓所拜會，並提出多場客席邀約。此外，台灣省立交響樂團團長陳澄雄也早早「下訂」，邀請林克昌赴台中霧峰，從八月十日起訓練省交，為期長達一個月。

在台灣諸多公家樂團中，陳澄雄是以實際行動表達對林克昌肯定的樂界人士之一，曾多次公開表示對林克昌指揮才華的欽佩，而且幾乎每個樂季都邀請林克昌客席演出。

<work>

此外，林克昌上個月赴俄羅斯與頂尖的「俄羅斯國家管絃樂團」灌錄了柴可夫斯基晚期三大交響曲以及「羅密歐與茱麗葉」、「斯拉夫進行曲」、「義大利隨想曲」，最後的DAT母帶也在前兩天寄回台灣。製作人楊忠衡與林克昌試聽之後，都對錄音效果—尤其音質表現感到相當滿意，楊忠衡估計呈現了七分林克昌指揮神韻。

<CV>

<private>

俄羅斯國家管絃樂團首席布魯尼日前也讚譽林克昌的指揮藝術有三大特點：一是控制自如的彈性速度；二是強烈的動態對比；三是宛如呼吸歌唱的旋律處理。這些對錄音師而言都構成很大挑戰。俄國錄音師雖然採用多軌混音，但定位、場面都有可觀之處。。

Folie von James Hendler (Univ. Maryland)

Versteckte Werte ...



Aber für die Maschine sehen die Tags in etwa so aus:

林克昌 根留台灣 可能增高

< ναμε >

< εδουχατιον >

在愛戴者熱心奔走之下，華裔名指揮家林克昌根留台灣的可行性又提升了幾分。兩廳院主任李炎、國家音樂廳樂團副團長黃奕明日前親赴林克昌、石聖芳寓所拜會，並提出多場客席邀約。此外，台灣省立交響樂團團長陳澄雄也早早「下訂」，邀請林克昌赴台中霧峰，從八月十日起訓練省交，為期長達一個月。

在台灣諸多公家樂團中，陳澄雄是以實際行動表達對林克昌肯定的樂界人士之一，曾多次公開表示對林克昌指揮才華的欽佩，而且幾乎每個樂季都邀請林克昌客席演出。

< ωορκ >

此外，林克昌上個月赴俄羅斯與頂尖的「俄羅斯國家管絃樂團」灌錄了柴可夫斯基晚期三大交響曲以及「羅密歐與茱麗葉」、「斯拉夫進行曲」、「義大利隨想曲」，最後的DAT母帶也在前兩天寄回台灣。製作人楊忠衡與林克昌試聽之後，都對錄音效果—尤其音質表現感到相當滿意，楊忠衡估計呈現了七分林克昌指揮神韻。

< Xς >

< πριωατε >

俄羅斯國家管絃樂團首席布魯尼日前也讚譽林克昌的指揮藝術有三大特點：一是控制自如的彈性速度；二是強烈的動態對比；三是宛如呼吸歌唱的旋律處理。這些對錄音師而言都構成很大挑戰。俄國錄音師雖然採用多軌混音，但定位、場面都有可觀之處。

Folie von James Hendler (Univ. Maryland)

Strukturierte Informationen



Structured content refers to information or content that has been broken down and classified using metadata.

Wikipedia

Broken down

into discrete concepts

Classified

as real-world things and relationships

Metadata

a structure readable by robots and people

Folie von Mike Atherton, BBC



Warum?

- manage content at scale
- meaningful navigation
- expose long-tail content
- reuse content assets
- bridge across subjects
- improve findability
- support social sharing
- improve SEO
- design for all devices
- robot friendly



Die 7 Informationsgesetze

1. Information is (Infinitely) Shareable
2. The Value of Information Increases with Use
3. Information is Perishable
4. The Value of Information Increases with Accuracy
5. The Value of Information Increases in Combination
6. More Is Not Necessarily Better
7. Information is Self-propagating

Daniel Moody, Peter Walsh: *Measuring the Value of Information: An Asset Valuation Approach*. In: European Conference on Information Systems (ECIS'99), Copenhagen Business School, Frederiksberg, Denmark, June 1999.



Leider sind die meisten Informationen unstrukturiert, weil das Strukturieren mit sehr viel (Hand-)Arbeit verbunden ist ...





Demo:

Ein Prozess, wie man basierend auf Open Source Software
Bedeutung aus unstrukturierten Inhalten gewinnen kann!



