



# 3<sup>RD</sup> INTERNATIONAL WORKSHOP ON SEMANTIC DIGITAL ARCHIVES

VALETTA, MALTA  
SEPTEMBER 26, 2013



## PUNDIT: CREATING, EXPLORING AND CONSUMING SEMANTIC ANNOTATIONS

Marco Grassi<sup>(1)</sup>, Christian Morbidoni<sup>(2)</sup>, Michele Nucci<sup>(3)</sup>,  
Simone Fonda<sup>(4)</sup>, Francesca Di Donato<sup>(5)</sup>



(1,2,3) DII - Department of Information Engineering. Polytechnic University of Le Marche, Ancona, Italy

(4) NET7 srl, Italy

(5) Scuola Normale Superiore, Italy

# OUTLINE

1.PUNDIT VISION

2.PUNDIT ANNOTATION MODEL

3.DISPLAYING ANNOTATION DATA

4.SOME EXAMPLES

5.CONCLUSIONS

# Pundit



## DEVELOPED BY:



**Semedia,**  
Università Politecnica delle Marche  
<http://semedia.dii.univpm.it>



**Net7 SRL**, Pisa  
[www.netseven.it/](http://www.netseven.it/)

## SUPPORTING PROJECTS:



**DM2E EU Project**  
<http://dm2e.edu/>

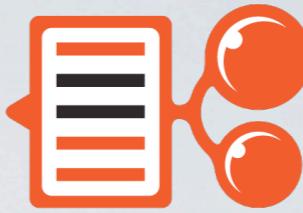


**AGORA EU Project**  
<http://project-agora.eu/>



**Semlib Project Eu Project**  
<http://semedia.dii.univpm.it>

# Pundit



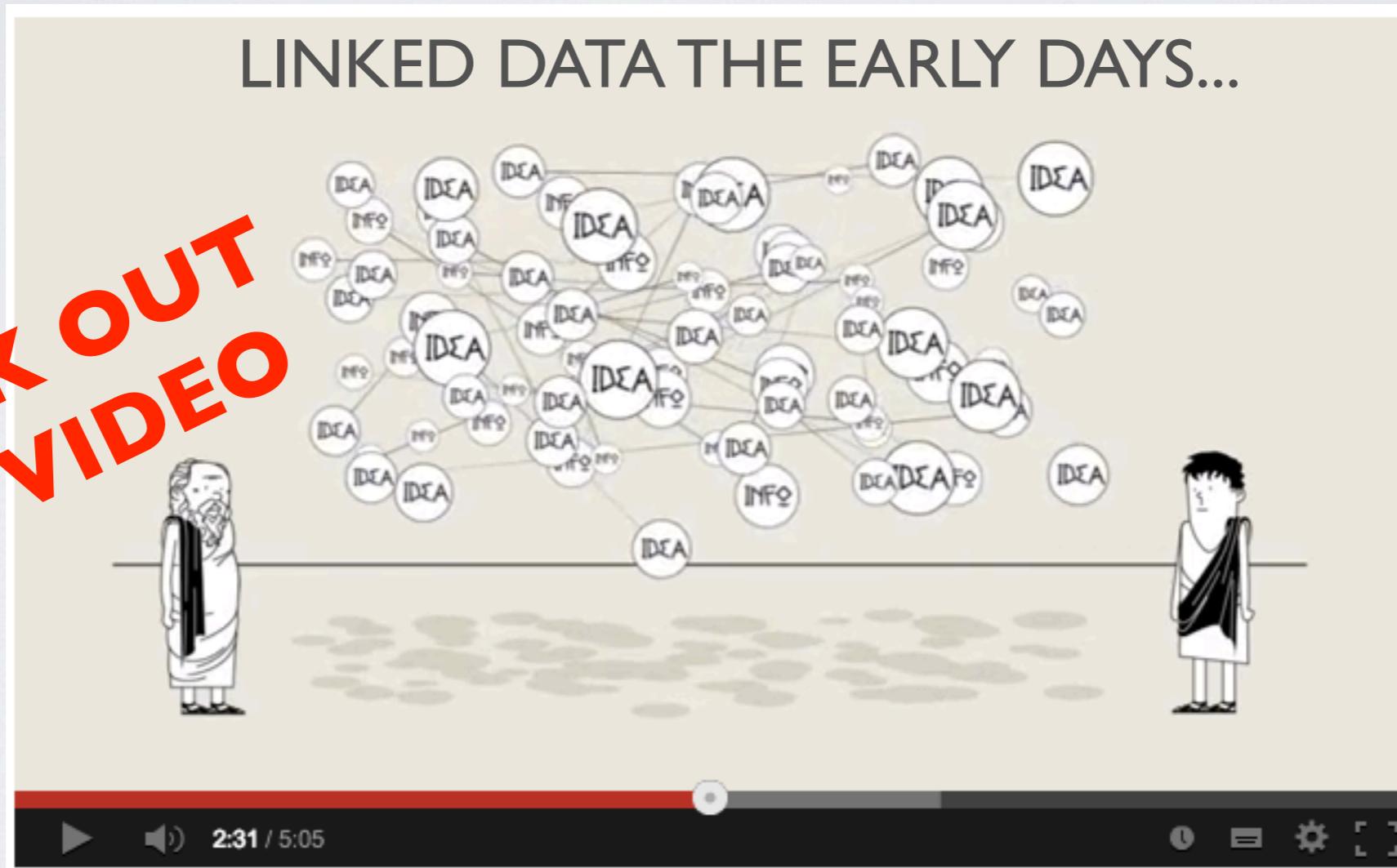
Pundit won the LODLAM Challenge 2013!

Data visualizations, tools, mashups for Linked Open Data in libraries, archives, and museums

<http://summit2013.lodlam.net>

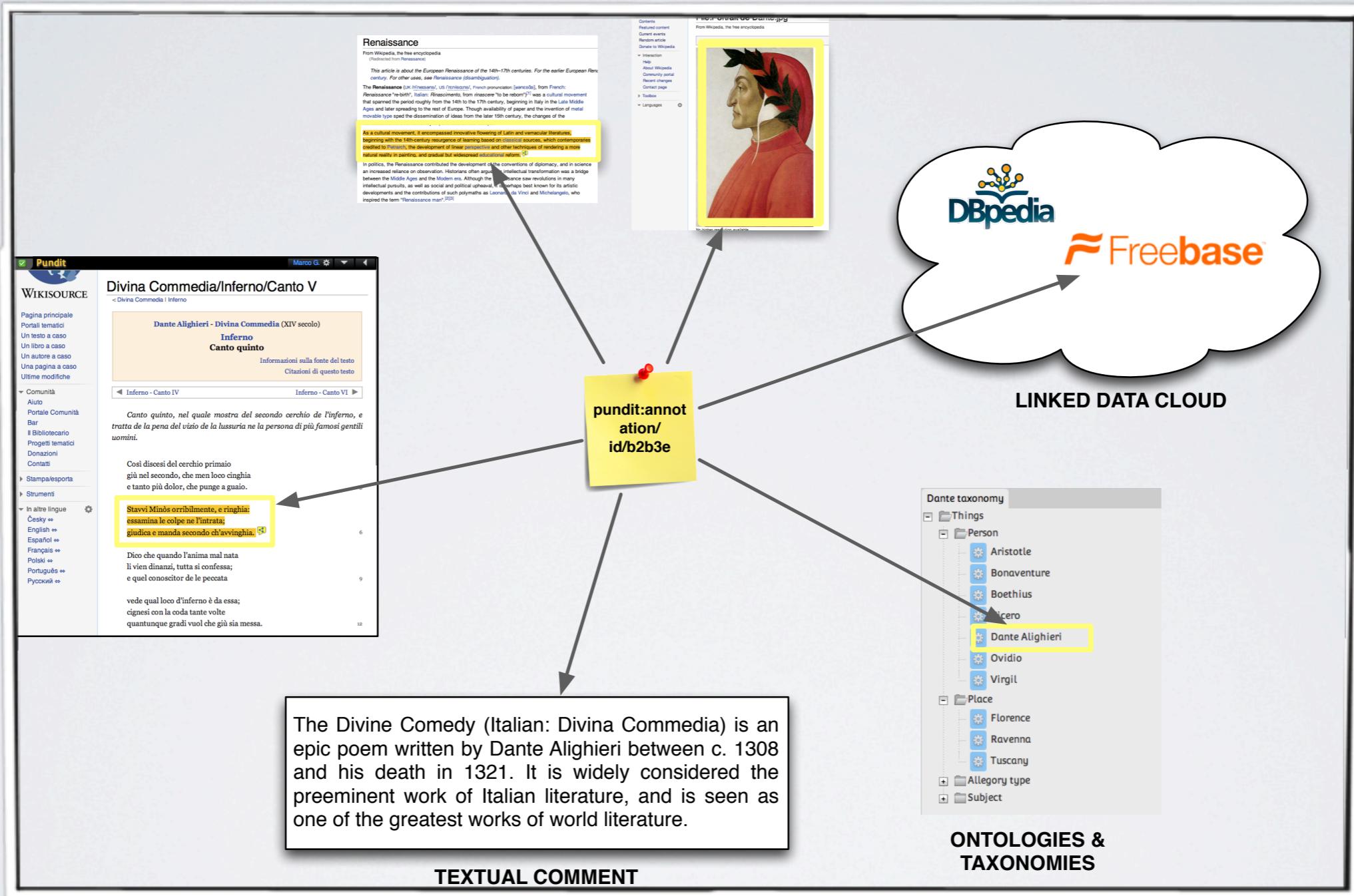
LINKED DATA THE EARLY DAYS...

