Time-Aware Entity Search in DBpedia

Search for the entity name "Irving" with time range in Feb. 2014 and the intended entity is **Kyrie Irving**, who won the NBA All-Star Game MVP Award on 2014-02-17.

The Total Count of **Irving, Texas** during 20140101 to 20141231 is 258

The Total Count of **Kyrie Irving** during 20140101 to 20141231 is 905711

<table>
<thead>
<tr>
<th>Resource</th>
<th>P(rl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irving, Texas</td>
<td>0.6338983050847458</td>
</tr>
<tr>
<td>Irving (band)</td>
<td>0.05084745762711865</td>
</tr>
<tr>
<td>Henry Irving</td>
<td>0.04745762711864407</td>
</tr>
<tr>
<td>Irving, New York</td>
<td>0.04406779661016949</td>
</tr>
<tr>
<td>Irving</td>
<td>0.023728813559322035</td>
</tr>
<tr>
<td>Irving Oil</td>
<td>0.023728813559322035</td>
</tr>
<tr>
<td><strong>Kyric Irving</strong></td>
<td><strong>0.023728813559322035</strong></td>
</tr>
<tr>
<td>Irving (Duluth)</td>
<td>0.020338983050847456</td>
</tr>
<tr>
<td>Washington Irving</td>
<td>0.013559322033898305</td>
</tr>
<tr>
<td>Irving Jacobson</td>
<td>0.010169491525423728</td>
</tr>
<tr>
<td>John Irving</td>
<td>0.006779661016949152</td>
</tr>
<tr>
<td>Irving, Illinois</td>
<td>0.006779661016949152</td>
</tr>
<tr>
<td>Irving, Wisconsin</td>
<td>0.006779661016949152</td>
</tr>
</tbody>
</table>
With just 2% true positives provided as labeled data, how can we get bang for buck?
SemNaaS
Add Semantic Dimension to the Network as a Service
Discovering Types in RDF Datasets

Our Goal
Providing the types describing a dataset when this information is incomplete or missing.

Our Approach
- Type discovery using density-based clustering
  - Groups entities according to the neighbors density
  - Robust to noise and deterministic, detects classes of arbitrary shape
- Type profiles generation
  - Properties with a level of confidence
- Overlapping types generation
  - Profile analysis
FOODpedia.tk

Food products and their ingredients as Linked Data dataset

Goals:
- multilingual knowledge graph (RU, EN)
- interlinked with LOD datasets (AGROVOC, DBpedia)
- crowdsourcing platform

Acknowledgements: We gratefully acknowledge funding from the US National Science Foundation under grant IIS-1344272.

Yolanda Gil, Felix Michel, Varun Ratnakar
Information Sciences Institute
University of Southern California

Matheus Hauder
Software Engineering for Business Information Systems
Technical University Munich

---

<table>
<thead>
<tr>
<th>Tool Comparison</th>
<th>Category</th>
<th>Tool</th>
<th>Community support</th>
<th>Mainly task oriented</th>
<th>Semantic organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forum</td>
<td></td>
<td>V-Bulletin Forum</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WordPress bbPress</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WordPress p2 Theme</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stackoverflow</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mathematics</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Wiki</td>
<td></td>
<td>MediaWiki</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Semantic MediaWiki</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Enterprise</td>
<td></td>
<td>Confluence</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Knowledge</td>
<td></td>
<td>Connections</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Management</td>
<td></td>
<td>Jive</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MS SharePoint</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Communote</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yammer</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ToDo Lists</td>
<td></td>
<td>Trello</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Keep</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Todoist</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Repositories</td>
<td></td>
<td>GitHub</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BitBucket</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Google Code</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Organic Data Science Wiki</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