Beispielszenario: IKS Blog



Blog Homebase Wiki About

THE SEMANTIC CMS COMMUNITY **IKS**

Join the Semantic CMS Community. Now!

January 28, 2013
by Adrian
Comments Off

Apache Stanbol Early Adopter using Ontology Manager and Rules components

Apache Stanbol provides different features that facilitate working with linked data, in the [netlabs.org early adopter proposal](#) I wanted to try features which were not much used before, like the [Ontology Manager](#) and the [Rules](#) component. The way I wanted to use them were not necessary the ways the Stanbol team was thinking of it in the first place but nevertheless I saw that it might solve some problems I have. [Continue Reading](#) →

[Tweet](#) [Like](#) 0 [Share / Save](#)

Categories: [IKS Early Adopers](#), [Ontonet](#), [Refactor Engine](#) | Tags: [Apache Stanbol](#), [early_adopters](#), [Ontology Manager](#), [Stanbol Rules](#) | [Permalink](#)

FOLLOW US

[RSS](#) [Facebook](#) [Google+](#) [Twitter](#) [YouTube](#)

CATEGORIES

Select Category ▾

ARCHIVES

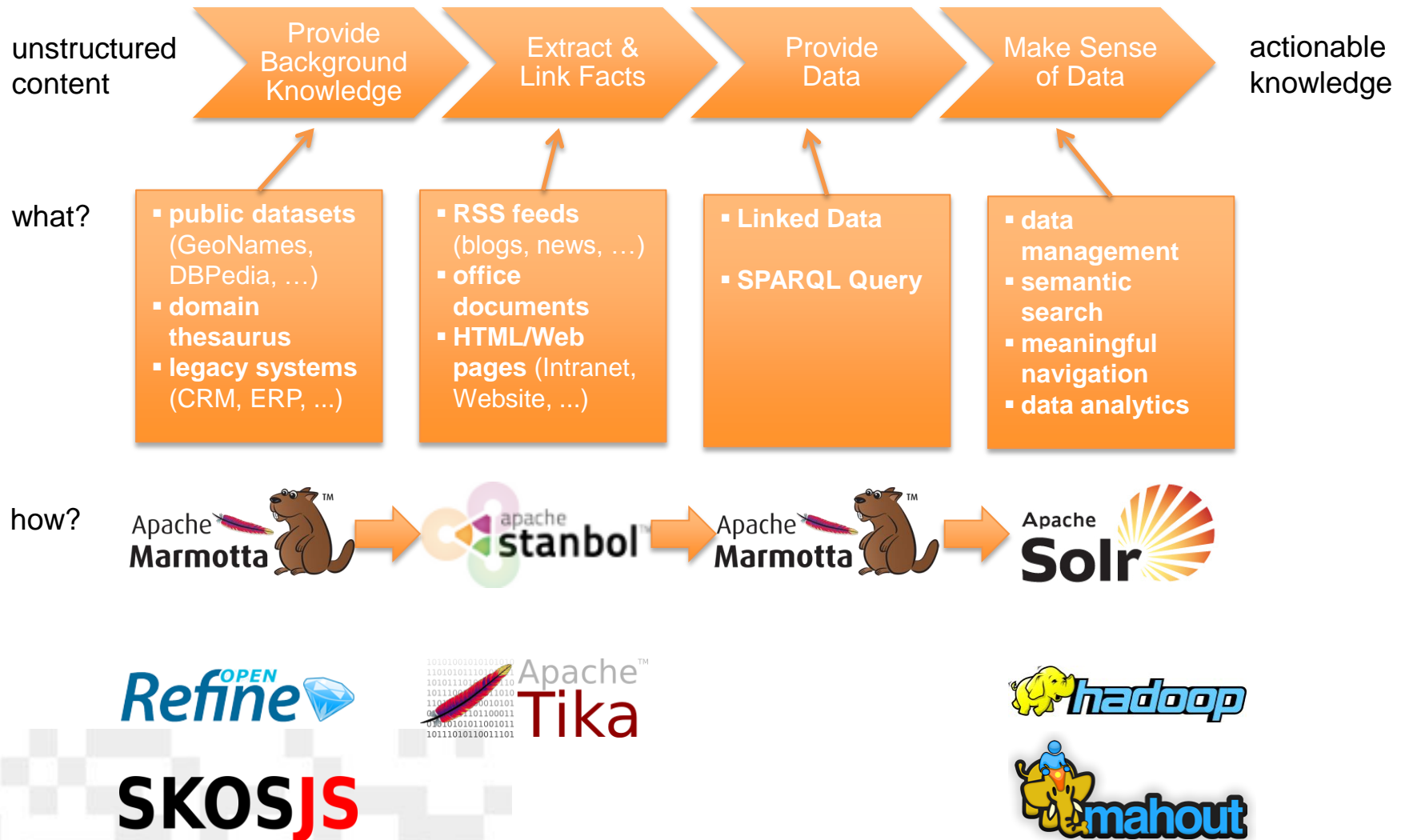
Select Month ▾

RECENT POSTS

- Apache Stanbol Early Adopter using Ontology Manager and Rules



Prozess: von Marmotta nach Stanbol und zurück





Hintergrundwissen bereitstellen

- **Ziel:** Domänenwissen maschinenlesbar bereitstellen, um Informationen darauf basierend zu analysieren und zu verknüpfen
- **Beispiele:**
 - Domänenthesaurus (z.B. Produkte, Abteilungen, Themen, ...)
 - Öffentliche Datensätze (z.B. Geodaten, Enzyklopädie, andere Datenbanken)
 - Andere Datenquellen (ERP, CRM, Telefonverzeichnis, ...)



Hintergrundwissen bereitstellen: Werkzeuge



- Thesaurus-Manager: SKOSJs

The screenshot displays the SKOSJs web interface. On the left, a hierarchical tree structure shows the following categories: **Persons**, **Salzburg Research** (containing **Rupert Westenthaler** and **John Pereira**), **Other** (containing **Henri Bergius**), **Uni Paderborn** (containing **Fabian Christ**), and **Free Concepts**. The main area on the right shows the detailed view for the concept **Salzburg Research**. This view includes several sections: **Broader Concepts** (with a search box containing 'Salzburg Research'), **Narrower Concepts**, **Related Concepts**, **Broad Match**, **Narrow Match**, **Close Match**, and **Exact Match**. On the far right, there are fields for **Preferred Label** (en: John Pereira, de:), **Alternative Label** (en: jpereira, de:), **Hidden Label** (en: , de:), **Author** (N/A), **Created** (2012-10-10T09:29:24), **Modified** (2012-10-10T09:29:24 and 2012-10-10T09:29:30), **Contributors** (N/A), and **Definition** (en:).