CHECK OUT  
THE VIDEO



[http://www.youtube.com/watch?v=6uUQ4f3z\\_E0](http://www.youtube.com/watch?v=6uUQ4f3z_E0)

# PUNDIT VISION



Pundit allows user to generate semantically structured data when they create annotations!

# PUNDIT VISION

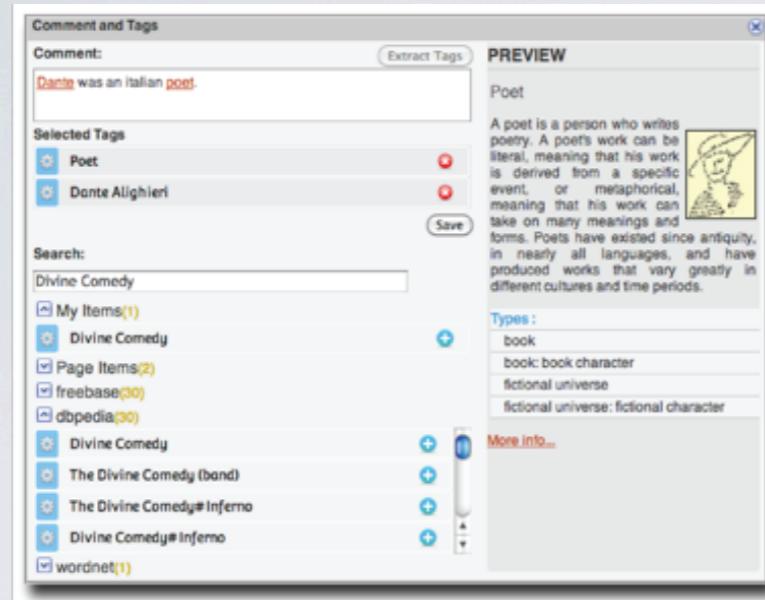
RESTful APIs are provided to consume created annotations (data):

