Gestione semantica ed adattativa per la fruizione dei Contenuti Culturali nel Progetto LC3

Beniamino Di Martino
Seconda Università di Napoli
Beniamino.dimartino@unina.it

Il Progetto LC3

- Laboratorio Publico Privato *per la Comunicazione delle Conoscenze Culturali in Rete*
- Finanziato dal MUR (azione FAR) insieme ad altri 19 nelle regioni Ob. 1
- Partners:
  - Seconda Università di Napoli
  - Università della Calabria
  - Università di Sassari
  - Atlantis s.p.a
  - Space s.p.a.
  - Unidata s.p.a.
- SubContractors:
  - Istituto Luigi Sturzo
  - Università Sapienza di Roma (UNIMED)
  - Università “Federico II” di Napoli
- Sito: http://lc3.spacespa.it

Obiettivi del Progetto

- definizione di una **metodologia** e di soluzioni per la gestione dei metadati analitici MAG collegati alle collezioni digitali, con definizione di estensioni applicative e semantiche;
- definizione di una **architettura di gestione di media digitali** basata su *contlets*
- definizione di un **modello di riferimento di annotazioni semantiche** e di soluzioni versatili per la generazione automatica o assistita di **ontologie di dominio**;
- definizione di soluzioni e per supportare il lavoro di **annotazione semantica** delle fonti documentali e l'**information retrieval semanticamente arricchito** sulle stesse;
- creazione di un **ambiente integrato di storytelling** che consenta l'impostazione di *outlines* narrative, il retrieval semanticamente assistito di *contlets* e l'associazione dei *contlets* ai supporti esplicativi in sequenze narrative composite;
- definizione di una metodologia e di soluzioni tecnologiche per la **produzione assistita di pubblicazioni** multilivello su web basate su XML e sulla gestione di profili degli utenti;
- definizione di **metodologie di profilazione degli utenti** per l'erogazione personalizzata dell'informazione;
- Definizione di metodologie e soluzioni di **geocoding** delle risorse documentali, orientate a criteri complessi quali attribuzioni areali, gradienti di pertinenza, contestualizzazioni condizionali e *bundles* di contestualizzazione con altre risorse affini;
- definizione di soluzioni, anche integrabili in piattaforme software esistenti, per la **contestualizzazione geografica** delle memorie culturali in rappresentazioni paesaggistiche 3D interattive;
- Analisi e ricerca delle soluzioni di **post processing e gestione dei contenuti** digitali multimediali;
- Analisi e ricerca di soluzioni per il riconoscimento automatico dei **protocolli necessari al delivery multidevice**;

Attività dell' Unità di Ricerca della SUN

- **Definizione di un modello uniforme per la creazione di annotazioni semantiche**
- **Derivazione Automatica e Costruzione Assistita di Ontologie di Dominio da corpora documentali**
- **Sistemi di Annotazione Semantica di Contenuti**
- **Tecniche e Strumenti per l'Information Retrieval Semantico**
- **Profilazione e riconfigurazione di terminali eterogenei mediante tecnologie ad agenti mobili**

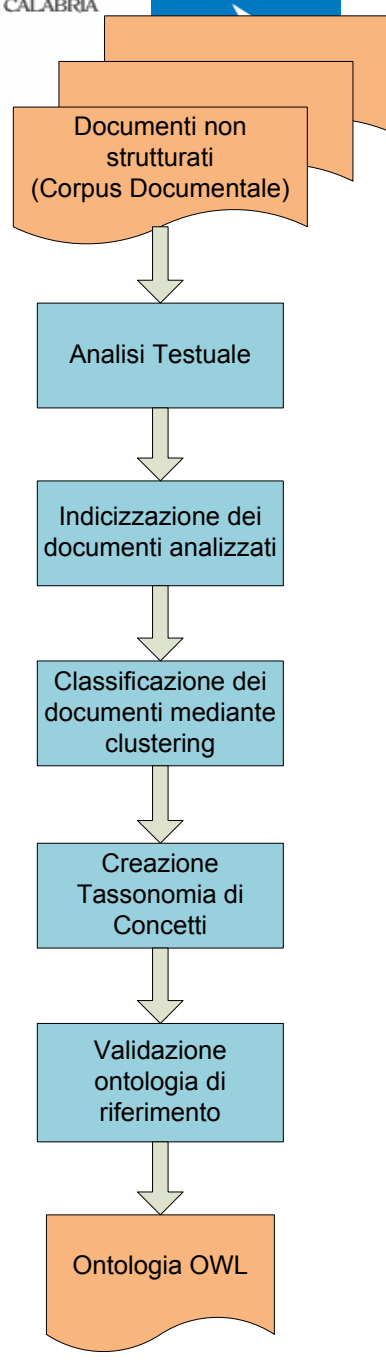
Costruzione Automatica o Assistita e Popolazione di Ontologie di Dominio

- Stato dell'Arte dei modelli e dei linguaggi standard per la rappresentazione di ontologie nel Semantic Web
- Rassegna ed Analisi delle tecnologie e degli strumenti per la costruzione Computer Assisted di Ontologie
- Rassegna ed Analisi delle metodologie, delle tecniche e degli strumenti per la derivazione automatica di Ontologie
- Definizione di una tecnica per la derivazione automatica di Ontologie da corpora documentali
- Realizzazione di uno Strumento Prototipale per la derivazione automatica di Ontologie da corpora documentali
- Definizione di una tecnica per la derivazione di Ontologie da strutture di documenti gerarchicamente o relazionalmente organizzate
- Realizzazione di uno strumento Prototipale per la derivazione di Ontologie da strutture di documenti gerarchicamente o relazionalmente organizzate
- Costruzione automatica di una ontologia in formato OWL sul dominio "Il rapimento di Aldo Moro"
- Realizzazione manuale di Ontologie di Dominio e relative Annotazioni Semantiche
- Definizione di un MAG Repository per lo storage sicuro dei contenuti multimediali granulari, in grado di gestire contenuti semanticamente annotati

Definizione di una tecnica per la derivazione automatica di Ontologie da corpora documentali

Fasi elaborative:

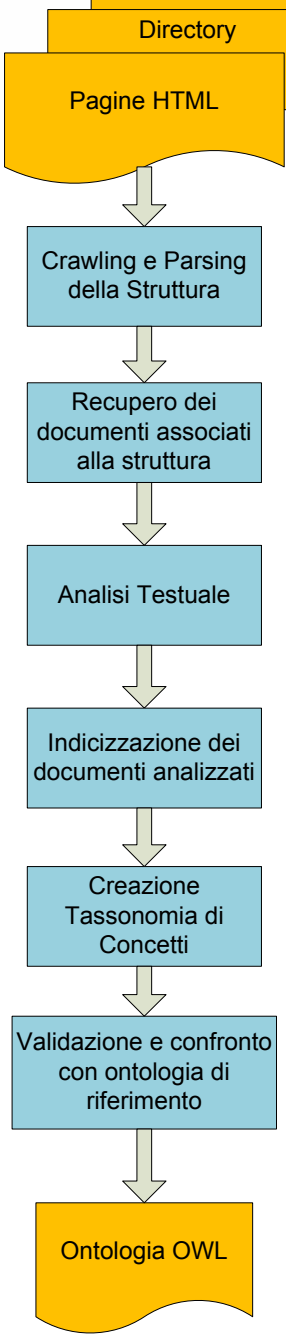
1. Analisi testuale dei documenti di un dato dominio
2. Indicizzazione dei documenti
3. Clustering dei documenti testuali indicizzati
4. Individuazione di una tassonomia di concetti e relazioni tra quest'ultimi
5. Validazione della tassonomia di concetti mediante confronto con ontologia di riferimento
6. Creazione ed esportazione di un'ontologia in linguaggio RDF/OWL



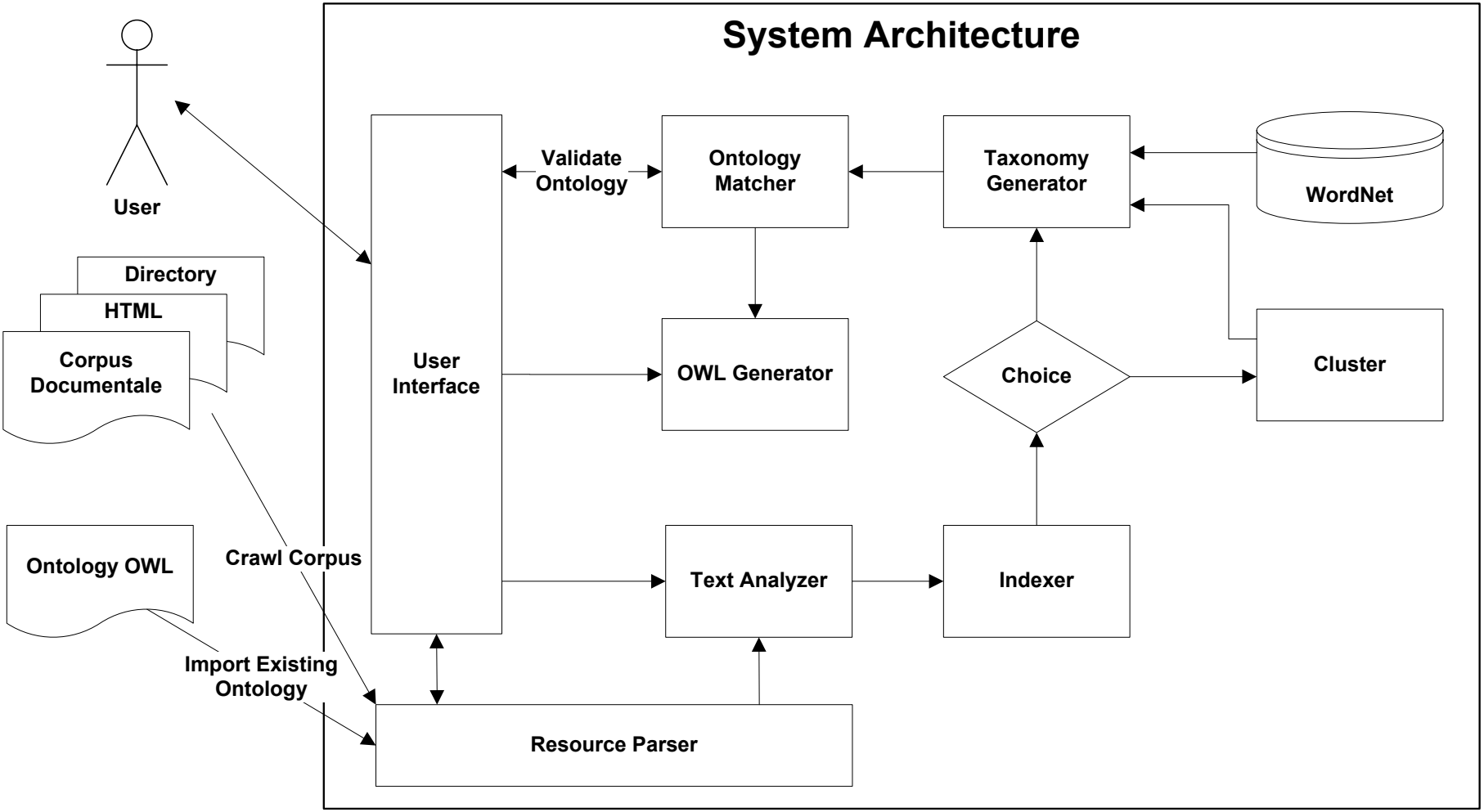
Definizione di una tecnica per la derivazione di Ontologie da strutture di documenti gerarchicamente o relazionalmente organizzate