---

1. **Task Representation**: Set up the GLM lake model for North Temperate Lakes
   - Collect downwelling longwave radiation data for North Temperate lakes
   - Collect air temperature data for North Temperate lakes
   - Collect relative humidity data for North Temperate lakes
   - Collect wind speed data for North Temperate lakes
   - Collect rainfall data for North Temperate lakes
   - Collect snowfall data for North Temperate lakes
   - Collect downwelling shortwave radiation data for North Temperate lakes
   - Format meteorological data for model input
   - Obtain meteorological data for North Temperate Lakes

2. **Task Metadata**: Type = medium, Start date = 1st Oct 2014, Target date = 1st Dec 2014, Owner = Jordan Read, Participants = Paul Hanson, Jordan Read, Craig Snorholm, Expertise = hydrodynamic modeling

3. **Task Navigation**: Select core lake and catchment models, Develop mathematical model of age of water, Implement the lake model for North Temperate Lakes, Document the GLM lake model, Set up the GLM lake model for North Temperate Lakes, Gather bathymetric and elevation data, Calibrate catchment models, Run the GLM lake model, Verify and Validate the model, Make predictions and projections to the future.

4. **Personal Worklist**: Document the GLM lake model, Develop a computational model for water, Couple the lake and catchment models.


6. **Timeline Navigation**: Your Overdue Tasks, Write about the evaluation a day ago.

7. **Task Alert**: Obtain meteorological data for North Temperate Lakes
   - Collect downwelling longwave radiation data for North Temperate lakes
   - Collect air temperature data for North Temperate lakes
   - Collect relative humidity data for North Temperate lakes
   - Collect wind speed data for North Temperate lakes
   - Collect rainfall data for North Temperate lakes
   - Collect snowfall data for North Temperate lakes
   - Collect downwelling shortwave radiation data for North Temperate lakes
   - Format meteorological data for model input
   - QAQC meteorological data

---

**Tool Comparison Table**

<table>
<thead>
<tr>
<th>Category</th>
<th>Tool</th>
<th>Community support</th>
<th>Mainly task oriented</th>
<th>Semantic organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forum</td>
<td>V-Bulletin Forum</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>WordPress bbPress</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>WordPress p2 Theme</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Stackoverflow</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Mathematics</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Wiki</td>
<td>MediaWiki</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Semantic MediaWiki</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Enterprise Knowledge Management</td>
<td>Confluence</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Connections</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Jive</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>MS SharePoint</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Communote</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yammer</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ToDo Lists</td>
<td>Trello</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Keep</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Todoist</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Repositories</td>
<td>GitHub</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>BitBucket</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Google Code</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Organic Data Science Wiki</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Towards Scalable Visual Exploration of Very Large RDF Graphs

- Interactive graph visualization
- Graph layout ⇒ Spatial indexing
- User exploration ⇒ Window queries
- Wikidata (300M nodes & edges)
Analysis of companies’ non-financial disclosures: Ontology learning by topic modeling

- Prior studies highlight Linked Data for the analysis of company financial disclosures.

- Non-financial disclosures (e.g. companies’ environmental footprints) remain unexplored.

- We employ a Latent Dirichlet Allocation (LDA) topic model. Terms are retrieved from a SPARQL endpoint and seeded as lexical priors.

- An ensemble tree evaluates the quality of disclosures using ratings published by the Carbon Disclosure Project (CDP) and Google Finance.
Telling stories about scientists, automatically

familiarity

- trivial

discovery +

sense making

randomness -

surprise

Physics

Einstein

Newton

Hume

Semantic Distance Scoring

0.62 via Physics

0.45 via Hume
SentiML++ : An Extension of the SentiML Sentiment Annotation Scheme

Saad Malik Missen, Mohammed Attik, Mickaël Coustaty, Antoine Doucet, and Cyril Faucher

• Semantic annotation of opinions = very important task for opinion mining
  ✓ Training machine learning approaches
  ✓ Evaluating opinion mining methods

• First this paper discusses, compares and identifies the pros and cons of existing annotation schemes
  ✓ SentiML
  ✓ OpinionMiningML
  ✓ EmotionML