→ SEARCH / EXPLORE / SHARE THE ANNOTATIONS

→ EASILY CREATE DYNAMIC DATA VISUALIZATIONS



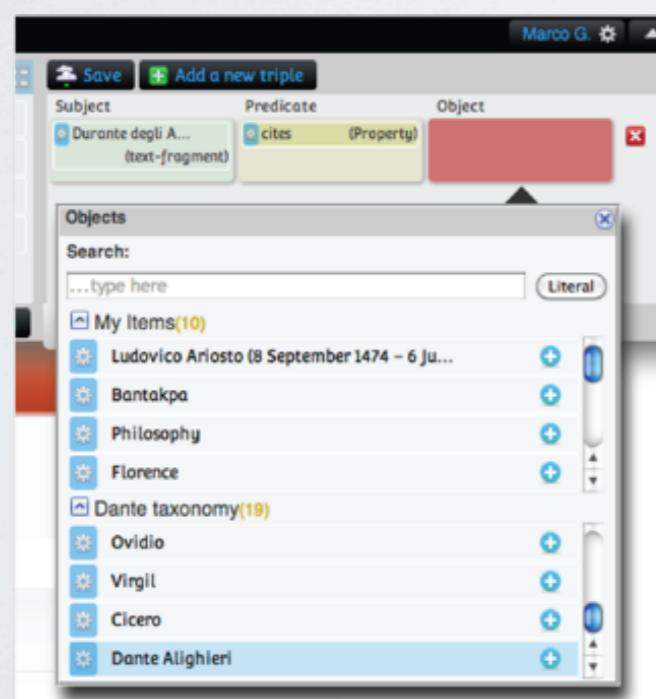
# CREATING SEMANTIC ANNOTATIONS



COMMENT/ SEMANTIC TAGS



TEXT RELATIONS



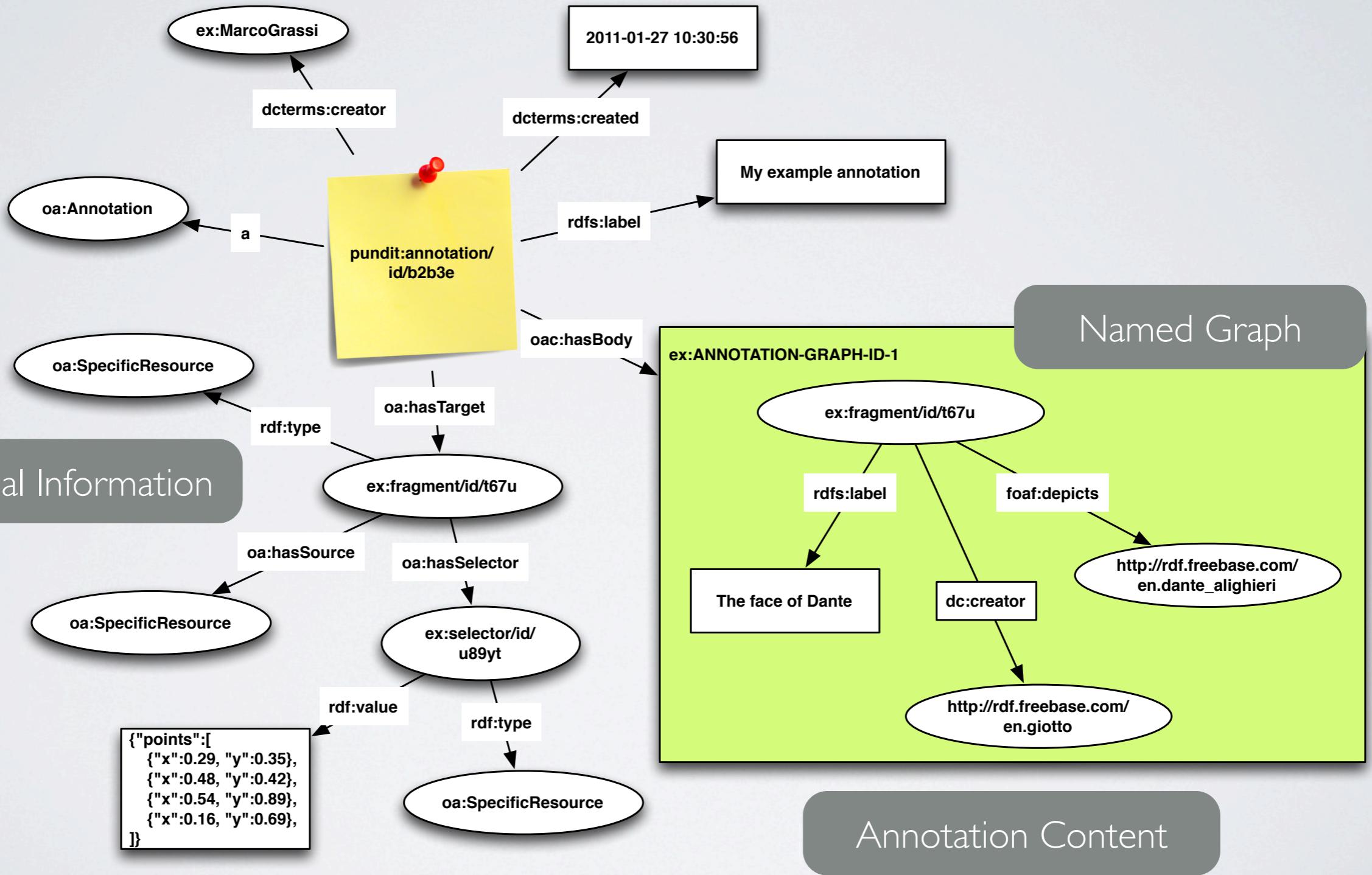
TRIPLE COMPOSER



IMAGE ANNOTATION

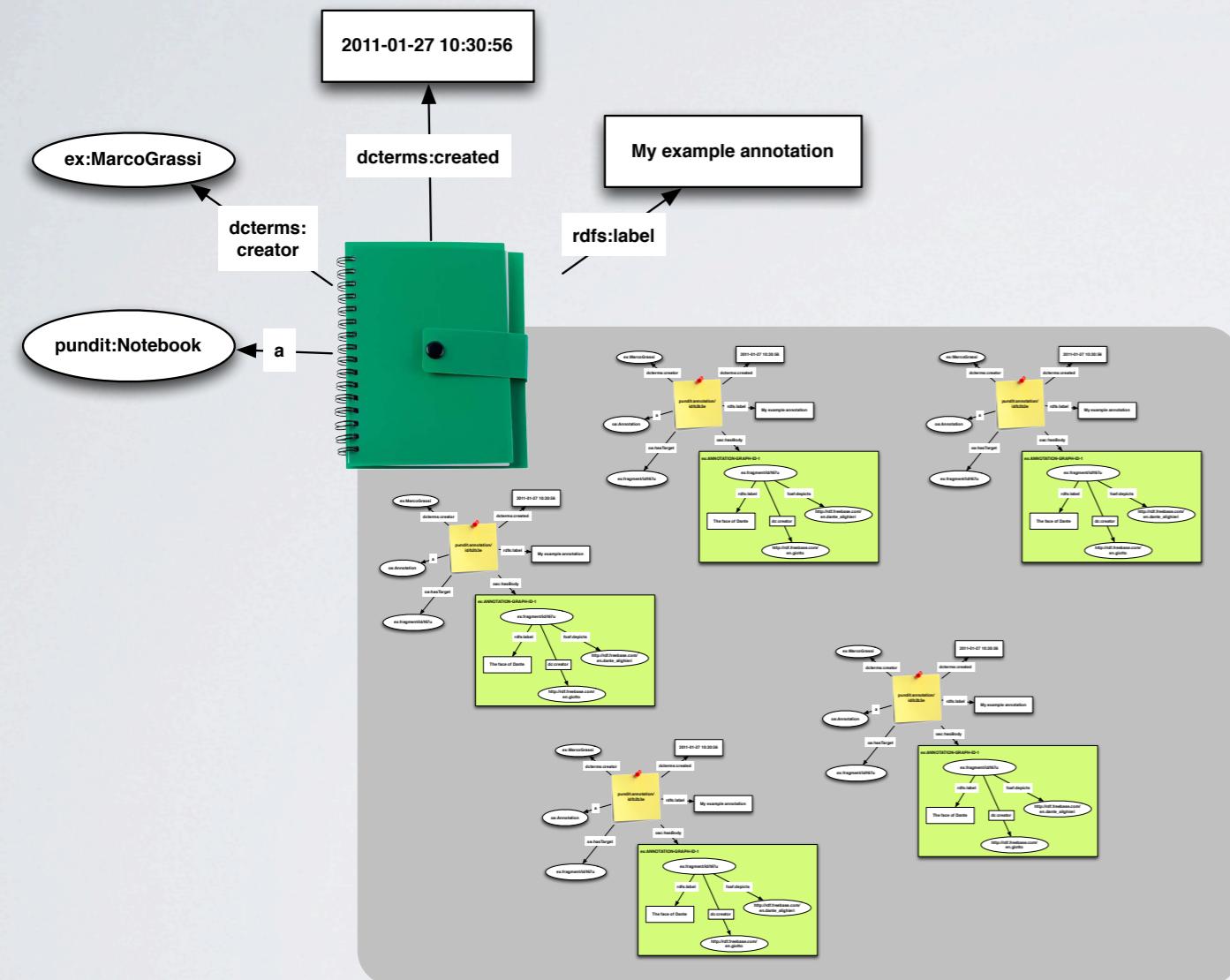
# ANNOTATION MODEL

- Based on Open Annotation Collaboration (OAC)



# NOTEBOOKS

- Annotations are collected in **notebooks**

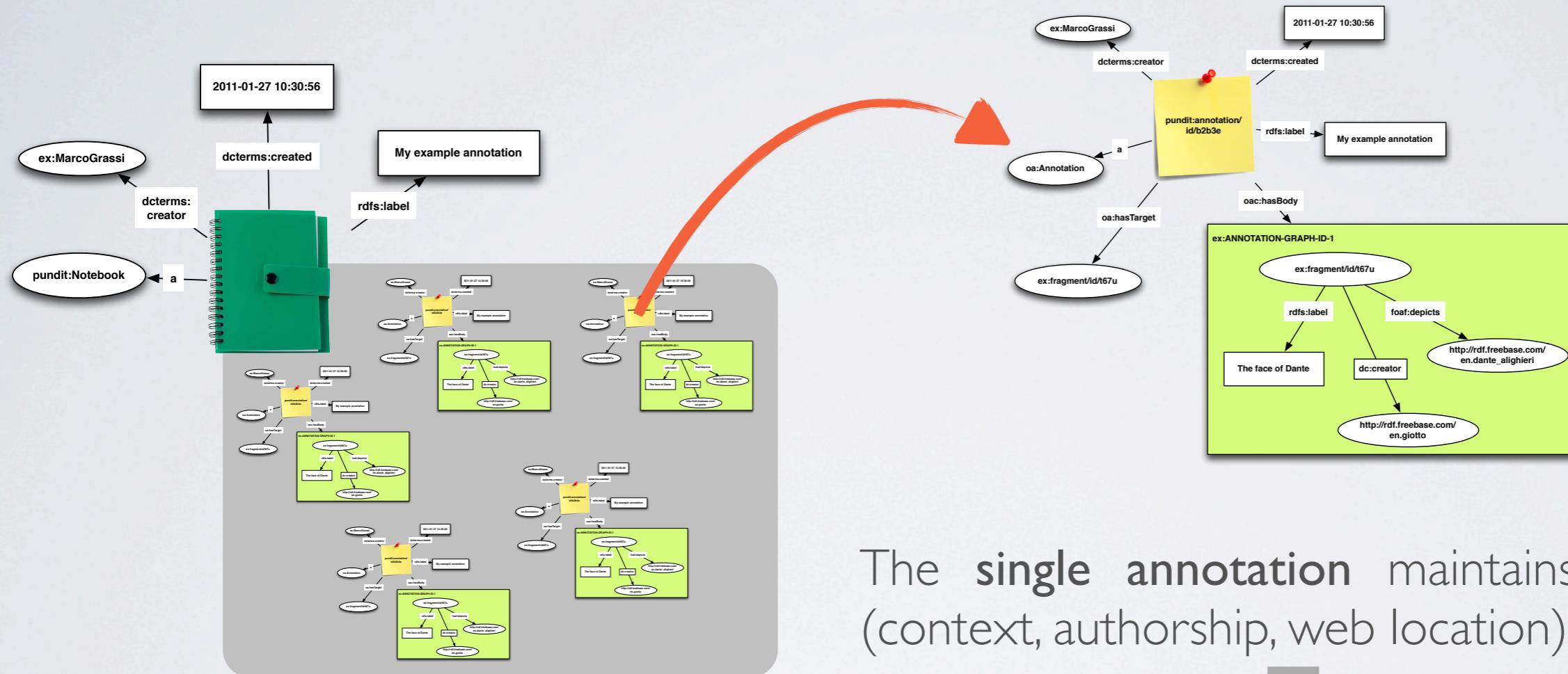


- Annotations are **shared** at notebook level.

- Users can **organize** their annotations in different notebooks
- Set as **Public/Private**
- Activate/Deactivate** to filter the amount of public annotations visualizing only those of interest.
- Identified by a (dereferenciable) **URI**

# DATA VISUALIZATION ANNOTATION CENTRIC

- play a fundamental role for data visualization:



The **single annotation** maintains its integrity  
(context, authorship, web location)



Explore collection of annotation grouped into notebooks

# ANNOTATION CENTRIC VISUALIZATION

# CONTEXTUAL VISUALIZATION

**PREVIEW**

Freebase DBpedia Word Net Europeana EDM bur-production

Save Add a new triple

Subject Predicate Object

Römisc... (Text Fragment) is relat... (RDF Property)

Types : topic person deceased\_person author

1 My Items 26 Page Items Vocabulary Triple Composer

hermann-kestner-1863-02-11 ⓘ

More info... Hannover den 11. Febr. 1863. Hochverehrtester Herr und Freund!

Längst hätte ich Ihnen für Ihren freundlichen Brief vom 15n Sept v J, (den ich übrigens erst am 7n Jan d. J. durch meine Tante erhalten habe), so wie für Ihre gütigen Bemühung wegen der Lösung einiger künstlerischer Rätsel gedankt, wenn ich nicht seit über zwei Monaten in die volle Widerwärtigkeit meiner halbjährigen Geldgeschäfte vertieft gewesen wäre, deren Abwicklung leider immer wieder immer jeder andern Beschäftigung vorgehen muß

Gott sei Dank habe ich nun einmal wieder ein paar ruhige Wochen und kann an solche erquickliche Dinge, wie es mir stets jede Seite von Kunst und Wissenschaft gewesen ist, ohne zu große Gewißensbisse wohl ein bisschen Zeit verschwenden

Von den Ihnen so unbescheiden von mir im Vertrauen auf unsere Römischen Sympathieen auferlegten Rätseln, habe ich seitdem mit

Hülfe eines jungen Reisenden aus Osnabrück Osnabrück, der schwerbeladen mit Photographien Photographeen heimkehrte, drei Gegenstände glücklich heraus gebracht, die ich, wenn es Sie, wie ich vermuthe, interessirt, näher bezeichne, nämlich Lit C Figuren aus der Kreuzig des Giottino in San Francesco d'Assisi Fresco Lit G Madonna auf dem Throne in San Domenico in Perugia von Fiesole, tempera Lit S Himmelfahrt von Fiesole in Florenz in S Marco Temp od Fresco

So hoffe ich denn allmählich auch mit den übrigen noch in's Reine zu kommen Die Photographie ist ein vortreffliches Hülfsmittel zu historischen Studien und bedauere ich nichts mehr, als daß man nach neueren Erfahrungen an der Haltbarkeit zweifelt.

Hoffentlich geht es Ihnen in den ange- nehmen angenehmen Kreisen und Beschäftigungen in denen Sie in Basel leben in

By Costanza Giannaccini on 2013-07-15  
 Created by: Costanza Giannaccini  
 On: 2013-07-15, 10:36:37  
 ID: 97d953fa  
 ein paar ruhige Wochen  
 period of dates ends at  
 1863-02-26

# ANNOTATION CENTRIC VISUALIZATION ASK (THE PUNDIT)

My Ask

Logout

Logged in as Marco G.  
— margra75@gmail.com

Create a notebook

Your notebooks

**public** Pundit Notes  
— 2012-09-22T16:55:31  
— 7 annotations  
[Browse](#) [Delete](#)

**public** Italian Literature  
— 2012-09-25T21:48:09  
— No annotations  
[Browse](#) [Delete](#)

**public** Smart Home  
— 2012-12-07T18:21:48  
— 18 annotations  
[Browse](#) [Delete](#)

**public** OWL to Java  
— 2012-11-28T12:46:34  
— No annotations  
[Browse](#) [Delete](#)

**BETA** Import your annotations here

Notebooks My Ask My N: Pundit Notes

**Pundit Notes fa0fe246**

Visibility public  
By Marco G.  
Created at 2012-09-22T16:55:31  
# annotations 7

[img: big-lebowski-still-hr-3.jpg?resize=506,340]



Marco G.  
2013-09-23T16:13:00  
jonesreportla.com/ailec\_eve ..

