

# Nepomuk

## The Social Semantic Desktop

Riccardo Chiumiento

21 ottobre 2010

# Struttura del talk:

## Top-down

- ▶ Web semantico e descrizioni formali basate su ontologie
  - ▶ Esempi
- ▶ Progetto Nepomuk
- ▶ Applicazioni che usano il framework sviluppato nel progetto Nepomuk
- ▶ Classi e metodi per un'applicazione di base
- ▶ Applicazione demo

# Web semantico

Da Wikipedia

Semantic Web is a group of methods and technologies to allow machines to understand the meaning of information on the World Wide Web.

# Semantica

un passo indietro

La semantica è quella parte della linguistica che studia il *significato* delle parole, delle relazioni fra parole, frasi e testi.

# Esempio di informazione non strutturata

*Una persona ha un'età e uno stato civile.*

Questa rappresentazione è sintetica perché ogni tipo di semantica non è esplicito, è sott'inteso. Queste informazioni sono (ancora) molto difficili da estrarre algorithmicamente.

## Da Wikipedia

Semantic Web is a group of methods and technologies to allow machines to understand the meaning - or semantics - of information on the World Wide Web.

## Esempio di informazione strutturata

```
<rdfs:Class rdf:ID="Person">
  <rdfs:comment>The class of people.</rdfs:comment>
  <rdfs:subClassOf rdf:resource="http://www.w3.org/2000/03/
</rdfs:Class>
<rdf:Property ID="maritalStatus">
  <rdfs:range rdf:resource="#MaritalStatus"/>
  <rdfs:domain rdf:resource="#Person"/>
</rdf:Property>
<rdf:Property ID="age">
  <rdfs:range rdf:resource="http://www.w3.org/2000/03/examp
  <rdfs:domain rdf:resource="#Person"/>
</rdf:Property>
<rdfs:Class rdf:ID="MaritalStatus"/>
<MaritalStatus rdf:ID="Married"/>
<MaritalStatus rdf:ID="Divorced"/>
<MaritalStatus rdf:ID="Single"/>
</rdf:RDF>
```

## Demenza artificiale?

Questo tipo di metadati strutturati sono tra l'altro una delle basi dell'attuale ricerca sull'intelligenza artificiale.

Ma è possibile racchiudere *tutti* i concetti di questo mondo tramite ontologie, tag, e relazioni fra di essi?

Eppure anche gli esseri umani non hanno nella loro mente un'ontologia che organizza *tutto*, ma piuttosto insiemi di concetti relazionati tra di loro.

Se leggete smartphone comunque, capite subito di che si tratta, senza che qualcuno debba spiegarvi com'è relazionato questo concetto con altri. Ma questo perché nel cervello c'è un tag che indica semanticamente cos'è uno smartphone -legando una serie di concetti atomici tramite sinapsi-.

# Mark David Chapman?

*David Chapman*: chi è costui? Probabilmente per molti di voi il nome non è così nuovo...

# Jigen Daisuke



# Jigen Daisuke

Ops, forse avevo forzato il legame col tag sbagliato :D



# Applicazioni informatiche basate su ontologie (1)



When is Albert Einstein born?



Input interpretation:

Albert Einstein date of birth

Result:

Friday, March 14, 1879

Date formats:

14/03/1879 (day/month/year)


[More formats/calendars](#)

# Applicazioni informatiche basate su ontologie (2)



pictures developed with coffee

About 11,500,000 results (0.12 seconds)

 Everything

More

Show search tools

[Coffee Graphics | Photos Developed With Coffee](#) ☆

**Photos Developed With Coffee**, Custom Coffee Art Using Coffee As Developing Agent  
Silver Salts As Discovered In Costa Rica In 1989.

[www.costaricacoffeeart.com/coffee\\_as\\_developer\\_for\\_silver\\_salts.php](http://www.costaricacoffeeart.com/coffee_as_developer_for_silver_salts.php) - Cached - Sim

[Coffee Art](#) ☆

**coffee pictures**. Your Company Can Enter and Capture a New Market ! ...

# Web semantico

Da Wikipedia

While the term Semantic Web is not formally defined it is mainly used to describe the model and technologies proposed by the W3C. These technologies include the Resource Description Framework (RDF), a variety of data interchange formats (e.g. RDF/XML, N3, Turtle, N-Triples), and notations such as RDF Schema (RDFS) and the Web Ontology Language (OWL), all of which are intended to provide a formal description of concepts, terms, and relationships within a given knowledge domain.