Fasi elaborative:

1. Crawling e Parsing della struttura in ingresso
2. Recupero dei documenti associati alla struttura
3. Analisi testuale dei documenti associati alla struttura
4. Indicizzazione dei documenti
5. Individuazione di una tassonomia di concetti e relazioni tra quest'ultimi
6. Validazione della tassonomia di concetti mediante confronto con ontologia di riferimento
7. Creazione ed esportazione di un'ontologia in linguaggio RDF/OWL



Architettura del Sistema



Componenti

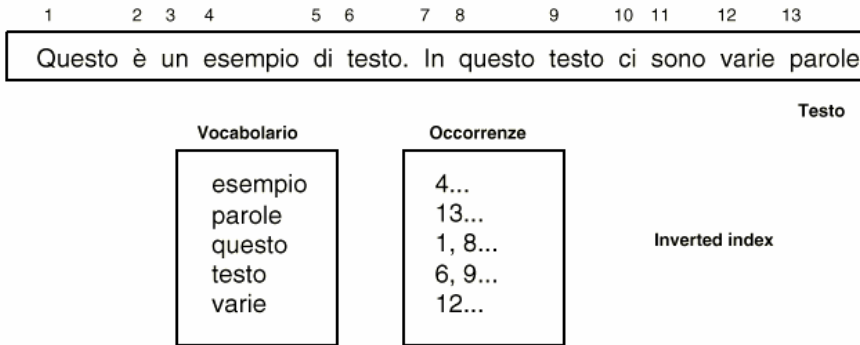
Text Analyzer

Pre-elaborazione del testo:

- Analisi lessicale
- Stemming

Indexer

- Inverted index: ad ogni documento viene associato un vettore di caratteristiche



Componenti

Clustering

Matrice di similarità o dissimilarità

- distanza Euclidea o Prodotto Scalare

Clustering

- Neighbour Joining

Creazione di un albero di cluster gerarchico

- gerarchia di documenti

Taxonomy Generator

Individuazione di una tassonomia di termini

Percolazione dell'albero di termini

- Conversione vettori di caratteristiche in bitSet
- Operazioni logiche di AND e di OR sui nodi

Associazione dei concetti tramite WordNet

Componenti

Ontology Matcher:

Input:

- tassonomia di concetti, legati tra loro da relazioni di sussunzione
- ontologia OWL di riferimento

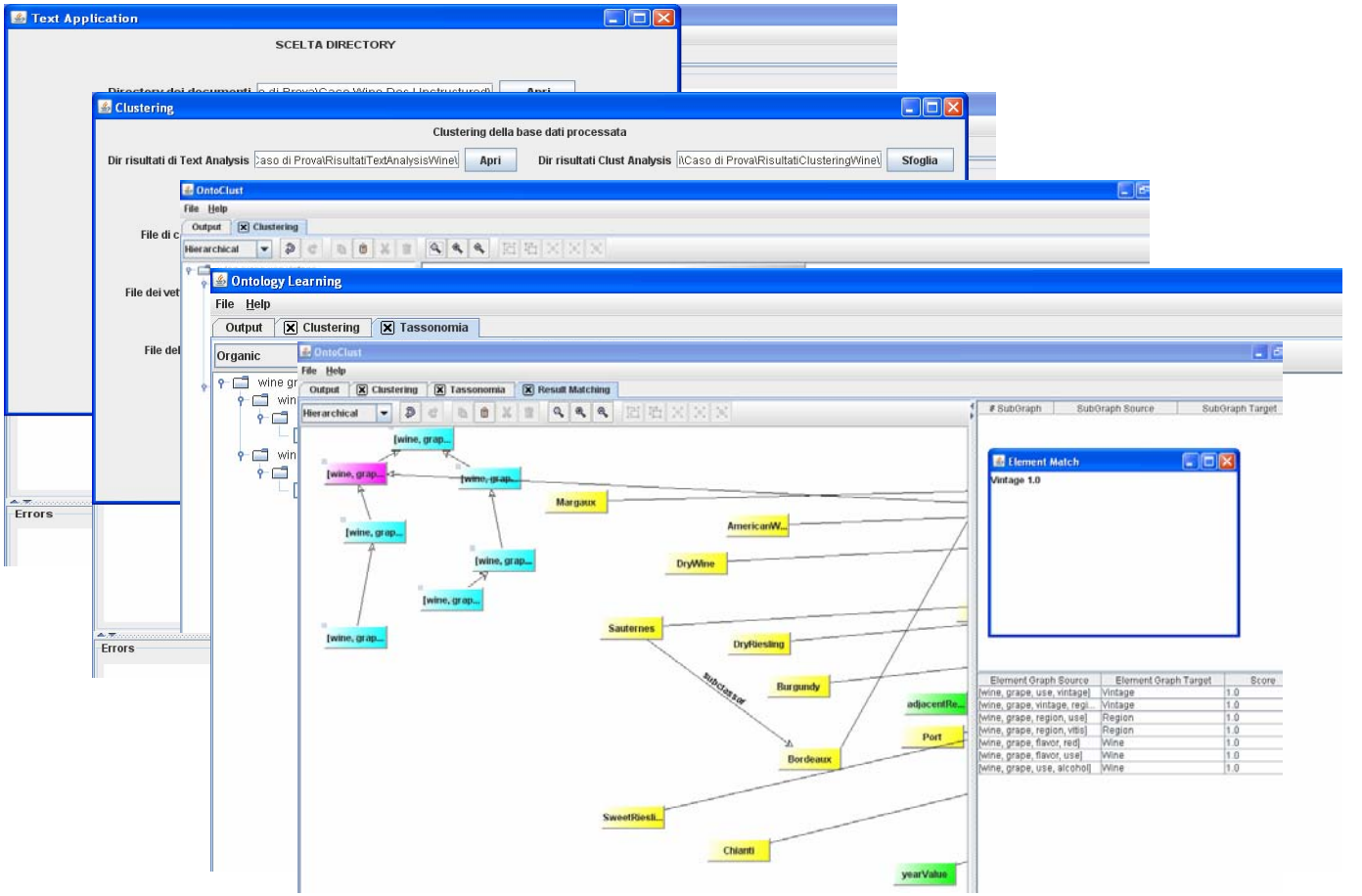
Match sintattico-semantico tra la tassonomia generata e l'ontologia di riferimento per validare la tassonomia

OWL Generator:

Esportazione in OWL

- Utilizzo di unparser (framework Jena)

Alcune Snapshot del Prototipo



Realizzazione di Ontologie di dominio

- Caso Moro
  - Contesto Politologico-Storico
- Progetto Michael
  - Progetto Digitalizzazione Dati
- Costituente
- MUSA
  - Museo Storico-Archeologico
- Palinsesti Televisivi e Web-TV
- Giurisprudenza
  - Atti Notarili

Caso Moro

CasoMoro_modif Protégé 3.3.1 (file:\C:\Documents%20and%20Settings\divby0\Desktop\LC3OWL\CasoMoro_modif.pprj, OWL / RDF Files)

File Edit Project OWL Code Tools Window Help

Metadata (CasoMoro.owl) | OWLClasses | Properties | Individuals | Forms

CLASS BROWSER

For Project: CasoMoro_modif

Class Hierarchy

- owl:Thing
  - Axiom_1
  - Caso_Moro
    - Biografia_di_Aldo_Moro
    - Il_contesto_storico-politico-sociale
      - Anni_di_piombo
        - Gruppi_estremisti
        - Strategia_della_tensione
          - Stragi_italiane (6)
        - Terrorismo
      - Situazione_economica (3)
      - Situazione_internazionale
      - Situazione_politica
      - Situazione_Sociale
    - Il_dopo_Moro
    - Il_sequestro
    - Le_date (99)
    - Nel_78 (1)
    - Persone_e_istituzioni
    - Professione (22)

INSTANCE BROWSER

For Class: Stragi_italiane

Asserted | Inferred

Asserted Instances

- Gioia_Tauro
- Piazza_della_Loggia
- Piazza_Fontana
- Portella_della_Ginestra
- Stazione_di_Bologna
- Treno_Italicus

CasoMoro_modif Protégé 3.3.1 (file:\C:\Documents%20and%20Settings\divby0\Desktop\LC3OWL\CasoMoro_m

File Edit Project OWL Code Tools Window Help

Metadata (CasoMoro.owl) | OWLClasses | Properties | Individuals | Forms

PROPERTY BROWSER

For Project: CasoMoro_modif

Object | Datatype | Annotation | All

Object properties

- hanno_fatto_parte_del_governo
- ha_sostenuto ↔ è_sostenuta_da
- è_sostenuta_da ↔ ha_sostenuto
- ha_titolo_di_studio
- è_del_partito ↔ ha_appartenenti
- ha_appartenenti ↔ è_del_partito
- ha_per_autore ↔ è_autore_di
- è_autore_di ↔ ha_per_autore
- è_stato_fatto ↔ ha_fatto
- ha_fatto ↔ è_stato_fatto
- è_destinatario_della_lettera ↔ indirizzata_a
- indirizzata_a ↔ è_destinatario_della_lettera
- Si_allontana_da
- si_ritrova_nel ↔ Parla_di
- Parla_di ↔ si_ritrova_nel
- è_stato_ucciso
- ha_editore ↔ è_editore_di
- è_editore_di ↔ ha_editore
- è_stato_trovato
- Ha_avuto_luogo_in ↔ è_il_luogo_del
- è_il_luogo_del ↔ Ha_avuto_luogo_in
- è_giornale_di_partito