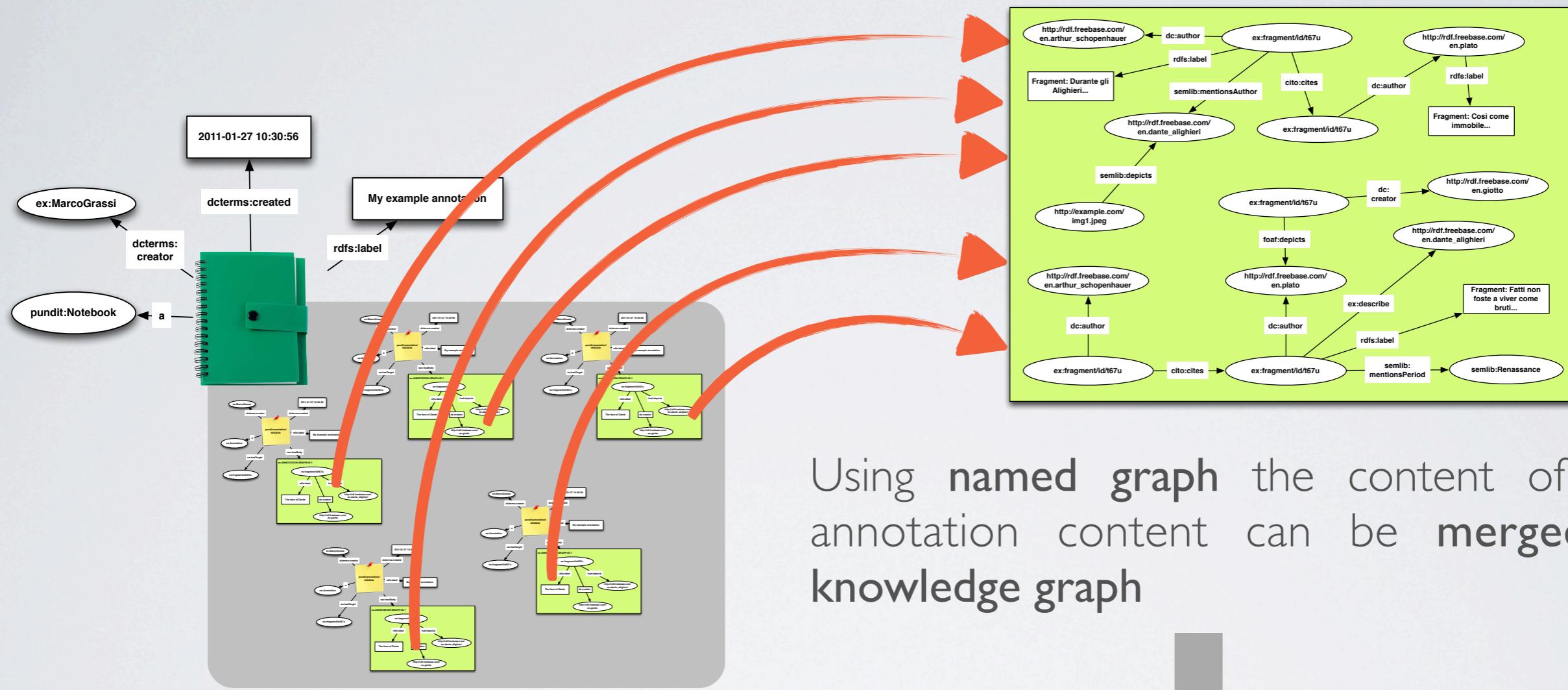
<http://schema.org/comment>

The dude takes it easy.

A portal to manage annotations, share them and explore public notebooks

<http://ask.thepund.it/>

# DATA VISUALIZATION ITEM CENTRIC



Using **named graph** the content of single annotation content can be **merged** into **knowledge graph**



Visualization on the **annotated items** and their **relations** with other items

**A-priori** knowledge on ontologies/custom vocabularies and relations used in annotations is **beneficial!**

# ITEM CENTRIC VISUALIZATION CONFIGURING PUNDIT

## Custom Vocabularies/Taxonomies:

- Online JSONP file (created manually or automatically from an ontology )
- Add URL to Pundit configuration

The screenshot shows the Pundit application interface. At the top, there's a navigation bar with tabs for Freebase, DBpedia, Korbo, Wordnet, HEO, Wordnet Affect, and Dante taxonomy. Below the navigation bar, there's a preview area for the item 'Florence'. The preview text describes Florence as a famous medieval European city, a center of trade, and the birthplace of the Renaissance. It includes a small thumbnail image of a building. To the right of the preview is a sidebar titled 'Pundit' with a tree view of categories: 'Things' (Person, Place), 'Allegory type' (Literal, Typological, Tropological (Moral), Anagogical), and 'Subject'. At the bottom of the sidebar, it says '20 My Items'.

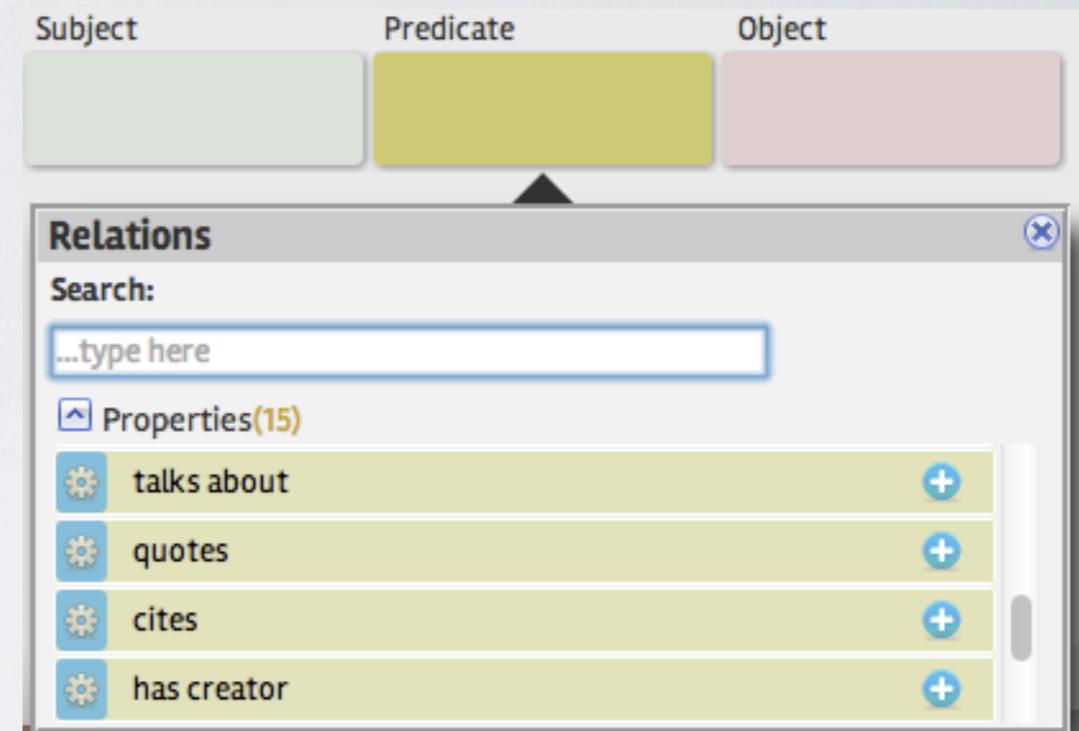
## Linked Data Selector:

The screenshot shows the Linked Data Selector interface. At the top, there are tabs for Freebase, DBpedia, Word Net, and Europeana EDM. A search bar contains the query 'dante ali'. Below the search bar, it says '30 RESULTS FOUND:' followed by a list of 30 results. Each result is a row with an icon, the name, and a description in parentheses. The results include various entities like Dante Alighieri, Dante, Inferno, Mos Def, and Dante Gabriel Rossetti, with their respective semantic types.

| Result                 | Description   |
|------------------------|---|
| Dante Alighieri        | (Topic, Context name, Musical Artist, ...)                  |
| Dante                  | (Football player, Athlete, Athlete extra, ...)              |
| Dante                  | (Topic, Fictional Character, Video Game Character)          |
| Inferno                | (Written Work, Work of Fiction, Book, ...)                  |
| Mos Def                | (Record Producer, Award Winner, Music video performer, ...) |
| Dante Gabriel Rossetti | (KWTopic, Client, Lyricist, ...)                            |
| Dante                  | (Person, Musician, Award Nominee, ...)                      |
| [...]                  |   |