• Secondly, this paper provides an extension of SentiML
  ✓ Add target (on sentence level), holder, topic and informal sentence identification as part of SentiML++
  ✓ SentiML++ adds flexibility to SentiML by giving freedom of choice for a taxonomy when annotating the topic of a sentence
ProtégéLOV - the missing plugin for ontology engineering
FAGI-gis: A tool for fusing geospatial RDF data

- Fusion of interlinked geospatial entities
- Merging of spatial/non-spatial properties according to several fusion actions
- Semi-automatic property matching and batch fusion actions
GERBIL - General Entity Annotator Benchmark

Poster 6 - http://gerbil.aksw.org

Diagram:
- GERBIL Core Controller
- Benchmark Core
- Persistent Experiment Database (Model)
- Dataset Wrapper
- DataHub.io
- Open Datasets
- Interface View
- Configuration (Model)
- Your Dataset
- Your Annotator
- Annotator Wrapper
Controlling and Monitoring Crisis

Crowd Crisis Control

Viendo informes de Dec 29, 2013 a Jul 25, 2014

Twitter @E112Andalucia Alert: fire Crisis
18:00 Jul 24, 2014
Los desalojados por el incendio forestal de Mijas, Málaga en Rincón del Hinojal ya han vuelto a sus casas. Weather: Max temp: 31.5°, Min...

Twitter @112madrid Alert: fire Crisis
17:00 Jul 24, 2014
Hoy también aprieta el calor, así que conviene recordar estos Consejos112 para evitar Incendios de vegetación http://t.co/06V3rachHf #BIF

Twitter @GobEx112 Alert: fire Crisis
17:00 Jul 24, 2014
#Consejos112 NO utilices vertederos y puntos de vertidos ilegales ya que con frecuencia son puntos que originan incendios forestales.

Twitter @112Asturias Alert: fire Crisis
17:00 Jul 24, 2014

Twitter @nswpolice Alert: fire Crisis
09:00 Jul 24, 2014 Springwood, Australia

Incendio

Descripción
Inquiries are continuing after a house fire at Springwood overnight. http://t.co/laBtCZGar, Weather: Max temp: 21.11°, Min temp: 17.78°, Mean temp: 19.32°, Precipitation: 0.0 mm/hour, Humidity: 54%, Pressure: 1021.67 hpa, Speed wind: 1.1 m/s, Forecast: scattered clouds]

Credibilidad: 🚧 🚧 🚧

Deja Un Comentario
Nombre:
Correo electrónico:
ViCoMap: Visual Analysis of Statistical Data on Maps using Linked Open Data

Making Sense of Data on the Web...

The Men Who Stare at Data

NO MAPS, NO GLORY.

VS.

Petar Ristoski, Heiko Paulheim
UP TO 90% OFF ON HTML CALLS!!

SPARQL endpoints overloading your servers? Data dumps not staying up to date?

This was solved by **Triple Pattern Fragments**!

But wait, there is more!

TPF now overloading your network with numerous HTTP calls?

All this is solved by our brand new patented and optimized **QUERY PLANNING** for TPF!!!
DataOps

- **Data Source**
  - Low-level data access

- **Mapper**
  - Translation into triples
  - Extract and manipulate data

- **Post Processor**
  - Reconciliation (merging)
  - Improve data quality
Coworkers are highly dynamic

SmartKeepers provides a flexible working collaborative platform for coworkers

Decentralized through a flexible and elastic architecture to follow the working relations of coworkers

Secure with fine-grained controls and high level of security

Collaborative tools to assist coworkers in their business activities

Interoperability through W3C standards

Demonstration through a deployment on commodity boxes (raspberry Pi)
VOCBENCH 2.0
Collaborative Development of Multilingual Thesauri with VocBench
(System Description and Demonstrator)

VocBench is an open source web application for editing of multilingual SKOS and SKOS-XL thesauri, with a strong focus on collaboration, supported by workflow management for content validation and publication.