PROPERTY EDITOR

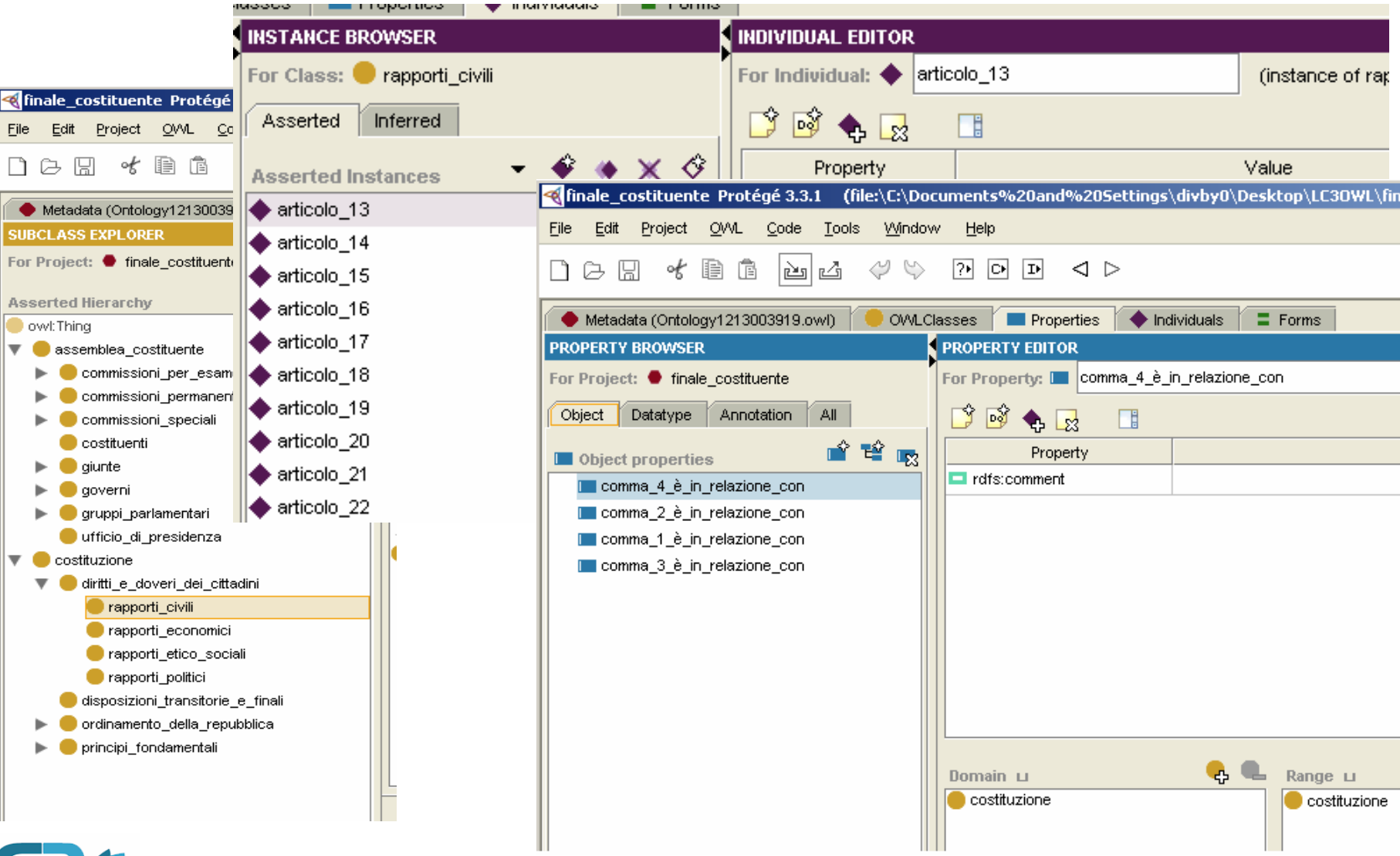
For Property: hanno_fatto_parte_del_governo