# ITEM CENTRIC VISUALIZATION CONFIGURING PUNDIT

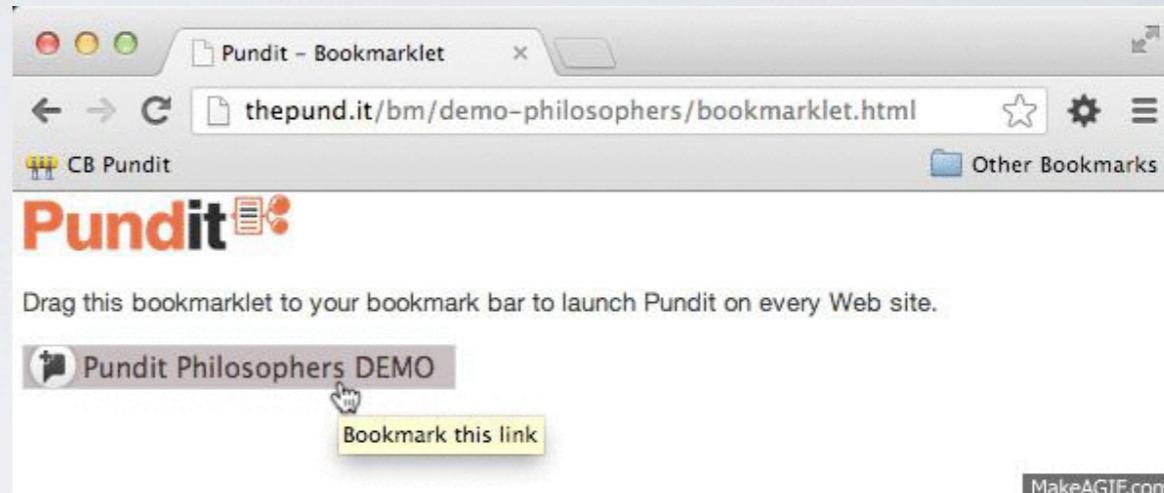
Predicates used in relations:



- address **different communities** and foster **data reuse**
- create **specific pattern** in the annotations to create **engaging user interfaces**

# ITEM CENTRIC VISUALIZATION DEPLOYING PUNDIT

- As a JavaScript library



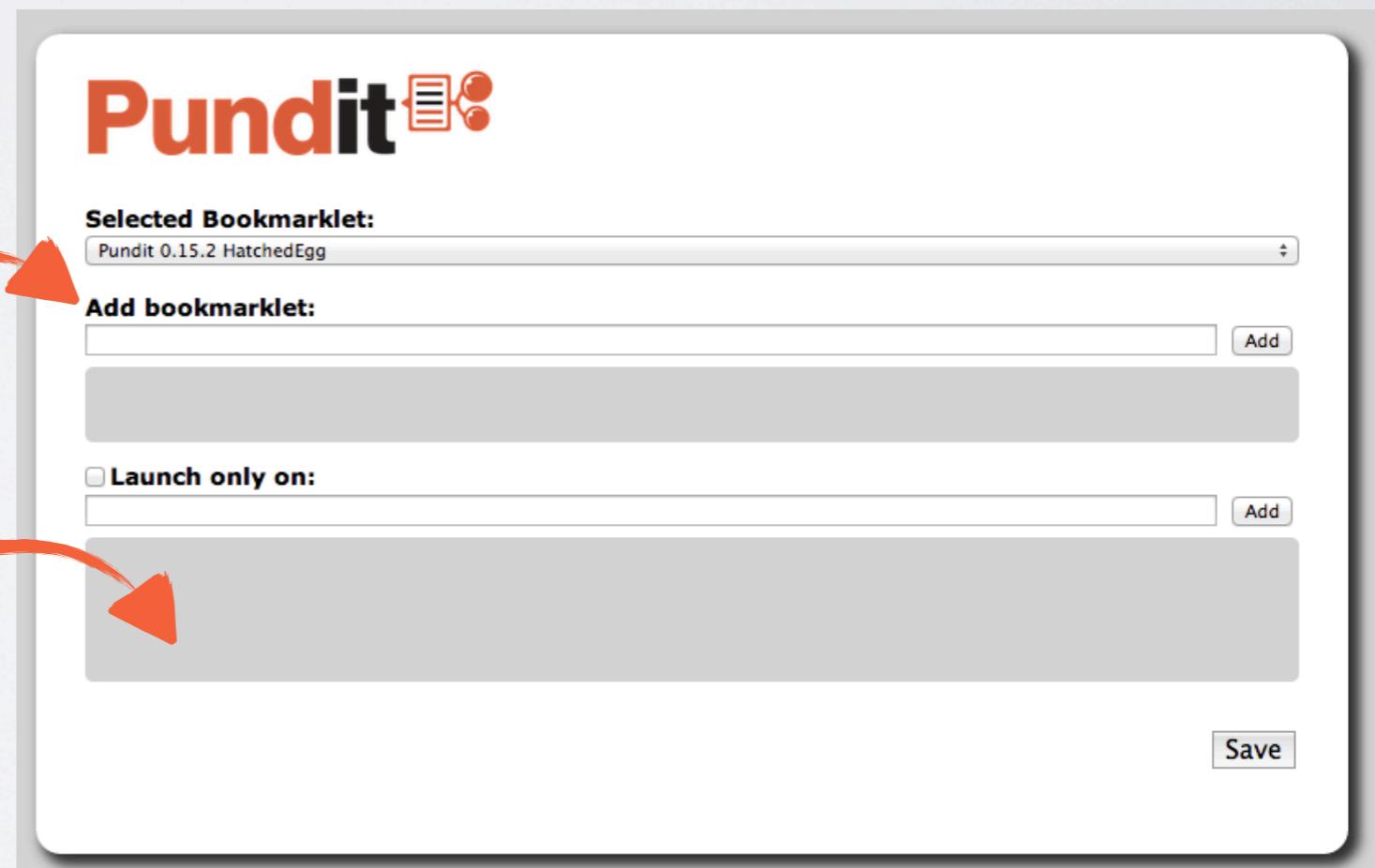
- As a bookmarklet

- As a Browser Extension  
(Chrome and Firefox)

Select between different instances of Pundit

Restrict the web pages where Pundit is automatically launched

[https://github.com/marcograssi/  
PunditBookmarklet](https://github.com/marcograssi/PunditBookmarklet)