Hintergrundwissen bereitstellen: Werkzeuge

- Datenaufbereitung: OpenRefine





Fakten extrahieren und verknüpfen

- **Ziel:** Fakten in unstrukturierten Inhalten erkennen und mit Hintergrundwissen verknüpfen
- **Beispiele:**
 - Orte im Text mit Geoinformationen (z.B. GeoNames)
 - Produkte im Text mit Firmenthesaurus oder Produktdatenbank
 - Personen im Text mit CRM-System oder Addressbuch



Fakten extrahieren und verknüpfen



apache **stanbol**™ The RESTful Semantic Engine [/enhancer](#) [/topic](#) [/entityhub](#)

Apache Stanbol Enhancer [Web View](#) [REST API](#)

Enhancement Chain: **Projects** all 7 engines available [< List of Enhancement Chains >](#)

Paste some text below and submit the form to let the Enhancement Chain Projects enhance it:

CreateJS **hackathon** in Berlin
July 26, 2012 by Henri
Last week we at IKS organized a two-day hackathon for developers interested in Create.js, CreatePHP, VIE, and in new tools for editing websites semantically. The original agenda for the event can be found from the event's wiki page, and I believe we got it covered quite well. I've been to quite a lot of hackfests, and usually the first day is mostly spent by people arriving, figuring out the WiFi setup, and getting to know the development tools. This time was different: people jumped straight into the code and started working.

The venue at Theseus Innovationszentrum, arranged by DFKI, was a nice one. We had food, coffee, and fast connectivity. And the place even had toys like big touchscreen computers to try Create.js on. On the first evening we continued the event at the c-base hackerspace, and the participants staying after Friday went to the Berlin GitHub drinkup.

More robust Hallo and Create

An important theme in the event was to improve both Hallo and Create. The Create.js theme got a lot nicer, and thanks to the contributions from Alkacon the next versions will run nicely also on older Internet Explorer and Opera versions.

Create's editor selection mechanism was also rewritten. Now you can easily set un custom editor conifaurations for manaaina different content tvnes. For instance:





Output format:

Enhancement Process Metadata

Execution of Chain **Projects** completed in **780ms**.