Property	
rdfs:comment	

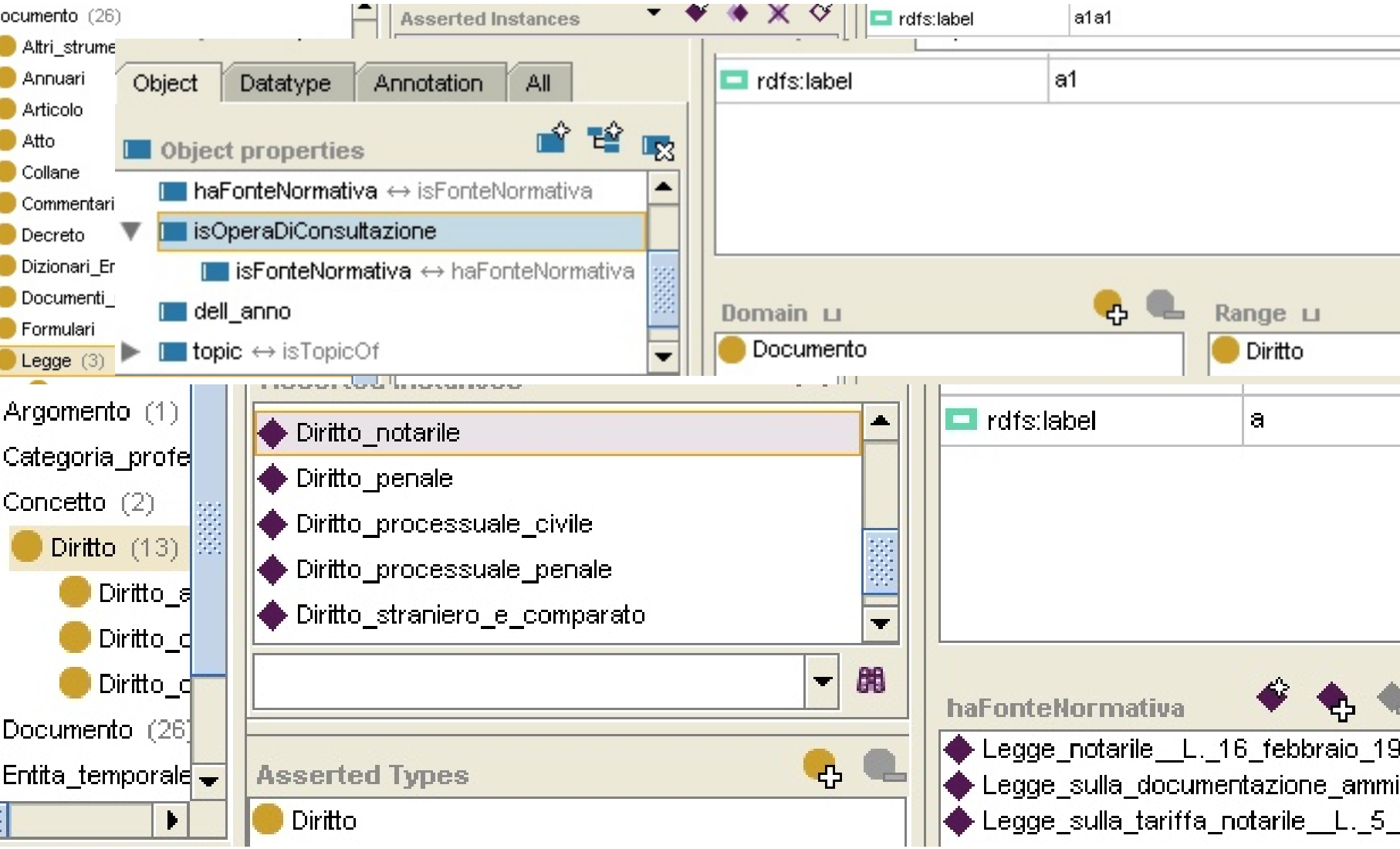
Domain: Ministri

Range: Governo_ed_elezioni

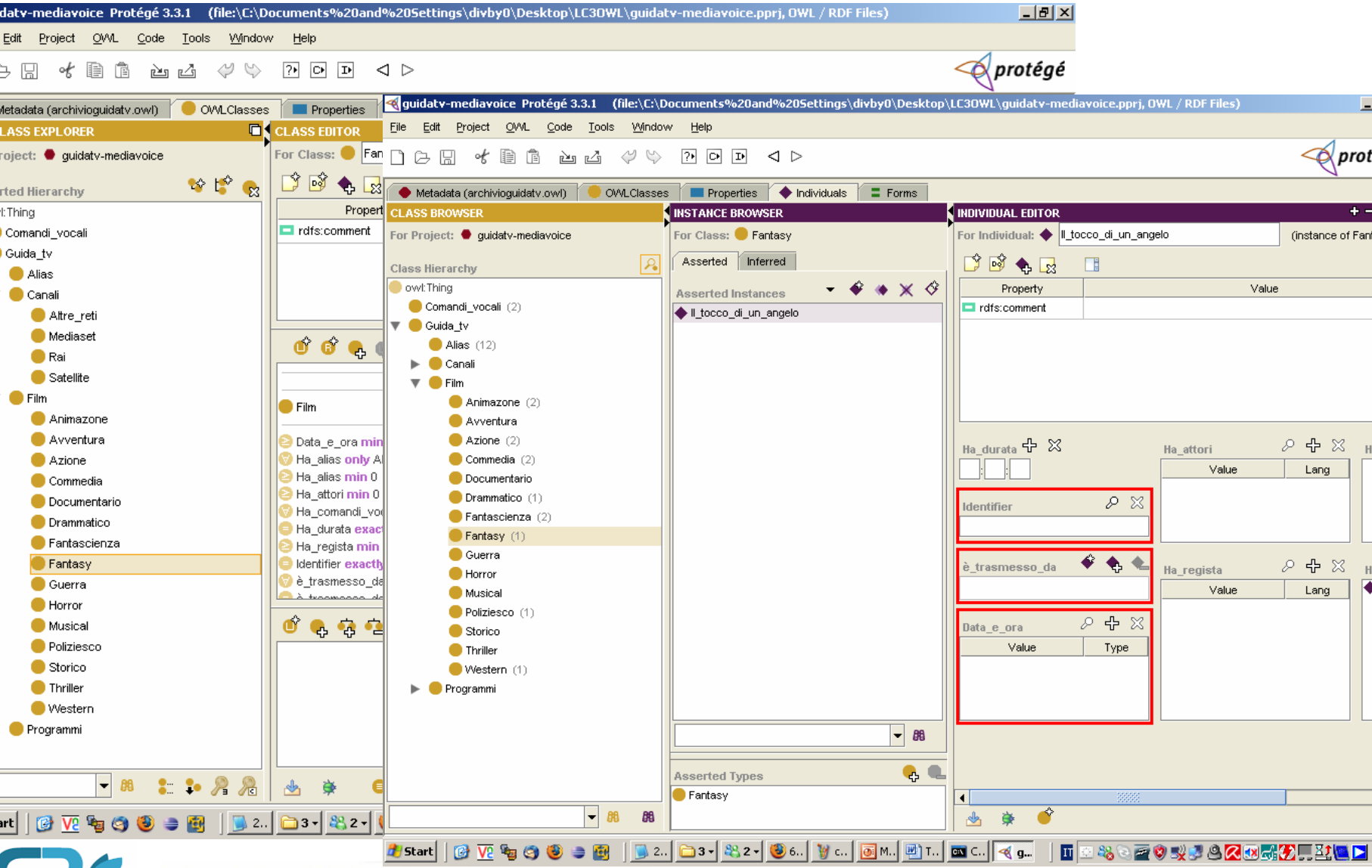
Costituente



Dominio Giuridico



Web-TV



Laboratorio pubblico-privato di ricerca sul tema
della Comunicazione delle Conoscenze Culturali

MUSA – Museo Storico Archeologico

agro Protégé 3.3.1 (file:\C:\Documents%20and%20Settings\divby0\Desktop\LC3OWL\agro.pprj, OWL / RDF Files)

File Edit Project OWL Code Tools Window Help

Metadata (Ontology1201087121.owl) | OWLClasses | Properties | Individuals | Forms

CLASS BROWSER

For Project: agro

Class Hierarchy

- owl:Thing
  - Cronologia
  - Materia_reperti (12)
  - Raffigurazioni
  - Reperti_archeologici
    - Abbigliamento
      - Accessorio
      - Armatura
        - Elmo (1)
      - Gioielli
    - Oggetti_vari
    - Oggetto_artistico
    - Vasellame
  - Stato_conservazione_reperti (4)
  - Tipologia_reperti

INSTANCE BROWSER

For Class: Elmo

Asserted | Inferred

Asserted Instances

- Reperto_n._8498

Asserted Types

- Elmo

INDIVIDUAL EDITOR

For Individual: Reperto_n._8498 (instance of Elmo)

Property	Value
rdfs:comment	

Altezza_m.

Value	Type
0.16	decimal

Note_alla_datazione

Value	Lang

Cat_De_Franciscis

Value	Lang

Numero_negativo_fc

Value	Type

Codice_a_matita

Value	Lang

Numero_oggetti_rafl

Value	Type

Start 11.57

Sistemi di Annotazione Semantica

- Stato dell'arte dei modelli ed Analisi di Sistemi di annotazione basati su ontologie
- Installazione, analisi e confronto funzionale dei sistemi di annotazione basati su ontologie
- Realizzazione di un'architettura per l' annotazione semantica e di uno strumento Prototipale
- Stato dell' arte ed Analisi delle metodologie e delle tecniche per la Classificazione ed il Clustering
- Realizzazione di uno strumento prototipale per l' annotazione semantica automatica di corpora documentali basato su tecniche di classificazione automatica
- Architettura generale e intercambio dei dati semantici tra il MAG Repository e lo strumento di annotazione semantica

Modelli di Annotazione

- Rispetto a dove vengono memorizzate
  - Embedded, esterne
- Rispetto ai linguaggi utilizzati
  - In genere dialetti di RDF e OWL
- Rispetto ai *link* agli elementi delle ontologie
  - Istanze, Concetti etc.
- Rispetto ai tipi di ontologie utilizzate per l'annotazione
  - Lightweight, heaveyweight: thesauri,vocabolari controllati, ontologie etc.
- Rispetto ai tipi di documenti annotati
  - Immagini, video, testi etc.

Modelli di Annotazione

- Esistono diversi framework/metodologie di annotazione in letteratura
  - MnM; OntoMat; Kim; SHOE; AeroSWARM etc.
- In genere ognuno supporta un proprio modello di annotazione
  - Spesso embedded
- Utilizzo predominante di OWL e RDF
- Raramente si annotano cose diverse dai documenti testuali
  - E i tool che supportano altri media annotano solo un tipo di documenti
- Nei sistemi di "annotazione" wide-spreaded (es: annotazione video Google, Facebook etc.) praticamente si supporta solo il tagging.
- Ogni Modello supporta o solo l'annotazione tramite istanze, o solo l'annotazione tramite link a concetti/relazioni
- Nessun supporto per l'annotazione di "interi documenti" o di "gruppi di documenti"
- Nasce L'esigenza, all'atto dell'annotazione, quando questa non è automatica, di introdurre la gestione delle "versioni"
  - Con problematiche di merging delle ontologie e delle annotazioni di versioni differenti.

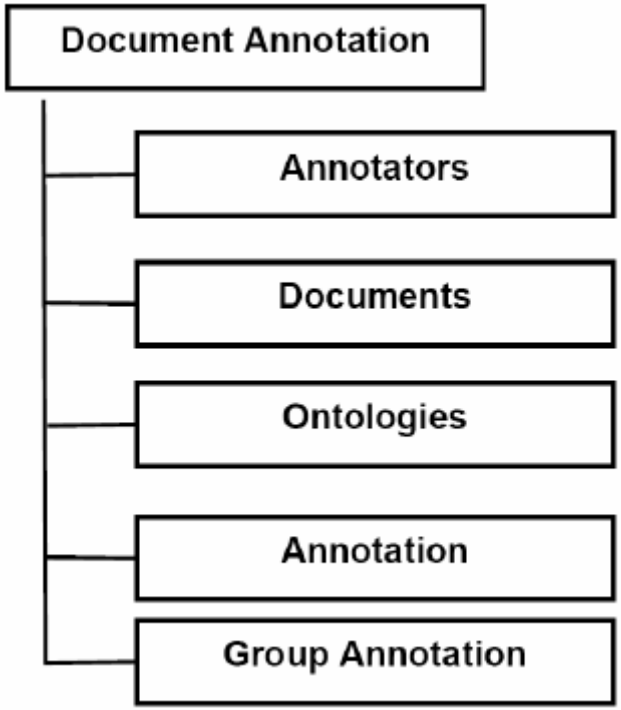
Considerazioni

- Dallo stato dell'arte, si percepisce l'esigenza di un modello di annotazione che supporti:
  - Annotazioni di documenti multimediali (audio, video, html, pdf, rtf, immagini etc.)
  - Gestisca "gruppi" di annotatori nel caso di annotazioni manuali
    - Gestisca l'accesso alle ontologie, ai documenti e alle annotazioni in modo appropriato per un sistema multi-utente
    - Gestisca Versioni di ontologie e annotazioni
  - L'annotazione di:
    - Parti di documenti
    - Documenti nella loro interezza
    - Gruppi di Documenti
    - Versioni di Documenti/parti/gruppi (es: documenti digitalizzati con diverse tecniche, annotazioni di gruppi differenti, annotazioni parziali etc.)
  - Annotazione con riferimento a
    - Concetti/relazioni di 1 Ontologia
    - Concetti/relazioni di Ontologie Multiple
    - Relazioni tra concetti di ontologie differenti scoperte durante il processo di annotazione e non durante quello di definizione dell'ontologia (quando gli utenti non hanno i permessi per creare ontologie)

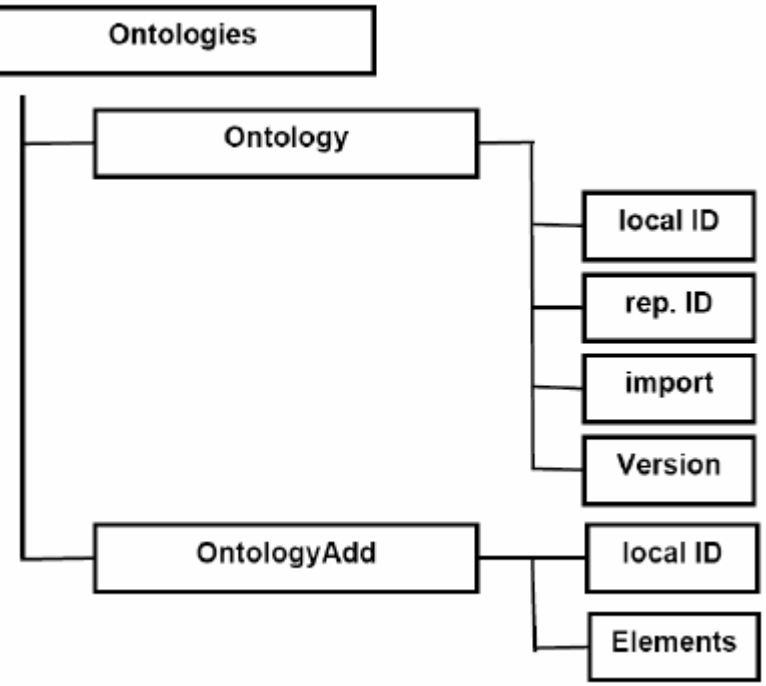
Il Modello di Annotazione definito

- Salva le annotazioni in file esterni per facilitarne l'esportazione
- Supporta annotazione di gruppi di documenti, documenti nella loro interezza e parti di documenti
- Supporta annotazioni di documenti di testo, pdf, rtf, word, immagini, video, audio.
- Supporto per il Versioning delle annotazioni
- Supporto per diversi Repository distribuiti
- Supporto per annotazioni
  - Linked
  - Istanze
  - Tag
  - RDF
  - OWL
  - Possibilità di introdurre un parametro di "rilevanza"
- Possibilità di Costruire Annotazioni complesse definendo nuove relazioni tra concetti di ontologie differenti