# Nepomuk

## The Social Semantic Desktop

Riccardo Chiumiento

21 ottobre 2010

# Nepomuk

[da.nepomuk.semanticdesktop.org](http://da.nepomuk.semanticdesktop.org)

“NEPOMUK (Networked Environment for Personalized, Ontology-based Management of Unified Knowledge) intends to realize and deploy a comprehensive solution - methods, data structures, and a set of tools - for extending the personal computer into a collaborative environment, which improves the state of art in online collaboration and personal data management and augments the intellect of people by providing and organizing information created by single or group efforts.”

# Consorzio

Fra gli altri...

- ▶ IBM
- ▶ HP
- ▶ Mandriva
- ▶ ...

Enti di ricerca:

- ▶ German Research Center for Artificial Intelligence DFKI GmbH
- ▶ Ecole Polytechnique Fédérale de Lausanne
- ▶ Università de la Svizzera Italiana
- ▶ ...

per un totale di 5.6 milioni di euro investiti

L'Unione Europea ha supportato il progetto con un ulteriore investimento di 11.5 milioni di euro.

Per un totale di *17.1 milioni di euro*. (!)

nooo...

Durata del progetto: dal 01/01/2006 al 31/12/2008: tre anni e “*stop*”. É previsto un ulteriore soltanto da un’organizzazione creata ad hoc, OSCAF (Open Semantic Collaboration Architecture Foundation), che ha la responsabilità sul mantenimento e sulla disponibilità delle ontologie sviluppate nel progetto Nepomuk.  
ma...

## Use case: comunità Mandriva

SWIM: Semantic Web Issue Manager.

“A typical Linuxs user or developer has to deal with a huge amount of information in order to get the most of his system and the project he’s interested in.”

Questionario finale agli utilizzatori: forte successo.

# Dal web semantico al desktop semantico

Utilizza tutte le tecnologie usate dal web semantico: ontologie, SPARQL per interrogare grafi RDF...  
Associa ogni file col proprio URI.

# Ontologie

Quattro ontologie differenti:

**NRL** rappresentazione della conoscenza, generica

**NAO** per annotazioni

**NIE** file: documenti, immagini, contatti...

**PIMO** informazioni personali

# Vision sull'ontologia PIMO (1)

PIMO è basato sull'ontologia NRL. Ogni concetto nel modello viene chiamato *Thing*. Esempio: *AGENT*

**Superclasses** pimo:ClassOrThing,  
pimo:ClassOrThingOrPropertyOrAssociation,  
rdfs:Resource, pimo:Thing

**Subclasses** pimo:Organization, pimo:Person, pimo:PersonGroup

**In domain of** pimo:createdPimo, pimo:isOrganizationMemberOf

**In range of** pimo:creator, pimo:hasOrganizationMember

**Description** An agent (eg. person, group, software or physical artifact). The Agent class is the class of agents; things that do stuff. A well known sub-class is Person, representing people. Other kinds of agents include Organization and Group. (inspired by FOAF). Agent is not a subclass of NAO:Party.

## Vision sull'ontologia PIMO (2)

Esempio: *CITY*

**Superclasses** pimo:ClassOrThing,  
pimo:ClassOrThingOrPropertyOrAssociation,  
pimo:Location, rdfs:Resource, geo:SpatialThing,  
pimo:Thing

**Subclasses** –

**In domain of** –

**In range of** –

**Description** A large and densely populated urban area; may include several independent administrative districts; Ancient Troy was a great city.

# GENERAL DISCLAIMER

!!!

Le stesse tecnologie usate nel web semantico al momento sono piuttosto sperimentali. Per questo l'inclusione in un sistema orientato all'utente finale come KDE *deve* essere limitata soltanto a quelle tecnologie abbastanza affidabili.

!!!

# Aperture

Aperture non è direttamente legato a Nepomuk, ma può essere usato e viene usato ampiamente per dare in pasto degli RDF a Nepomuk. “Aperture is a Java framework for extracting and querying full-text content and metadata from various information systems.”