Extracted entities

Concepts

-  [CreateJS](#)
-  [CreatePHP](#)
-  [Django](#)
-  [Rails](#)



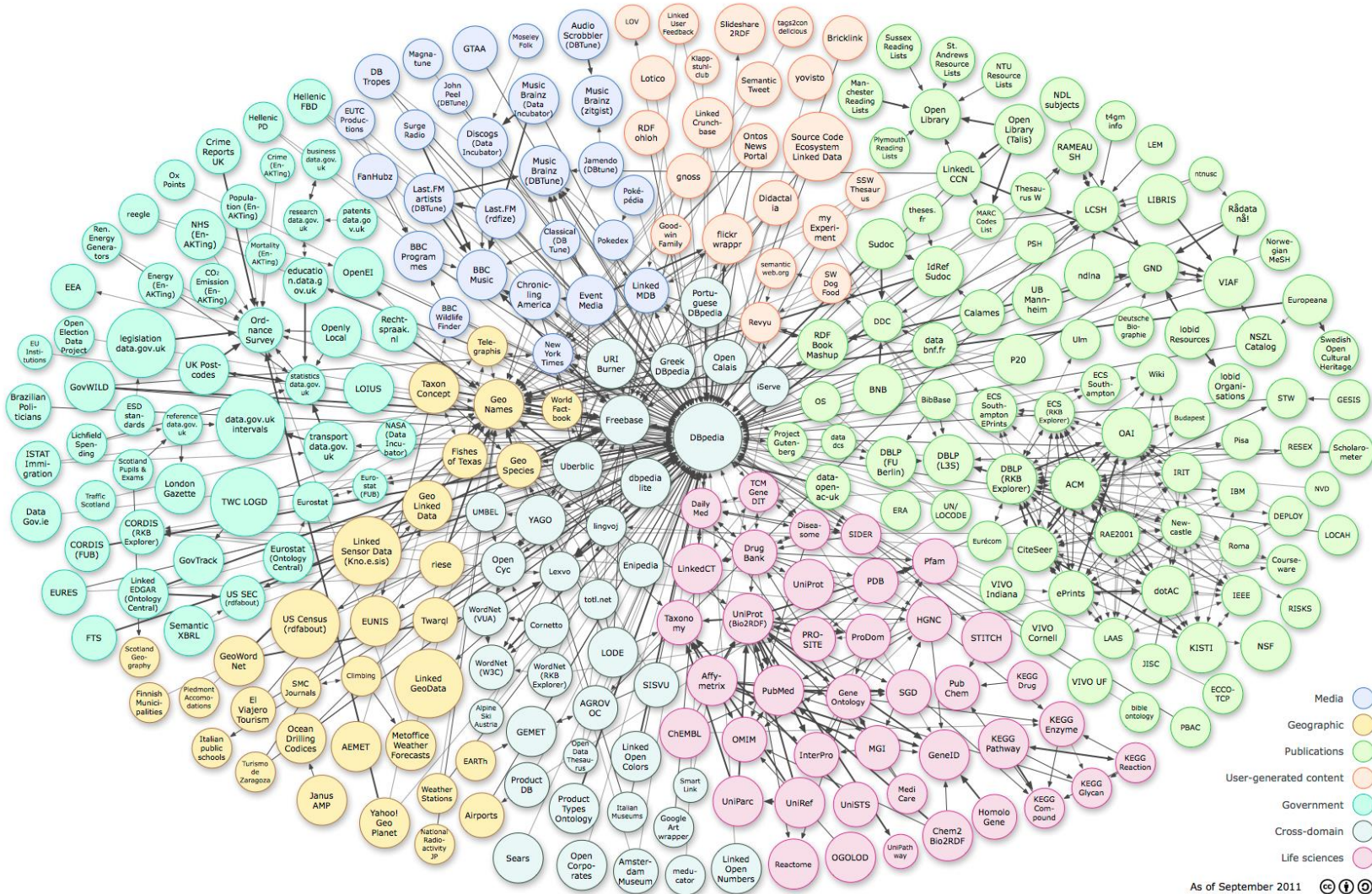
Daten bereitstellen

- **Ziel:** Daten in einheitlichem Datenmodell anderen Anwendungen zur Weiterverarbeitung zur Verfügung stellen