Modello Annotazione: Componenti

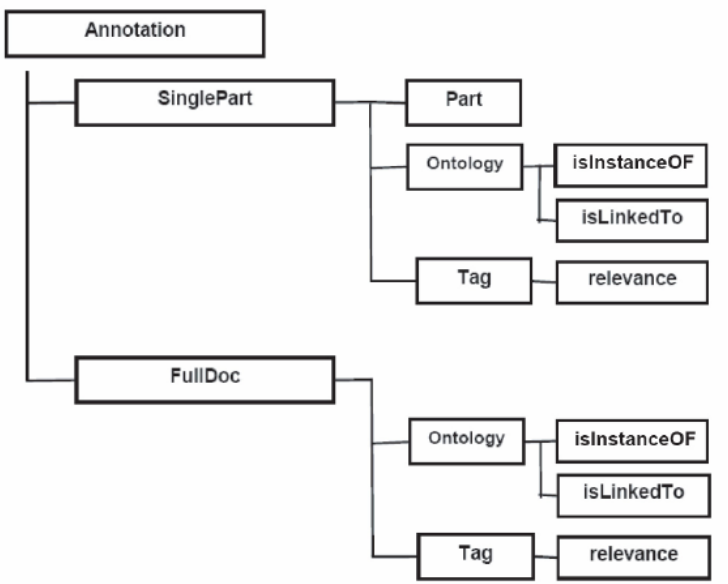


Modello Annotazione: Componenti




```
<OWLOntologies localID="ID01">
 <Ontology localID="A" version="2">
  <owl:imports rdf:resource="http://www.w3.org/TR/2004/REC-owl-guide-20040210/A/>
 </Ontology>
 <Ontology localID="B" version="1">
  <owl:imports rdf:resource="http://www.w3.org/TR/2004/REC-owl-guide-20040210/B"/>
 </Ontology>
 <OntologyAdd name="C" localID="C" uri="file:///./C">
  <Element>
   <owl:ObjectProperty rdf:about="C#R1" >
    <rdfs:domain rdf:resource="B#C2" />
    <rdfs:range rdf:resource="A#C1" />
   </owl:ObjectProperty>
  </Element>
 </OntologyAdd>
</OWLOntologies>
```

Modello Annotazione: Componenti




```
.....
<Annotation ID="10101" version="1">
 <FullDoc id="Doc0001-f01"  idRef="Doc0001">
  <Ontology idref="PersonalData">
   <isInstanceOf> #BirthCertificate</isInstanceOf>
  </Ontology>
 </FullDoc>
 <SinglePart id="Doc0001-s01" idRef="Doc0001">
  <Part id="part01">
   <TextPosition from="30"/>
   <TextValue>John Doe</TextValue>
  </Part>
  <Ontology idref="PersonalData">
   <isInstanceOf>#FullName </isInstanceOf>
  </Ontology>
 </Singlepart>
 <SinglePart id="Doc0001-s02" idRef="Doc0001">
  <Part id="part02">
   <TextPosition from="80"/>
   <TextValue>Rome</TextValue>
  </Part>
  <Ontology idref="Cities">
   <isLinkedTo>#Rome</isLinkedTo>
  </Ontology>
 </SinglePart>
</Annotation>
...
```

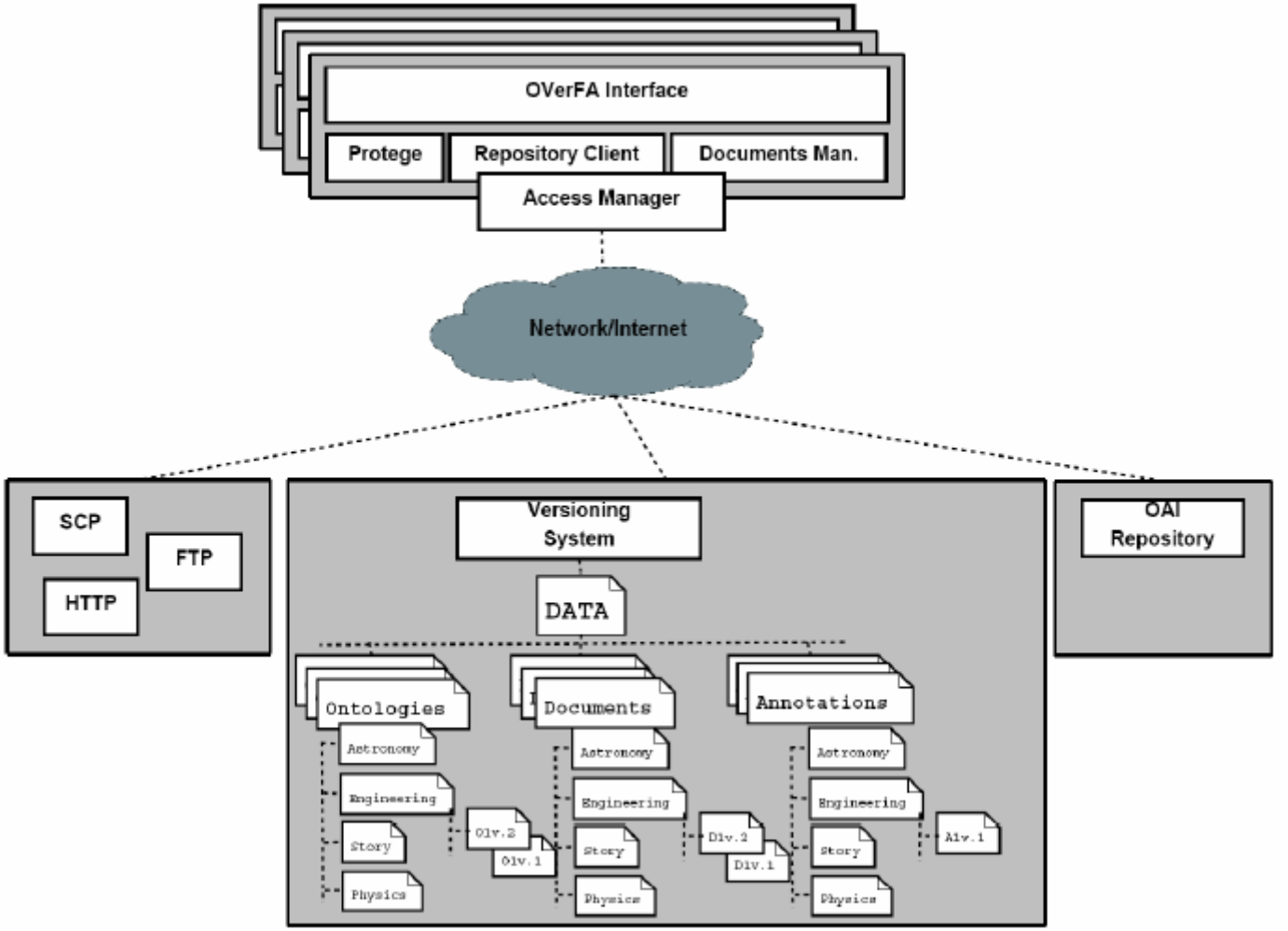
Tools di Annotazione

- Diversi tool di Annotazione in Letteratura:
  - OntoMat; IAnnotate; SMORE; MnM; Annozilla; Cohse; Trellis; Melita; Gate; KIM; etc.
- Problemi:
  - Spesso annotazione embedded
  - Nessun supporto per repository distribuiti (KIM usa SESAME)
  - Ognuno implementa un proprio modello di annotazione
  - Annotano un solo tipo di media (quasi sempre testo/html)
  - Nessun supporto al versioning

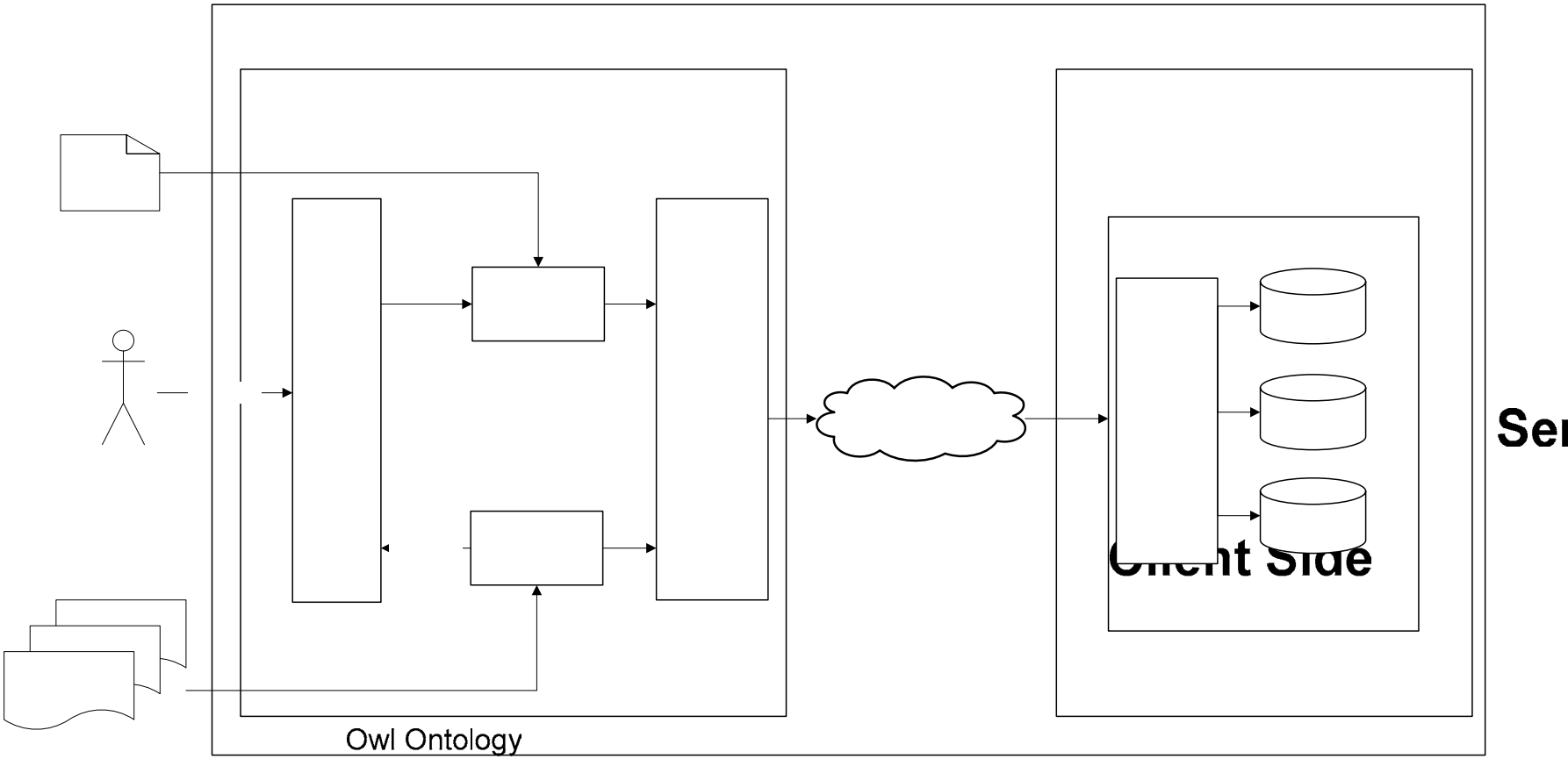
Lo strumento di annotazione definito

- Strumento Collaborativo per annotazioni manuali
  - Versioning
  - Annotazioni
    - Documenti
    - Gruppi di Documenti
    - Parti di Documenti
  - Relazioni con
    - Classi
    - Istanze
    - Proprietà
  - Diversi tipi di Media
    - Al momento:Plain text, immagini JPG, rtf/word, ppt, pdf
  - Annotazioni Collaborative (multi-utente)