Esempi di applicazioni che usano Aperture

[deliciouscrawler](#) crawls delicious accounts, extracts bookmarks and tags

[flickrcrawler](#) crawls flickr accounts, extracts tags, and photos metadata

[imapcrawler](#) crawls remote mailboxes accessible with IMAP

[webcrawler](#) crawls websites

# Strigi

“Strigi is a daemon which uses a very fast (?) and efficient crawler that can index data on your harddrive.”

ogni utente KDE nel bene e nel male ci ha avuto a che fare...

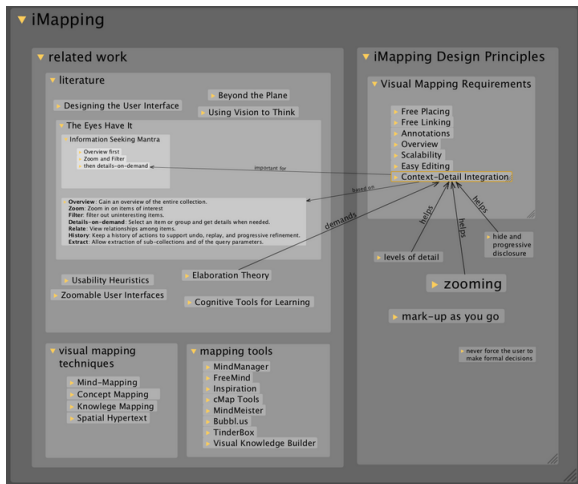
# IBM Galaxy

“IBM Galaxy performs semantic analysis of natural language texts and provide recommendations.”

Viene usato dal componente Comp-StrucRec di Nepomuk.

# iMapping

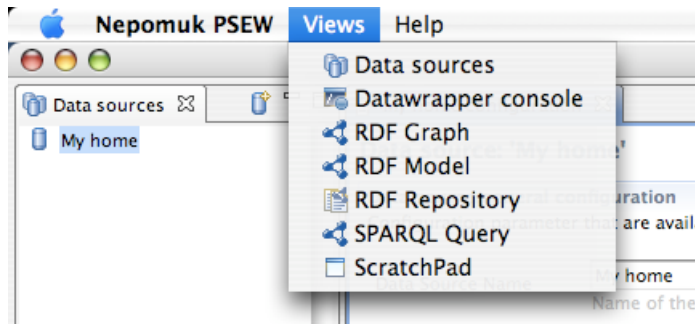
Un approccio grafico per strutturare i concetti



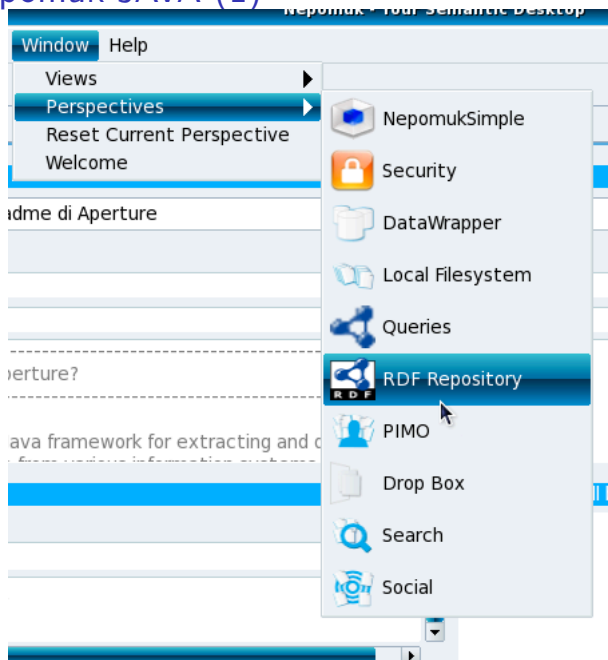
*iMapping si auto-illustra*  
Equivalente CL: Quickey.

# PSEW

P2P Semantic Eclipse Workbench



# Nepomuk JAVA (1)



# Nepomuk JAVA (2)

The screenshot displays a Linux desktop environment with a dark blue background. On the left, there is a sidebar with icons for a folder, a star, and a trash can, labeled "pythonext-2.5.2.20080611", "Benvenuto", and "Cestino" respectively. A terminal window is open in the foreground, showing a shell prompt and some output.

The main window is titled "Nepomuk" and has a menu bar with "File", "Actions", "Window", and "Help". The window title bar shows "Drop Box" and "DropBox X". The main content area displays the following information:

**You dropped this file:**

Name	Soltanto il readme di Aperture
Title	README.txt
Type	text/plain
Snippet	[0. What is Aperture? Aperture is a Java framework for extracting and querying...