All editing actions inside VocBench undergo a validation and publication workflow.

The full history of actions per each project is stored inside VB and can be inspected by users (clients may also subscribe to its associated RSS feed).

SPARQL Editing with syntax completion and highlight

http://vocbench.uniroma2.it
DaCENA
SERENDIPITOUS NEWS READING WITH DATA CONTEXTS

Read a news article while visually exploring a serendipitous data context

Extraction of semantic associations from linked data (the article data context)

Ranking and selection of associations by serendipity (relevance & unexpectedness)

Web-based Interactive Visualization

M. Palmonari*, G. Uboldi°, M. Cremaschi*, D. Ciminieri°, F. Bianchi*

Play around with DaCENA online at dacena.densitydesign.org
QueryVOWL

Visual Composition of SPARQL Queries
The Russian Linked Culture Cloud Project

A knowledge graph bringing together the sources of open data on Russian Cultural Heritage

- The largest and most interlinked knowledge base for Russian culture
- Data from museums, libraries, archives
- Knowledge graph management for domain experts

- End-user tools: timelines, enriched text, interactive maps, pivot views, etc.
- A mobile app for museum goers
- Data navigation, analysis and powerful visualization

www.culturecloud.ru
Visualization example

Showing the dataset summary. Interactively. Just with SPARQL.

Just another RDF dataset visualization app? Come and see the demo to decide yourself.

- As a graph of **typical** class/property/datatype combinations.
- See vocab. usage: nodes are color coded by namespace.

- Changing the level of detail to see only more frequent combinations.
- Retrieving example instantiations of a selected subgraph.

100 million triples summarized in 5 minutes.

Needs just SPARQL endpoint access.

Web app ready to use.

Works on 50% of endpoints. But there are workarounds!

http://lod2-dev.vse.cz/LODSight
Supporting Real-Time Monitoring in Criminal Investigations
powered by Pundit - [http://thepund.it](http://thepund.it)
Contacts: christian.morbidoni@gmail.com
WHAT IF I TOLD YOU

http://csarven.ca/this-paper-is-a-demo
**Indications and Usage**
Recommended when you have to maintain a triplification $T$ of data stored in a frequently updated relational database $R$.

**Directions**
Do not triply the relational data again. Use RUBYA to incrementally maintain $T$ as described in Figure 1.

---

**Warnings**
You must follow our strategy to create $T$ as a materialized RDF view of $R$.
You must be able to add a few triggers to $R$, synthesized by RUBYA.
You must add a view controller to the triplestore storing $T$, also provided as part of RUBYA.

---

**MADE IN BRAZIL**
Distributed by UFC - Federal University of Ceará
What types of resources are there in a data set? How are they described?

What properties are used to link different types of resources? How frequently?

Are data described as prescribed by the ontology?

**Linked Data Summaries**

*Frequency of Abstract Knowledge Patterns* `<Type,Property,Type>` extracted from the data

Represented in RDF, queryable, navigable

explore summaries online at http://abstat.disco.unimib.it:8880

“141000+ Music Artists are linked to Tracks by the property made”
How to Stay Ontop of Your Data: Databases, Ontologies and More

Big Data Challenges:

-ontop- is a platform to query databases as virtual RDF Graphs using SPARQL, and exploiting OWL 2 QL ontologies and mappings.
LIDSEARCH

User query: Names and books of writers born in Paris

Data:
- Name: Anatole France, Book: Thais
- Name: Anatole France, Book: Crainquebille
- Name: Maurice Druon, Book: Five Miles to Midnight
- Name: Maurice Druon, Book: Vertigine d’amore

Recommended Services:
- Service: Amazon, Description: Book prices
- Service: Fnac.com, Description: Buy books
- Service: Facebook, Description: Friends who read these books

Services from different repositories

M.L. Mouhoub – Univ. Paris Dauphine
Distributed Linked Data Business Communication Networks