Architettura

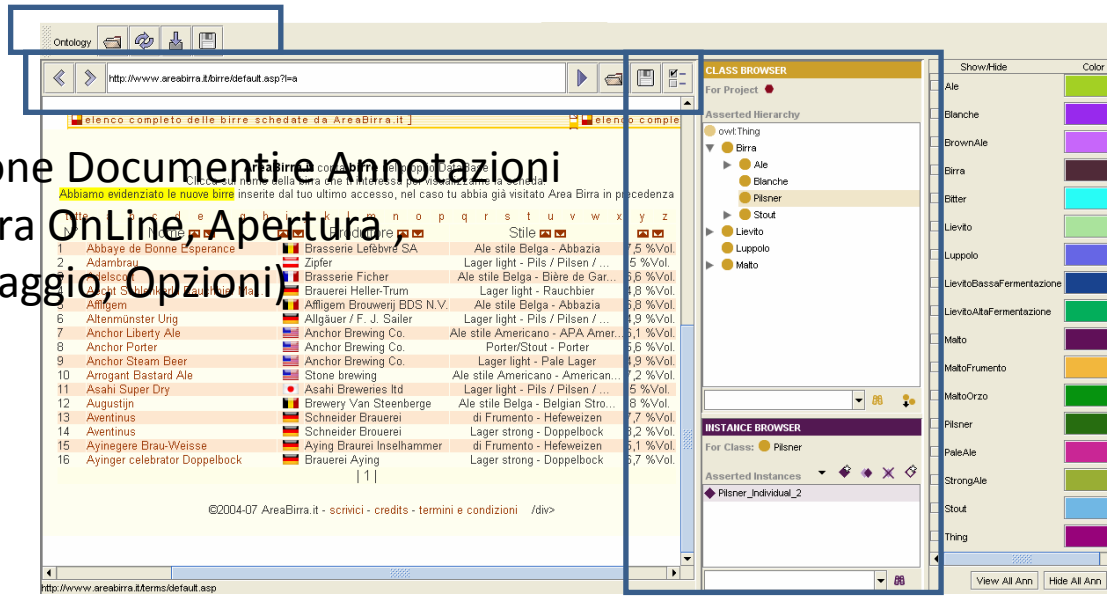


Architettura



Interfaccia

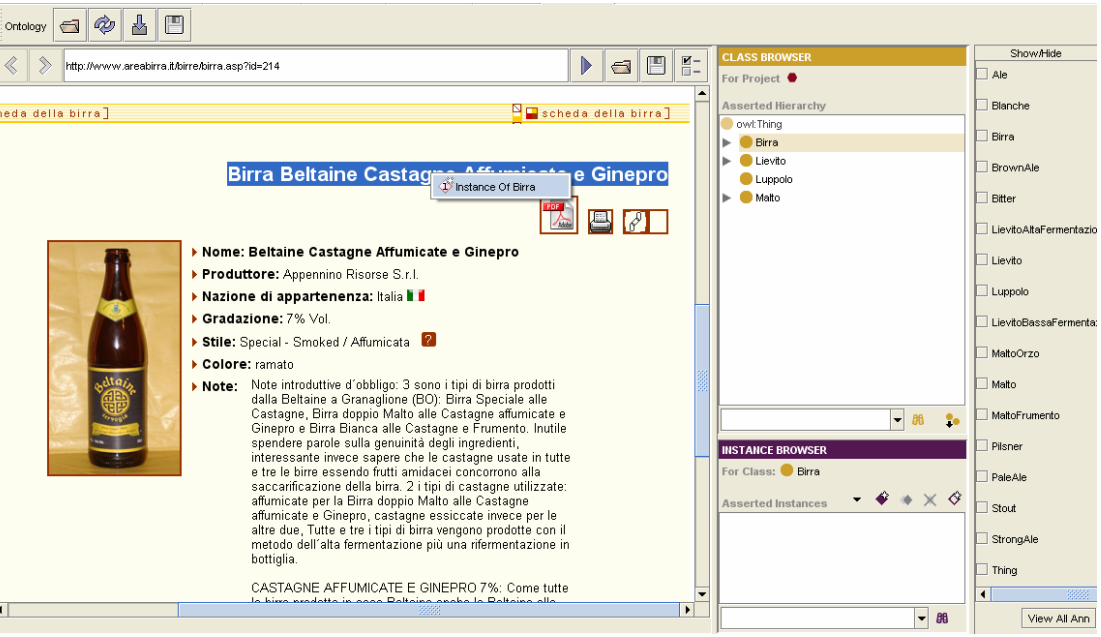
Gestione Ontologie (Apertura, Update, Creazione Repository, Salvataggio (Locale/SVN)



estione Documenti e Annotazioni
ettura OnLine, Apertura,
alvataggio, Opzioni)

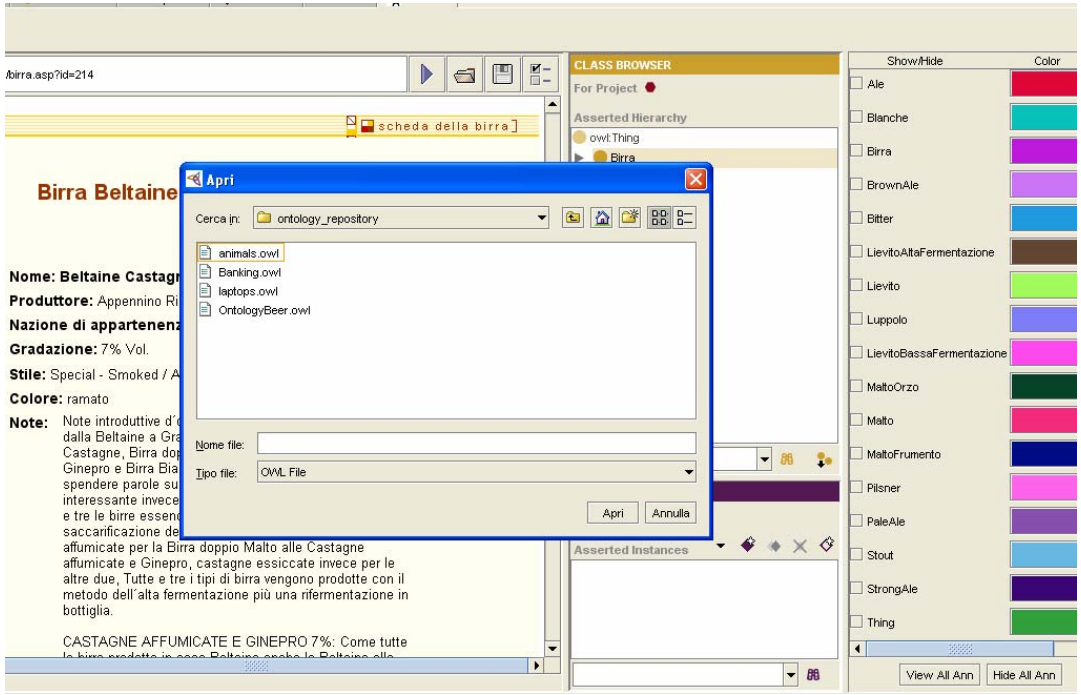
Browser Classi e Istanze di Ontologie OWL

Annotazione con Concetti



1. Si carichi il documento da annotare, digitando l'url nella barra degli indirizzi o usando il pulsante "Open" per caricare un file di testo;
2. Si selezioni il concetto da annotare nel Class Browser del pannello centrale;
3. Si selezioni il testo da annotare nel browser;
4. Si effettui un click col tasto destro del mouse e si scelga l'opzione <I> Instance Of <nome_classe>, che crea un'istanza per il concetto selezionato.

Annotazione con più Ontologie

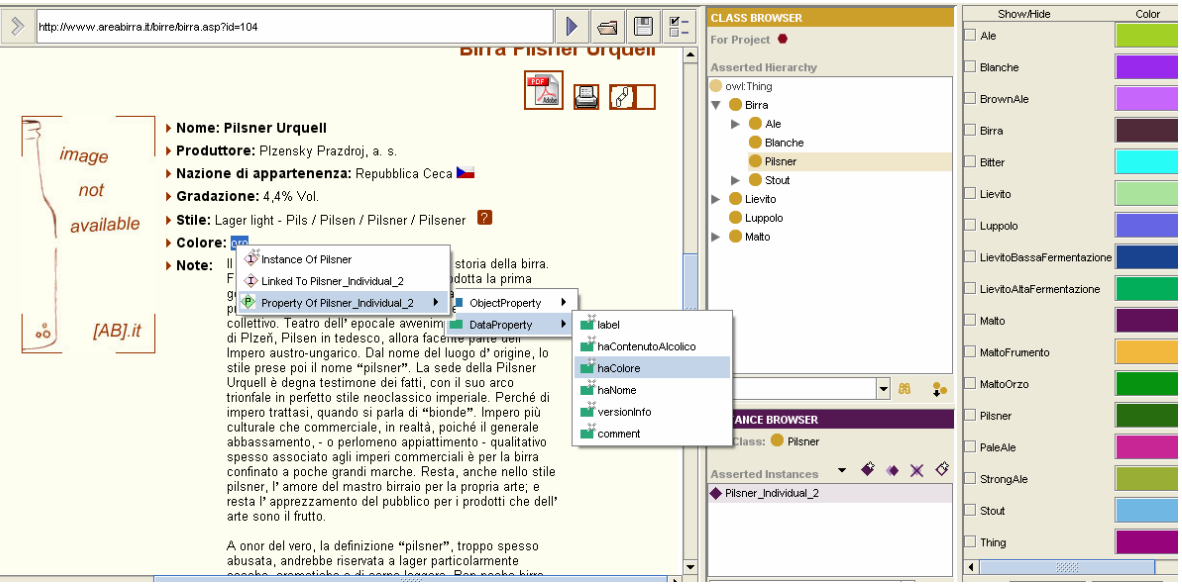


1. Si carichi il documento da annotare, digitando l'url nella barra degli indirizzi o usando il pulsante "Open" per caricare un file di testo;
2. Si selezioni il concetto da annotare nel Class Browser del pannello centrale;
3. Si selezioni il testo da annotare nel browser;
4. Si effettui un click col tasto destro del mouse e si scelga l'opzione <I> Instance Of <nome_classe>, che crea un'istanza per il concetto selezionato.
5. Tramite la toolbar "Ontology" si carichi una nuova ontologia utilizzando il pulsante e si proceda di nuovo dal passo 2.