# ITEM CENTRIC VISUALIZATION CREATING DATA VISUALIZATION

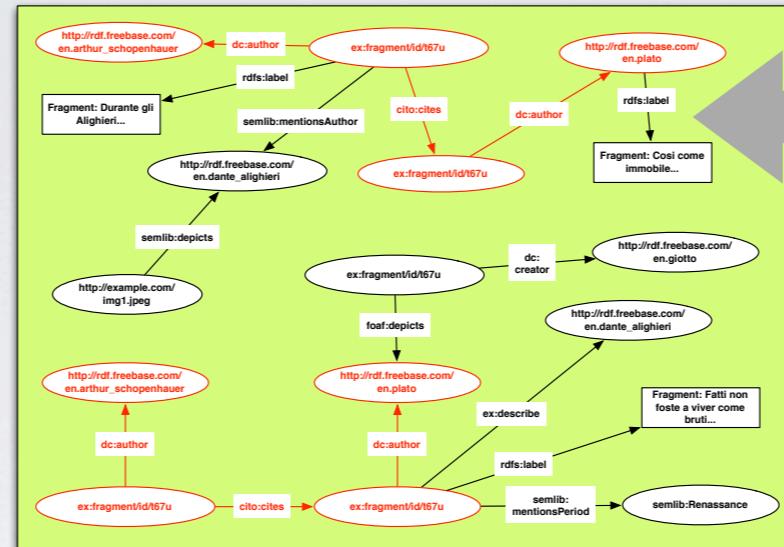
COMMUNITY USERS



PUNDIT  
custom



PUNDIT  
custom

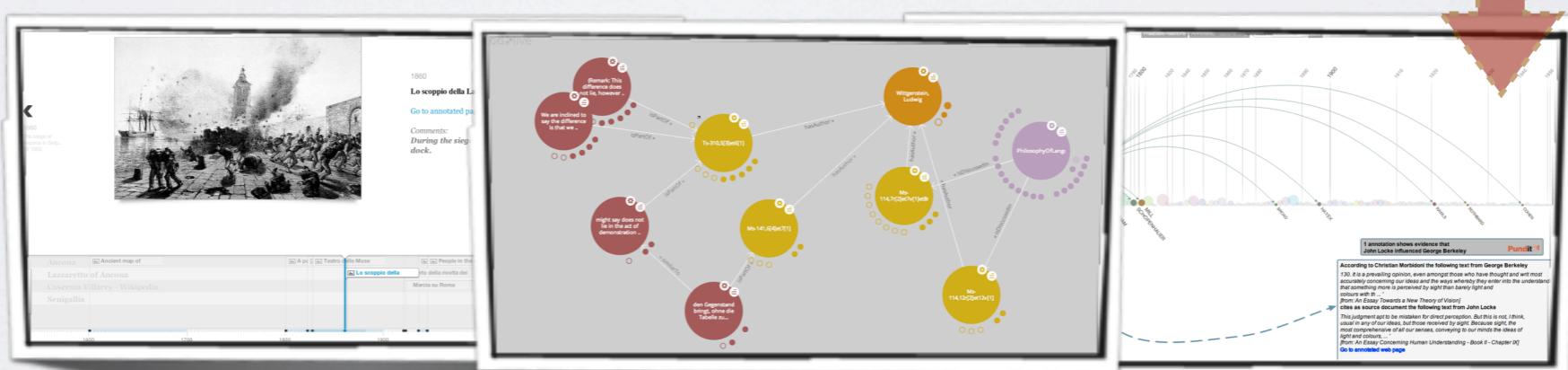


REST API + SPARQL

Specific patterns &  
Standardized vocabularies

Other data sources

Third party  
libraries and tools

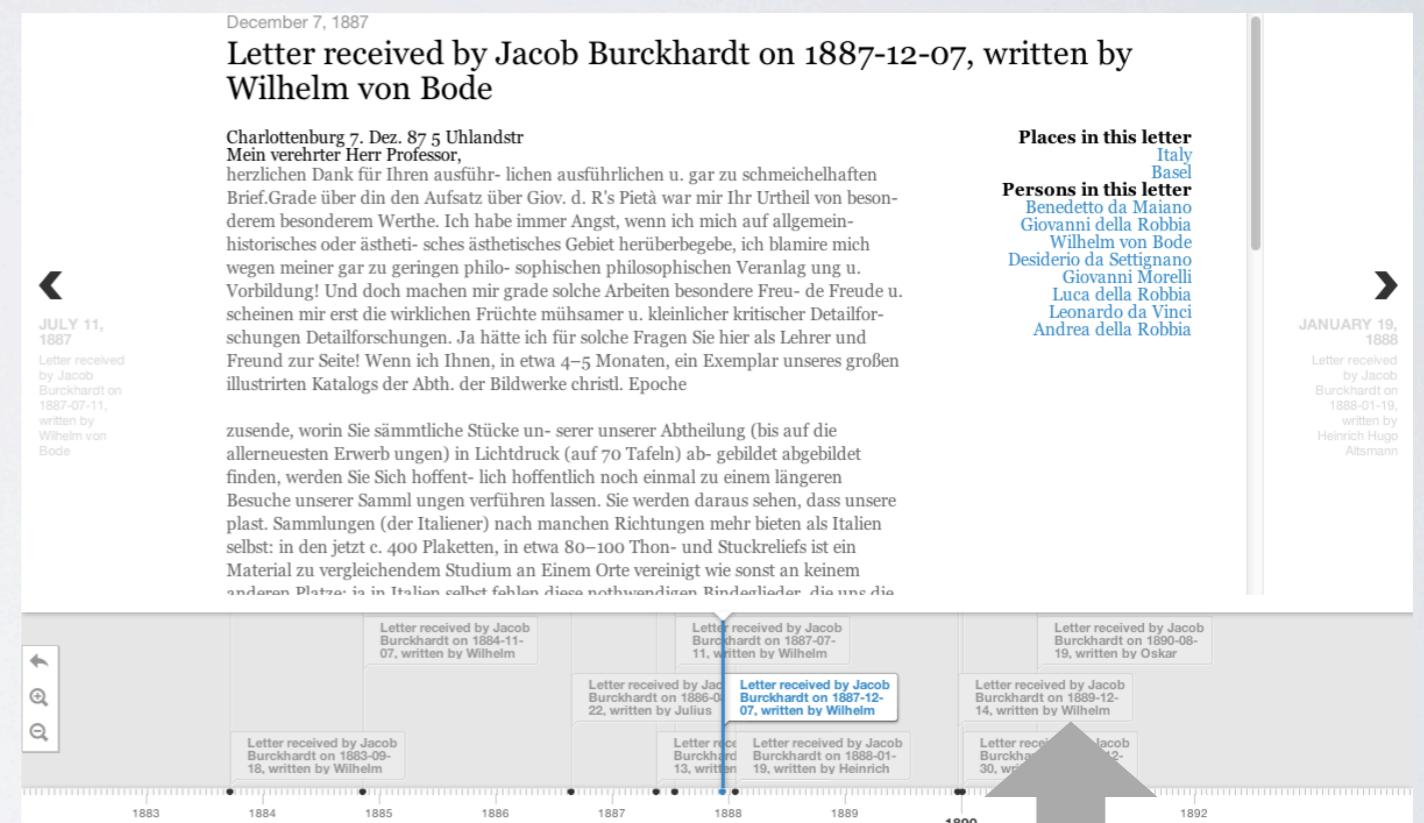
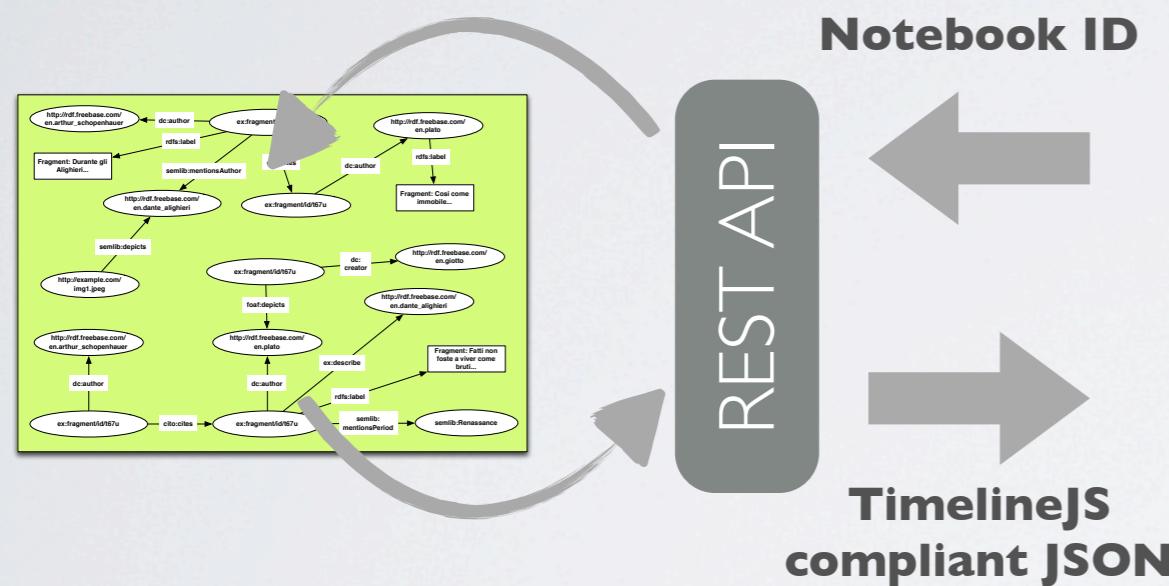


## SPECIALIZED DATA VISUALIZATION

# TIMELINE

## EASILY CREATE INTERACTIVE TIMELINE USING PUNDIT...

- Create annotations in a notebook of text fragment or images, containing a date, title and creator.



- Reuse TimelineJS (<http://www.timeline.verite.co/>)

**Timeline JS**

# TIMELINE

...JUST LIKE THIS ONE I'VE CREATED DURING THE WORKSHOP

11:35 AM  
BlogNEER:  
Applying Named  
Entity Evolution  
Recognition on  
the  
Blogosphere?

Semantic Web Vision  
[...] open up the knowledge and workings of humankind to meaningful analysis by software agents [...] [Garcia-Luna, T., Hinsler, J., Klemenz, O., et al. The semantic web. Scientific American. 2013; 305\(5\):28-37.](#)

2BW Page 2

12:10 PM  
September 26, 2013  
Semantification of Query Interfaces to Improve Access to Deep Web Content

[Go to annotated page >>](#)

Comments:  
by Arne Martin Klemenz and Klaus Tochtermann

2:00 PM  
Towards  
Preservation of  
semantically  
enriched  
Architectural  
Knowledge

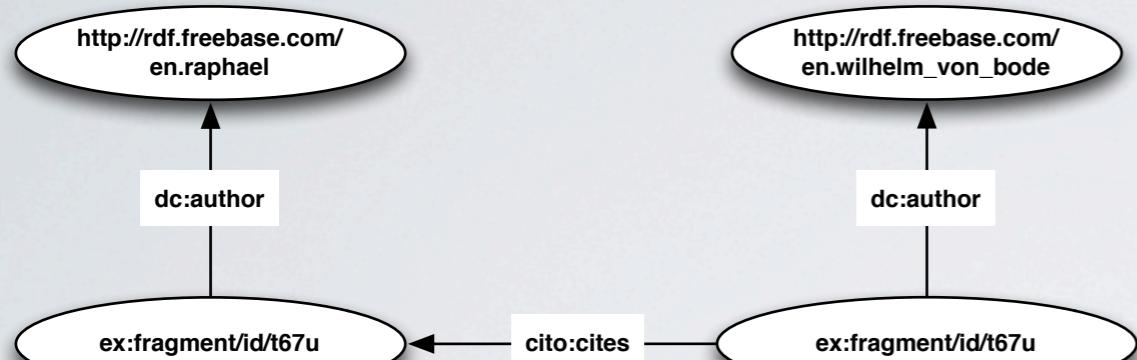
10:00 AM  
11:00 AM  
12:00 PM  
1:00 PM  
2:00 PM  
3:00 PM

Digital Geolinguistics: on the use of Linked Data  
Elevating Natural History Museums' Cultural Collections  
The Role of Language Evolution in Digital Archives  
BlogNEER: Applying Named Entity Evolution  
Semantification of Query Interfaces to Improve Access to Deep Web Content  
Towards Preservation of semantically enriched Architectural Knowledge  
A Storage Ontology for Hierarchical Storage Management  
A linked open data architecture for contemporary cultural heritage management