The LUCID Endpoint

Linked data managed (1,2), published & consumed (3,4,5) using LUCID endpoints

Revisioning data modelled in RDF, endpoints exchange revisioning messages (3) via Pubsubhubbub

Demo Use-Case: Master Data Management
The Island of Misfit Entities

Enforce quality and coverage in KGraphs WikiData/DBPedia: notability criteria
What about tail entity structured representation
Leveraging people diversity: Push Crowdsourcing

dbpedia:Noma_(restaurant)

Hiša dobrot, gostinske storitve, d.o.o.
Vilfanova ulica 10
6320 PORTOROŽ

The long tail of entity popularity
The Research Articles in Simplified HTML (RASH) is an HTML(+RDF) format for writing academic research articles. Please do not hesitate to contact me for additional information about the RASH Framework or if you would like to use RASH for allowing HTML submissions to your academic event.

Silvio Peroni
silvio.peroni@unibo.it  @essepuntato
Dept. of Computer Science and Engineering
University of Bologna, Bologna, Italy

Venues that have already adopted RASH:

- Learning in the Cloud LC 2015
- SAVE-SD 2015
- LINKED 2015
- 1st International Workshop on LINKed Education at the ISWC 2015
- Semantic Publishing Challenge 2015

Future works: converters into/from additional formats
Future works: word processor extensions
Conversion from ODT format
Conversion into LaTeX styles
ACM ICPS and Springer LNCS

Publisher

Author

Reader

Future works: converters into/from additional formats

Conversion into
ODF

Conversion from
ODT format
Ibuprofen was extracted by the research arm of Boots Company during the 1960s ...

Automatically learned enrichment pipeline

1. Dereferencing
2. Authority Conformation
3. NLP
4. Filter
5. Predicate Conformation
6. Dereferencing
7. Authority Conformation
8. NLP

• Fully managed RDF DBaaS available 24/7
• Ideal for small / moderate data & query volumes
• Instantly deploy new databases whenever needed
• Zero administration
  – Automated operations, maintenance & upgrades
• OpenRDF REST APIs
• Designed for high availability & resilience
• Use if for free at http://s4.ontotext.com
The Self-Service Semantic Suite (S4)

- Available anytime, anywhere
- Scalable & reliable
- Simple REST APIs
- GATE & UIMA plugins
- SDKs
- Use it for free at s4.ontotext.com

RDF Data

Knowledge Graphs

- Access to knowledge graphs
- DBP, GN, FB/WD, biomedical

Text Analytics

- News analytics
- Biomedical analytics
- Twitter analytics
- Entity disambiguation & linking

Deputy Defence Minister Kostas Isichos told the BBC northern Europe must do more to rescue and shelter migrants.

- RDF database as-a-service
- Self / fully-managed
International Literature (English)

Country-Specific Literature (multilingual)

Electronic Health Records (multilingual)

Cloud-Based Solution Provider

Kconnect Cloud Services

Source KBs

Update/New Languages

Professional Services Community

Foreseen Adopters: Publishers, Pharma, Medical Institutes, ...

Early Adopters: Search Engine Providers (TRIP, HON)

Search Logs

Fast Systematic Reviews

Early Adopters: Hospital/County Admin (KCL, QUL)

Kconnect Local

Analysis Results

check out KConnect.eu
Anna Wróblewska, Data Scientist  
anna.wroblewska@allegro.pl
Research Cases

What would research look like...

...reenlightened?

http://researchcases.org
#TripleProv

the first provenance-aware triplestore

Linked Data is linked data.
Linked Data is heterogenous data.

Different Provenance

➢ store and track provenance (WWW 2014)
➢ tailor queries with provenance (WWW 2015)
➢ a demo of TripleProv (VLDB2015)
HCLS Dataset Descriptions

W3C Interest Group Note
http://www.w3.org/TR/hcls-dataset/