Annotazione con Istanze già presenti in Ontologia

1. Si carichi il documento da annotare, digitando l'url nella barra degli indirizzi o usando il pulsante "Open" per caricare un file di testo;
2. Si selezioni l'istanza da annotare nell'Instance Browser del pannello centrale;
3. Si selezioni il testo da annotare nel browser;
4. Si effettui un click col tasto destro del mouse e si scelga fra le opzioni la terza di esse <I> Linked To <nome-individuo>, la quale crea un collegamento all'istanza per il testo selezionato.

Annotazione con Proprietà di una istanza

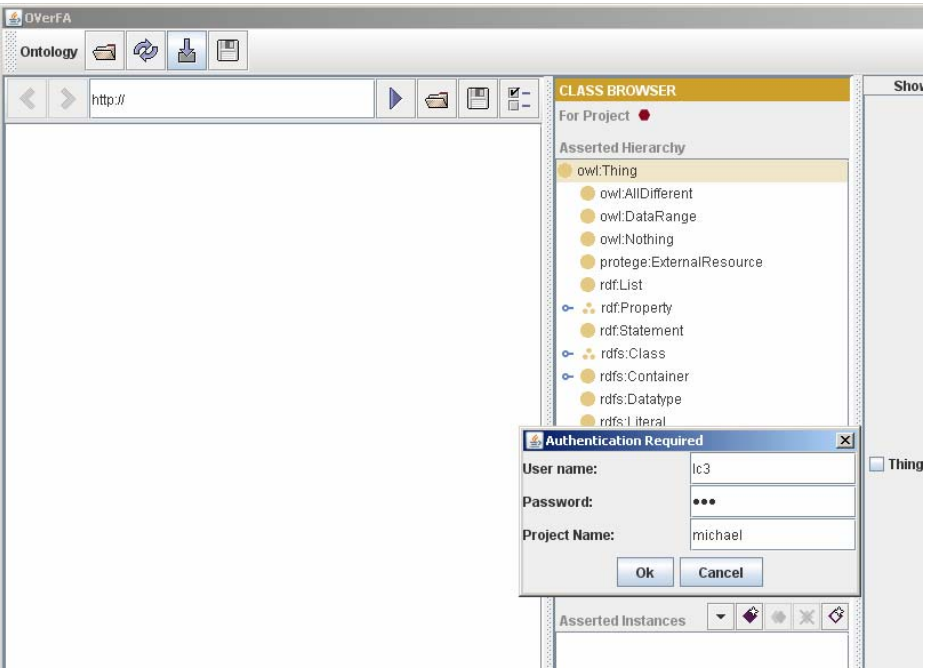


1. Si carichi il documento da annotare, digitando l'url nella barra degli indirizzi o usando il pulsante "Open" per caricare un file di testo;
2. Si selezioni il concetto da annotare nel Class Browser del pannello centrale;
3. Si selezioni il testo da annotare nel browser;
4. Si effettui un click col tasto destro del mouse e si scelga l'opzione <I> Instance Of <nome_classe>, che crea un'istanza per il concetto selezionato;
5. Una volta creata l'istanza di un concetto è possibile esplorarne le sue proprietà facendo doppio click sulla istanza appena creata nell'Instance Browser.
6. Si voglia ora associare a tale istanza altre istanze utilizzando le sue proprietà:
    1. Si effettui un click col tasto destro del mouse e si scelga la terza opzione (nome dell'istanza), si aprirà una tendina con le proprietà dell'istanza selezionata al passo precedente;
    2. Si effettui un click col tasto sinistro del mouse sulla proprietà da utilizzare;
    3. Ora l'istanza ha associata, nel campo proprietà, una nuova istanza come da figura.

Annotazione con Istanze create durante l'annotazione

1. Si carichi il documento da annotare, digitando l'url nella barra degli indirizzi o usando il pulsante "Open" per caricare un file di testo;
2. Si selezioni l'istanza da annotare nell'Instance Browser del pannello centrale;
3. Si selezioni il testo da annotare nel browser;
4. Si effettui un click col tasto destro del mouse e si scelga fra le opzioni la terza di esse <I> Linked To <nome-individuo>, la quale crea un collegamento all'istanza per il testo selezionato.

Accesso Repository Subversion



OverFA e Modello di Annotazione Semantica

- OverFA permette
  - Annotazione di parti di documento
  - Annotazione di un documento nella sua interezza
  - Annotazione di file di testo, ppt, immagini e di annotazioni
- OverFA al momento non supporta:
  - Annotazioni con concetti di diverse ontologie tra loro associate.
  - Annotazioni di Gruppi di Documenti
  - Il Modello di Annotazione permette comunque di produrre tali annotazioni

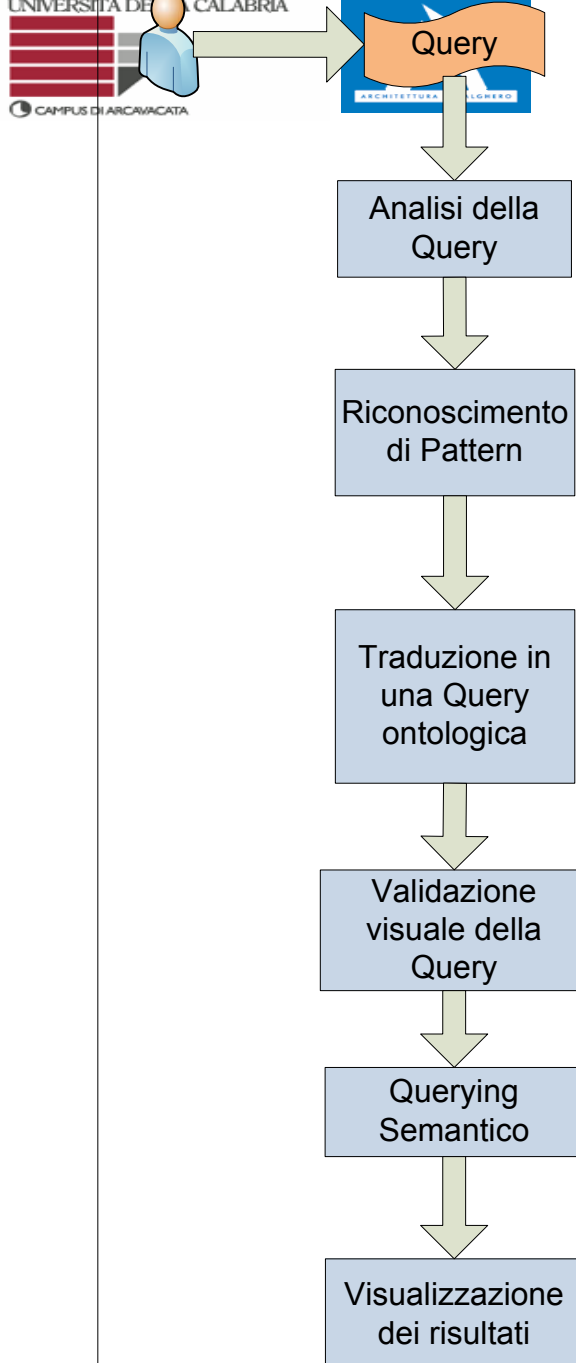
Definizione di tecniche semantiche per l'Information Retrieval e realizzazione di uno strumento Integrato

- Stato dell'arte ed analisi delle metodologie e delle tecniche, delle Tecnologie e degli strumenti per il Natural Language Processing
- Stato dell'arte ed analisi dei modelli e delle metodologie, delle tecnologie e fegli strumenti per l'Information Retrieval
- Stato dell' arte ed analisi delle metodologie, dei linguaggi standard e delle tecniche per la Rappresentazione della Conoscenza ed Inferenza in ambito Ontologico
- Stato dell'arte ed analisi delle metodologie, dei linguaggi standard e degli engines per il querying semantico
- Stato dell'arte ed analisi delle metodologie e delle tecniche per il matching di ontologie
- Definizione di una tecnica, di una architettura per il Semantic Information Retrieval
- Realizzazione di uno strumento prototipale per il Semantic Information Retrieval
- Definizione di una tecnica e realizzazione di uno strumento prototipale per il confronto (matching) di ontologie