<http://semedia.dii.univpm.it/news/16-having-fun-with-pundit-sda-2013>

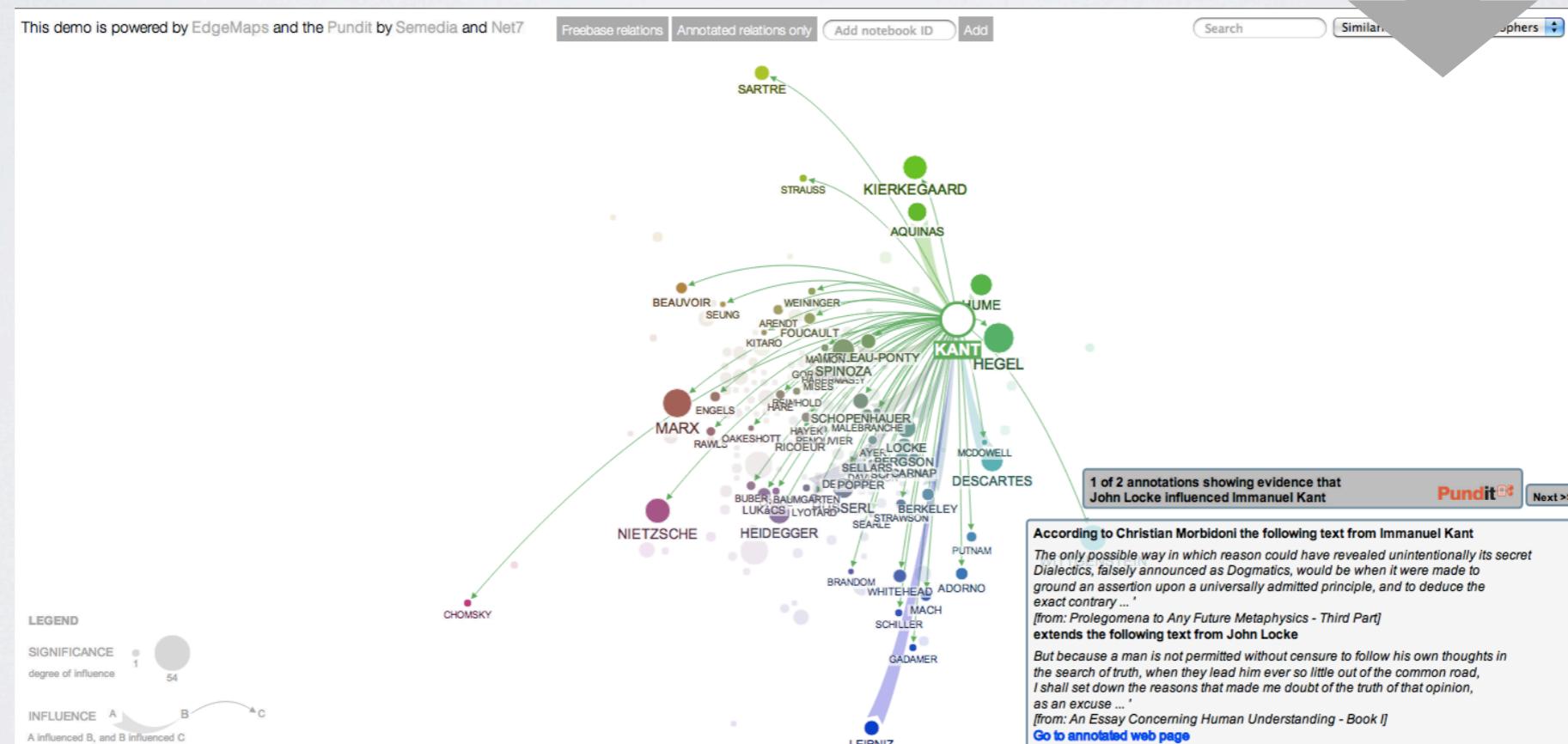
# EDGEMAPS VISUALIZATION

Every-time that an annotation like this is created



A relation is generated between the authors

**influenced**

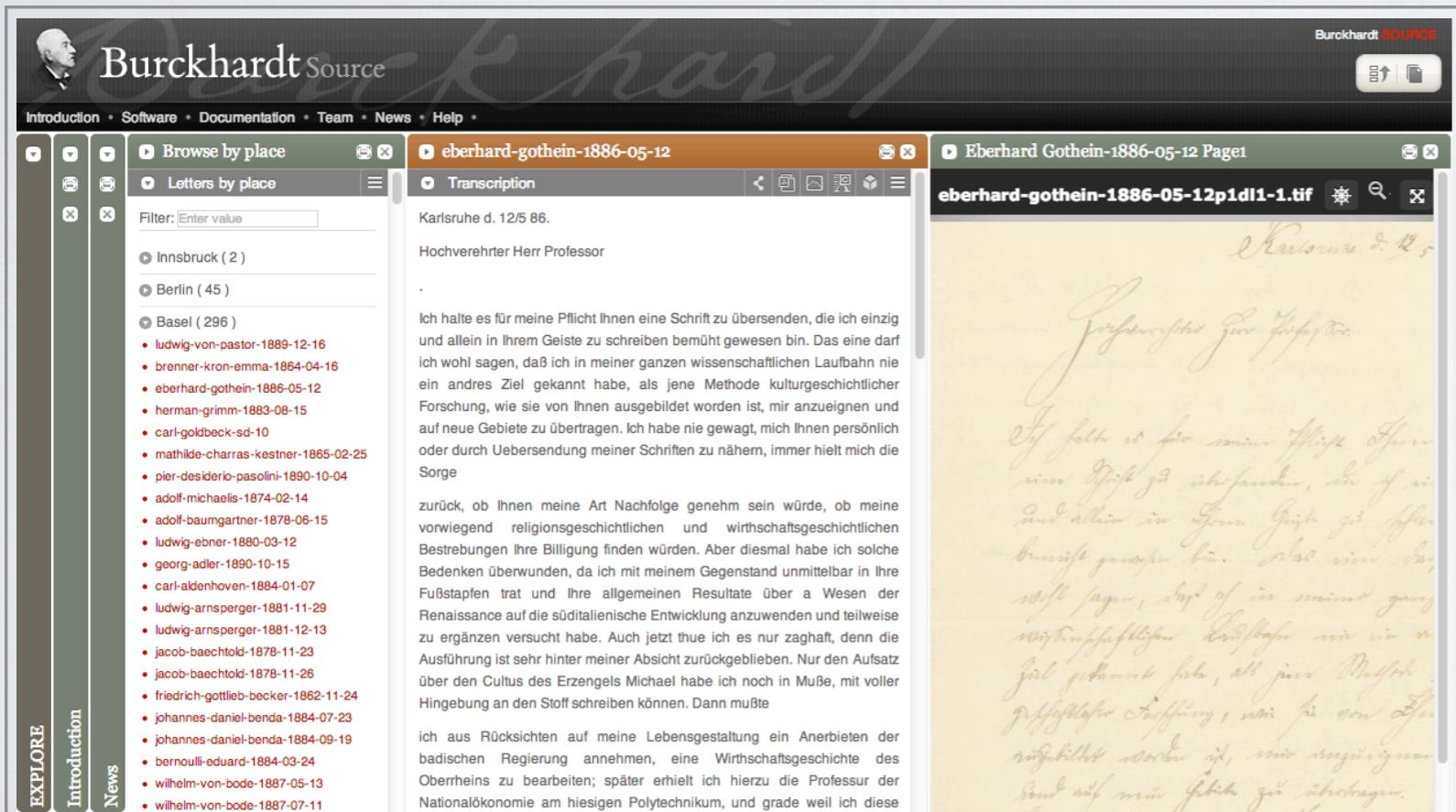


Reuse <http://mariandoerk.de/edgemaps/>

# BURCKHARDT SOURCE

<http://burckhardtsource.org/>

- Burckhardtsource.org platform aims at mapping and producing a critical edition of the extensive correspondence of 400 European intellectuals with Jacob Burckhardt over a period of more than half a century from 1842 to 1897.



- Resources of interest:  
Person, Places, Works of art
- Named entities source:  
Freebase
- Missing resources have been  
added to Freebase