**Select multiple tags**

Search for more tags

readme

readme(new tag)

# Semnotes

The screenshot shows the SemNotes application window. The title bar reads "SemNotes - Semantic Notes". The menu bar includes "File", "Game", "Edit", "Move", "View", "Go", "Bookmarks", "Tools", "Settings", and "Help". The toolbar contains icons for "New", "Open", "Save", "Print", "Help", "Relations", "Linking", and "Show Main Window".

The main content area displays a list of notes:

- A note with a green arrow icon, a star rating of five stars, and the text "modified 0 second ago".
- A note titled "Pennello rosso" with a green arrow icon, the text "rosso", a star rating of five stars, and the text "modified 1 hour ago".
- A note titled "Penna" with a green arrow icon, a star rating of five stars, and the text "created 1 hour ago".

The "Penna" note is expanded to show its content:

- Text: "rosso"
- Text: "prova 2"

On the left side of the application, there is a vertical sidebar with a search bar and a list of items, including "a Vis", "700)", "alho", "ord", "stin", "w3.o", "/ont", "'pr", "v5 <", "3.or", "score (?v6) . } . } . }", and "prova") (6, "2"))".

# Estensione Firefox

essione Visualizza

o su open

USE

rogetto openSU  
enSUSE crea un  
sparente ed ami  
urce

rogetto è contro  
me tester, redatt  
braccia una vast  
ferenti lingue ed  
vertiti...

**Nepomuk comment manager for Firefox**

Comment Management for Web pages

commento|

Save/Update Remove Exit

Rate: ☆☆☆☆☆

# Nepomuk

## The Social Semantic Desktop

Riccardo Chiumiento

21 ottobre 2010

## CMakeLists.txt

```
find_package(KDE4 REQUIRED)
find_package(Nepomuk REQUIRED)
target_link_libraries(nomeesequibile ${NEPOMUK_LIBRARIES} )
```

# ResourceManager

*ResourceManager* è implementata come una sottoclasse figlia di *QObject*. Tramite i metodi forniti da questa classe è possibile visualizzare e accedere alle risorse di Nepomuk.

PS. `#include <Nepomuk/ResourceManager>`

Un metodo a caso:

```
QList<Resource> Nepomuk::ResourceManager::allResources ( )
```

*Retrieve a list of all resource managed by this manager.*

## Metodo init():

```
int Nepomuk::ResourceManager::init ()
```

*This method will initialize the communication with the local Nepomuk-KDE services, ie. the data repository. It will trigger a reconnect to the Nepomuk database. NOTA: questo metodo è utilizzato da molti altri e in genere è consigliato non usarlo esplicitamente, ma è pubblico e pertanto può essere richiamato anche manualmente: lo useremo :D.*

Nel nostro caso semplifica la vita.

# Resource

Gli oggetti della classe *Resource* sono “quelli centrali”, quelli con cui generalmente lavora Nepomuk. In effetti, la classe *File* (del namespace Nepomuk!), la classe *Tag*, la classe *Thing* usata ampiamente da PIMO, non sono nient'altro che specializzazioni della classe *Resource*!

NOTA: non è necessario includere altri file header.

## Namespace Query

Il namespace `Nepomuk::Query` è stato aggiunto in KDE 4.4 ed è l'unico modo fornito dal framework per interrogare Nepomuk-KDE senza conoscere l'*SPARQL*, il linguaggio specifico per le query di RDF.

```
Nepomuk::Tag myTag = getOurFancyTag();
QString query
    = QString("select distinct ?r where { ?r %1 %2 . }")
      .arg( Soprano::Node::resourceToN3(Soprano::Vocabulary::NAME))
      .arg( Soprano::Node::resourceToN3(myTag.resourceUri()));
```

diventa. Notare: query SPARQL trasparente.

```
Nepomuk::Query::ResourceTerm tagTerm(myTag);
Nepomuk::Query::ComparisonTerm term(Soprano::Vocabulary::NAME);
Nepomuk::Query::Query query(term);
QString queryString = query.toSparqlString();
```

# Tag

```
#include <Nepomuk/Tag>
```

Tag è una classe derivata di Resource. Questa classe mette a disposizione pochi metodi, nel nostro caso basta il costruttore.