- **Technologien:**
 - Linked Data / RDF
 - SPARQL Anfragesprache




Daten bereitstellen: Linked Data



Daten bereitstellen: SPARQL





SPARQL - snorql login

SPARQL Query Form

Querying <http://localhost:8080/sparql/select>

```
PREFIX mao: <http://www.w3.org/ns/ma-ont#>
PREFIX yt: <http://gdata.youtube.com/schemas/2007#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX ldp: <http://www.w3.org/ns/ldp#>
PREFIX lmftypes: <http://www.newmedialab.at/lmf/types/1.0/>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX sioc: <http://rdfs.org/sioc/ns#>
PREFIX dcat: <http://www.w3.org/ns/dcat#>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX context: <http://localhost:8080/context/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX local: <http://localhost:8080/resource/>
```

17 SELECT DISTINCT ?property ?hasValue ?isValueOf
18 WHERE {
19 { <urn:enhancement-0043807f-6e67-c1d8-5ac4-3fc1dfbf551f> ?property ?hasValue }
20 UNION
21 { ?isValueOf ?property <urn:enhancement-0043807f-6e67-c1d8-5ac4-3fc1dfbf551f> }
22 }
23 ORDER BY (!BOUND(?hasValue)) ?property ?hasValue ?isValueOf

Results:

Description of urn:enhancement-0043807f-6e67-c1d8-5ac4-3fc1dfbf551f:

property	hasValue
<http://fise.iks-project.eu/ontology/confidence>	"0.6400000190734865"
<http://fise.iks-project.eu/ontology/entity-label>	"FISE"@en
<http://fise.iks-project.eu/ontology/entity-reference>	<http://localhost:8080/LMF/resource/Egl8XBPo>
<http://fise.iks-project.eu/ontology/entity-type>	skos:Concept
<http://fise.iks-project.eu/ontology/extracted-from>	<http://localhost:8080/enhancement/sha1-011ffe9b959978141d66e72d8645b6b18b154bd>
dct:created	"2013-04-16T12:04:26.000Z"
dct:creator	"org.apache.stanbol.enhancer.engines.entitylinking.engine.EntityLinkingEngine"
dct:relation	<urn:enhancement-fc3fe90f-741b-8e89-6e9e-02fec751879>
<http://stanbol.apache.org/ontology/entityhub/entityhub#site>	"lmfSiteCb6dea305eaa5fa26fcb67b808697b6"
rd:type	<http://fise.iks-project.eu/ontology/Enhancement>
rd:type	<http://fise.iks-project.eu/ontology/EntityAnnotation>



Nutzen stiften

- **Ziel:** aus den Daten neues Wissen gewinnen und daraus Handlungen ableiten.

- **Beispiel:**
 - Semantische Suche
 - Datenanalyse
 - Wiederverwendung von Inhalten



Nutzen stiften: Semantische Suche



salzburgresearch

Tools

<input type="checkbox"/> Wordpress	6
<input type="checkbox"/> Drupal	5
<input type="checkbox"/> Alfresco	4
<input type="checkbox"/> OpenCMS	4
<input type="checkbox"/> Firefox	2
<input type="checkbox"/> Google Chrome	2
<input type="checkbox"/> Internet Explorer	2
<input type="checkbox"/> Nuxeo	2

Projects

<input type="checkbox"/> Apache Stanbol	55
<input type="checkbox"/> VIE	26
<input type="checkbox"/> CreateJS	5
<input type="checkbox"/> CreatePHP	2

Locations

no facets

Persons

<input type="checkbox"/> Henri Bergius	6
<input type="checkbox"/> Salzburg Research	3
<input type="checkbox"/> Rupert Westenthaler	2
<input type="checkbox"/> Fabian Christ	1
<input type="checkbox"/> John Pereira	1

Active Filters

Currently no active filters

Shown results: 1 to 15 of 89

results/page sort by

IKS Technology Stack – Introduction and Overview Video

by Fabian

The IKS community workshop in Salzburg was a great event. As a member of the IKS project it was very inspiring to see all those CMS people who are really interested in the IKS technology. I got the honor to ... [Continue reading](#) →

created: 07/05/12

The Enterprise Search Market – What should be on your radar? Keynote IKS 2012 Workshop

by asharpe

I thoroughly enjoyed speaking and interacting with the attendees at this years IKS workshop. This sector does not get that much attention, and has fallen a long way from the spotlight that it had on it in the dot com ... [Continue reading](#) →

created: 07/04/12

Zaizi present Alfresco Semantic Solution at IKS Semantic Enterprise Technology Workshop

by rharo

Zaizi recently participated in the 7th IKS workshop that took place in Salzburg, Austria on 12-13 June. We joined the IKS community to share our experience in the use of Semantic Technologies within the enterprise and how IKS technology stack ... [Continue reading](#) →

created: 07/09/12


Apache Stanbol Stress Test Tool

by Rupert Westenthaler

The Apache Stanbol Enhancer provides a RESTful API for extracting information from parsed content