# TIMELINE



<http://metasound.dibet.univpm.it/timelinejs/examples/bur-bode.html>

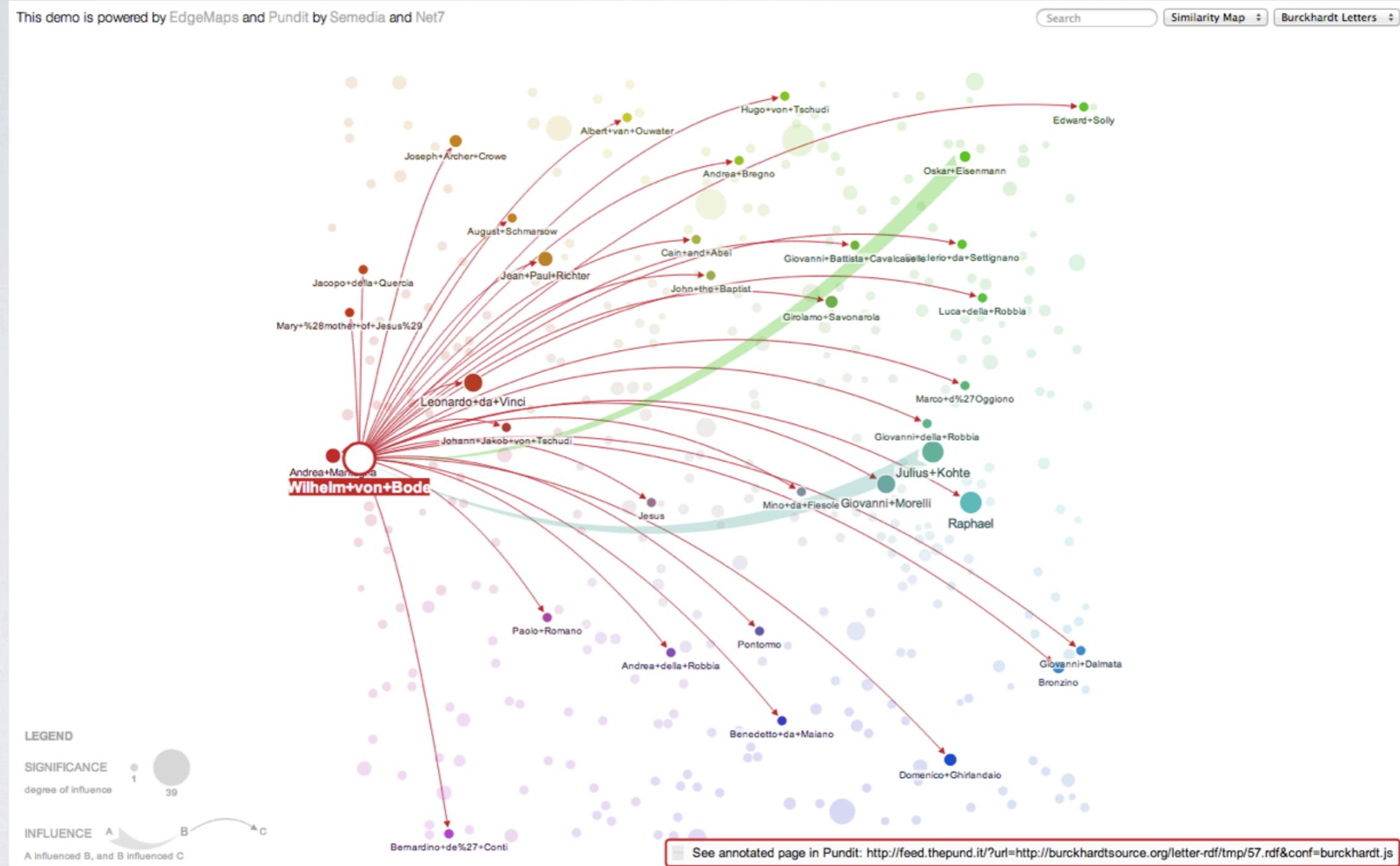
# EDGEMAPS VISUALIZATION

This demo is powered by EdgeMaps and Pundit by Semedia and Net7

Search

Similarity Map

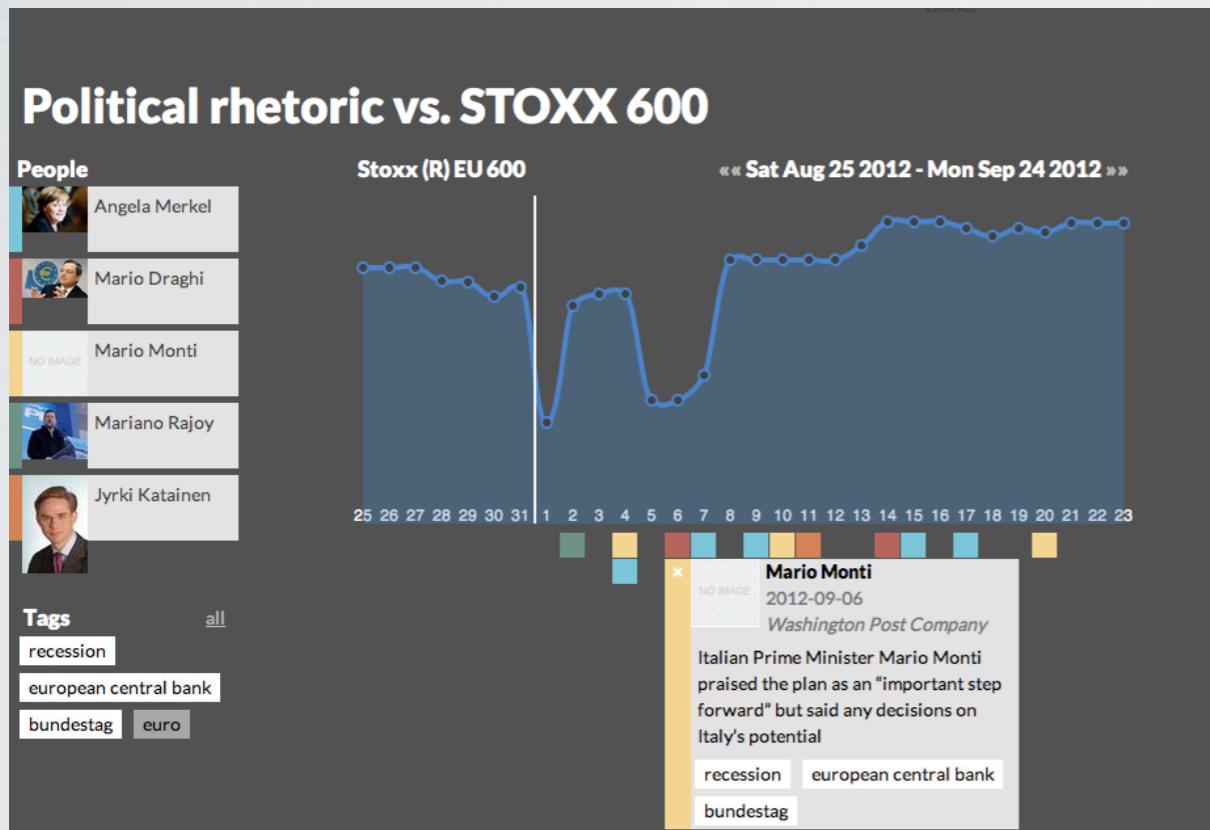
Burckhardt Letters



[http://metasound.dibet.univpm.it/edgemaps/maps/test.html#letters;map;;](http://metasound.dibet.univpm.it/edgemaps/maps/test.html#letters;map;)

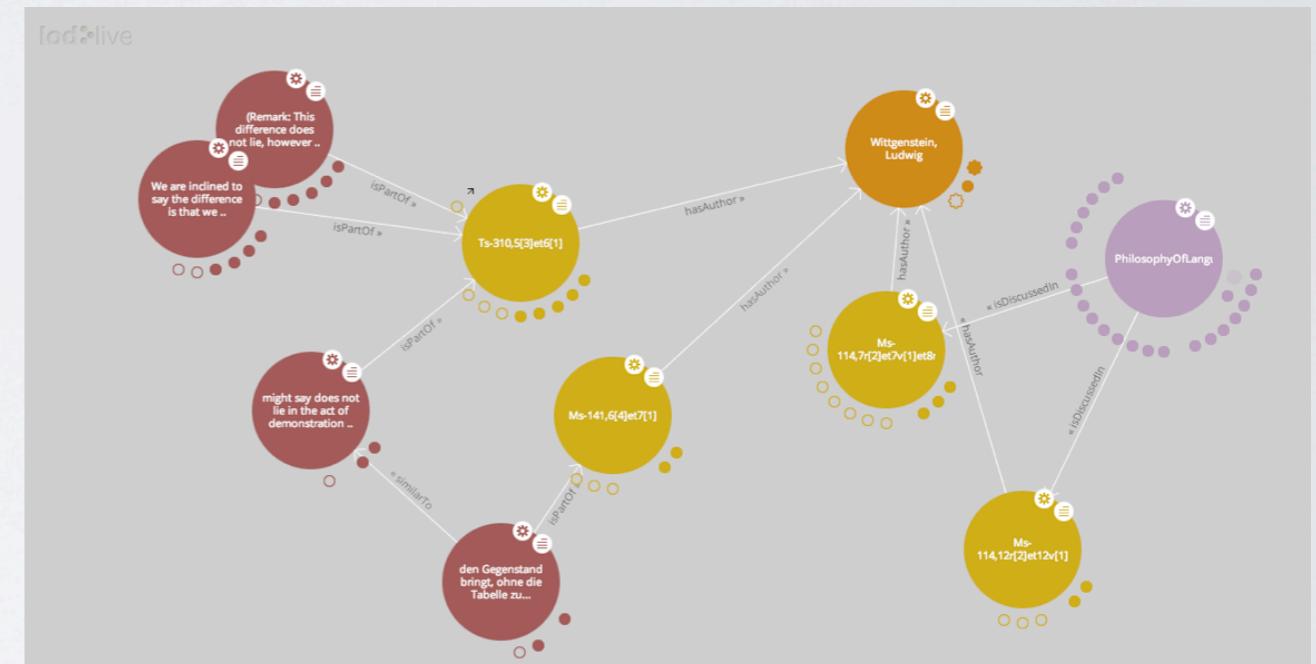
# MORE ON <http://thepund.it>

## DATA JOURNALISM



<http://ask.thepund.it/#/timeline/31951d93>

## LOD LIVE



# CONCLUSIONS

- Pundit is a **customizable** and **flexible semantic web annotation tool**.
- **Custom instances** can be created for specific use scenarios...
- ...and easily **deployed to users** (bookmarklet or browser extension).
- Specific data visualization can be created also using **third party applications**.

# THANK YOU!

**Pundit** 

<http://thepund.it>



**Semedia**  
SEMANTIC WEB AND MULTIMEDIA

<http://semmedia.dii.univpm.it>



[www.netseven.it/](http://www.netseven.it/)



SCUOLA  
NORMALE  
SUPERIORE

[www.sns.it/](http://www.sns.it/)



Seplib Project Eu Project

<http://www.semilibproject.eu/>



DM2E EU Project

<http://dm2e.edu/>



AGORA EU Project

<http://project-agora.eu/>