Nutzen stiften: Datenanalyse





Core Services - dataview login

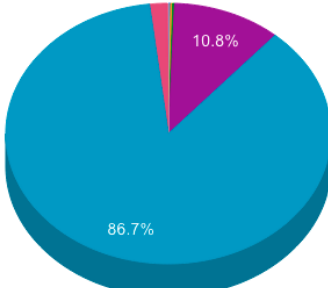
Dataview

Overview

System contains 2859 subjects with overall 28754 triples in 7 graphs

Graphs

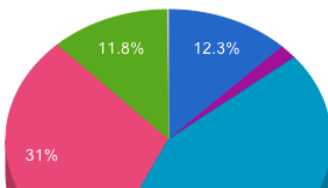
All Graphs in the system. Click to select!



Graph URL	Percentage
http://localhost:8080/context/tools	86.7%
http://localhost:8080/context/projects	10.8%
http://localhost:8080/context/persons	
http://localhost:8080/context/default	
http://localhost:8080/enhancer/iks	
http://localhost:8080/context/inferred	
Other	

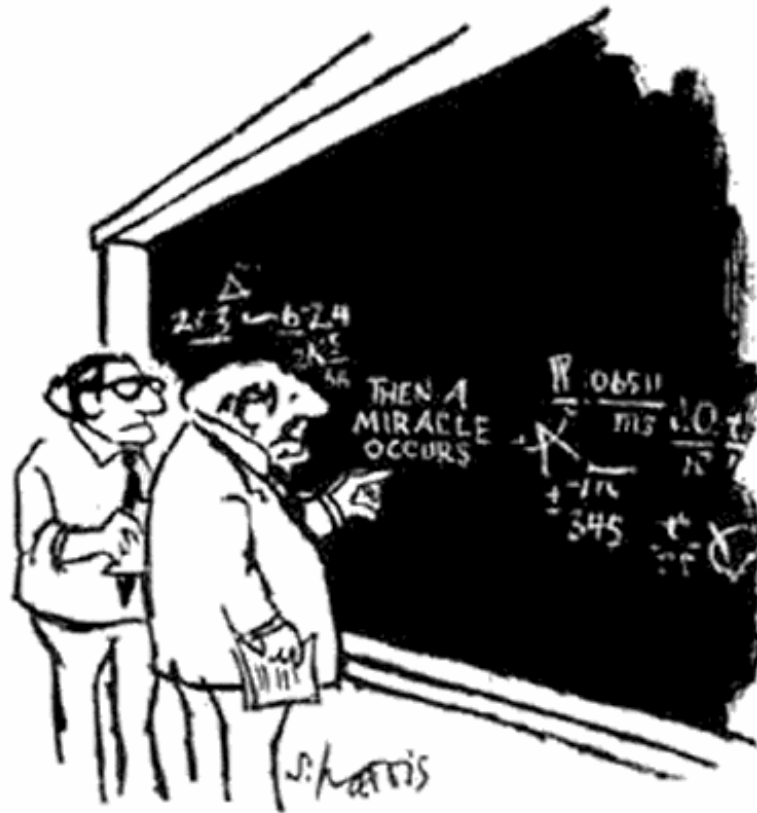
Types

Classes in all graphs



Class URL	Percentage
http://www.w3.org/2004/02/skos/core#Co...	12.3%
http://rdfs.org/sioc/ns#Post	
http://fise.iks-project.eu/ontology/Enhance...	11.8%
http://fise.iks-project.eu/ontology/TextAnn...	31%
http://fise.iks-project.eu/ontology/EntityAn...	
Other	

Vielen Dank!



"I THINK YOU SHOULD BE MORE EXPLICIT
HERE IN STEP TWO."

© 1988, 1979-81 (1982) -

Distributed By Oxford-Engineering Ltd

Kontakt



- Forschungsfragen:
sebastian.schaffert@salzburgresearch.at

salzburgresearch

- Umsetzung, Wartung, Support:
sebastian.schaffert@redlink.co

REDLiNK





Thanks for your Attention!

salzburg**research**

Dr. Sebastian Schaffert

Head of Knowledge and Media Technologies
Salzburg Research Forschungsgesellschaft m.b.H.
Jakob-Haringer-Straße 5/III | Salzburg, Austria
Tel. +43 662 2288-423| Fax +43 662 2288-222
sebastian.schaffert@salzburgresearch.at