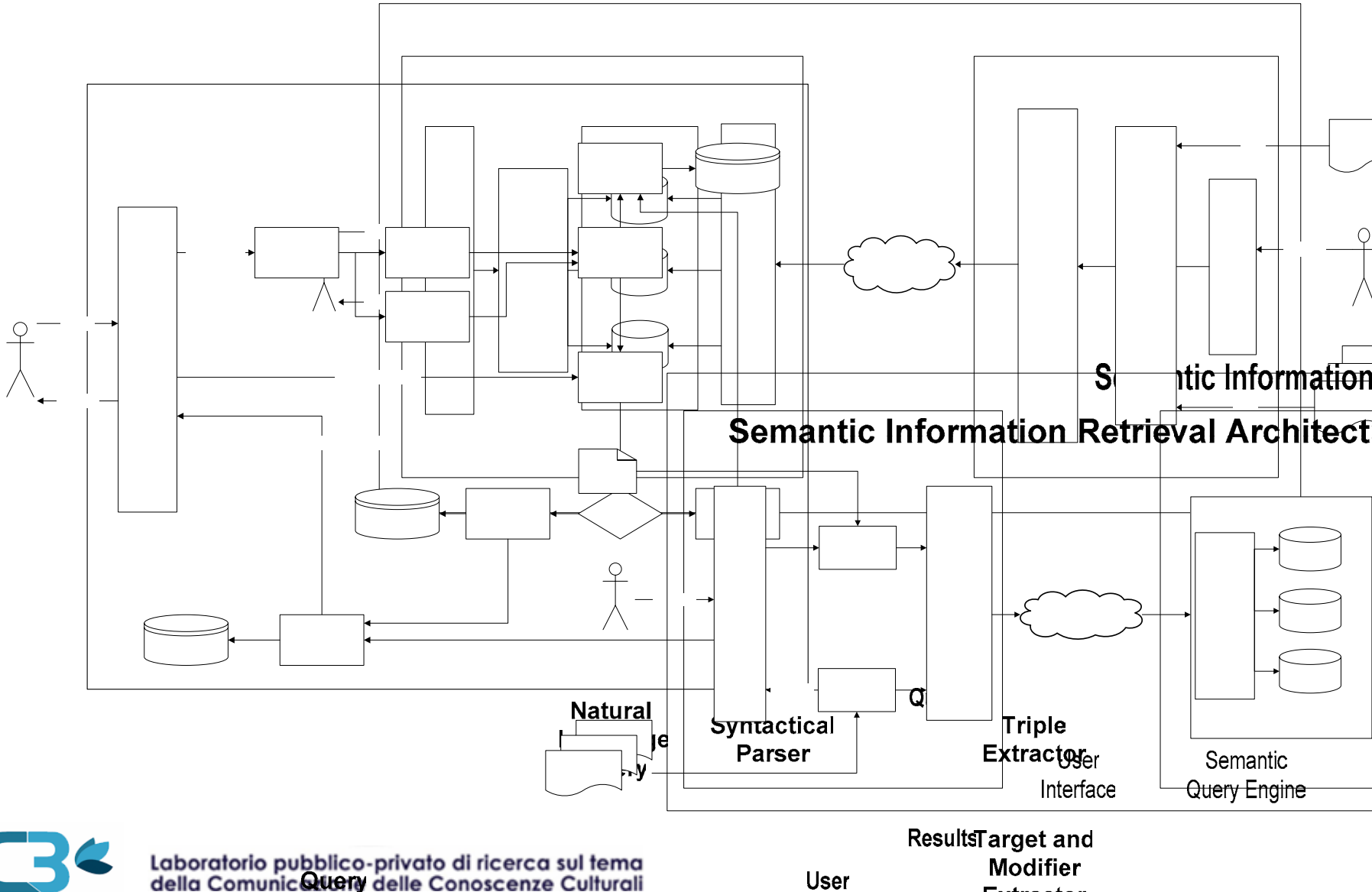
Definizione di una tecnica per il Semantic Information Retrieval

<u>Fasi elaborative</u>:

1. Analisi grammaticale e sintattica della query in linguaggio naturale
2. Estrazione dei sintagmi e riconoscimento di pattern euristici
3. Mapping semantico su una o più ontologie
4. Visualizzazione della query e validazione
5. Traduzione della query in un linguaggio di query semantico
6. Visualizzazione dei risultati



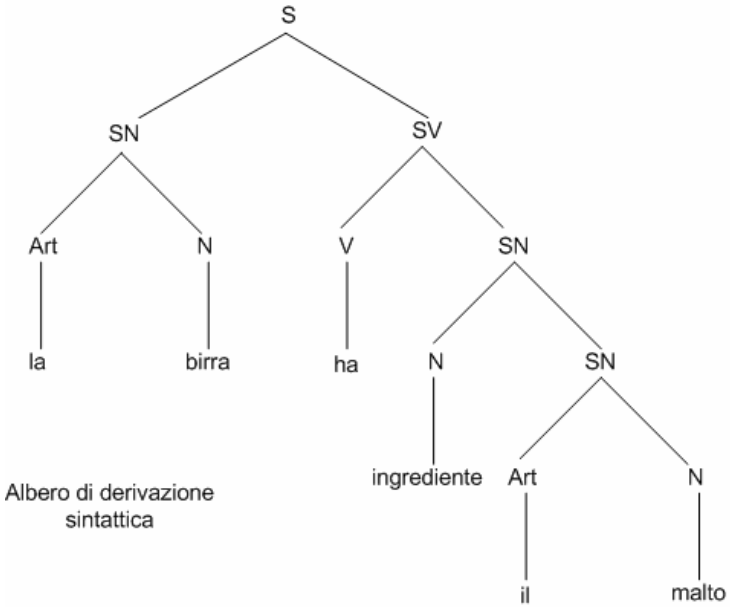
L' Architettura



I Componenti

Syntactical Parser

Effettua il parsing della query in linguaggio naturale, utilizzando l'analisi grammaticale e sintattica, e ne genera il corrispondente albero sintattico



Albero di derivazione sintattica

Regole:
1. S -> SN
2. SN -> A
3. SN -> N
4. SV -> V

Lessico:
la Art
birra N
ingrediente
ha V
il Art
malto N

Triple Extractor

Estrae uno "scheletro" dall'albero sintattico della query, formato da tutti i sintagmi riconosciuti in essa, ne ricava tutti i sintagmi nominali e ne ricerca le relazioni esistenti tra ogni coppia di essi.
Queste triple sono individuate tramite una ricerca di pattern euristici

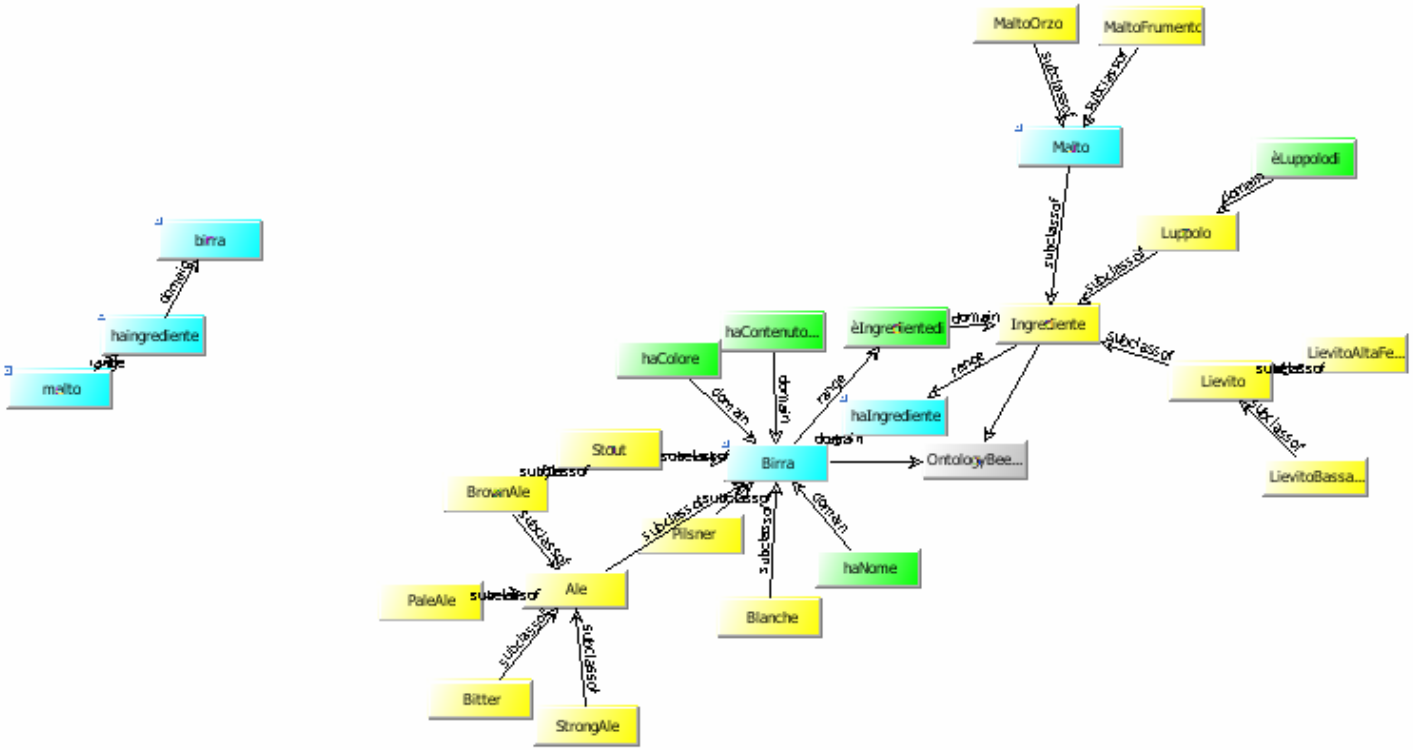
SN	SV , SN	SN
La birra	ha , ingrediente	il malto

I componenti

Semantic Mapper

Mappa le triple estratte dalla query in linguaggio naturale sulle triple definite nelle ontologie e ne estrae i sottografi (o path) in comune

Cost



ù

orre al

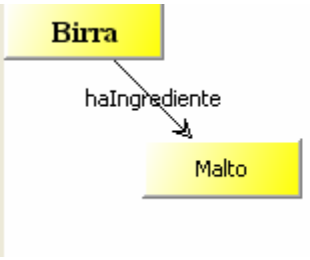
I Componenti

Ontology Querying

Il Query Graph viene sottoposto ad un motore di query semantico, a tale scopo sono stati scelti:

- il motore di query ARQ di Jena
- il motore di query del reasoner Pellet

in entrambi i casi il grafo rappresentante la query deve essere tradotto nel linguaggio SPARQL




```
PREFIX rdf:  <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT  ?Birra
WHERE {
  ?Birra rdf:type <http://www.owl-ontologies.com/Beer.owl#Birra> .
  ?Malto rdf:type <http://www.owl-ontologies.com/Beer.owl#Malto> .
  ?Birra <http://www.owl-ontologies.com/Beer.owl#haIngrediente> ?Malto .
}
```

I Componenti

Graph Matching Querying

In questo caso il Query Graph viene sottoposto ad un motore di query semantico appositamente sviluppato, il quale cerca le corrispondenze tra questo e gli Annotation Graph

Il confronto è effettuato tramite tecniche di schema matching sia di tipo sintattico-semantico che strutturale

Le etichette dei nodi dei due grafi sono confrontate secondo metriche di similarità tra stringhe e l'utilizzo di Wordnet

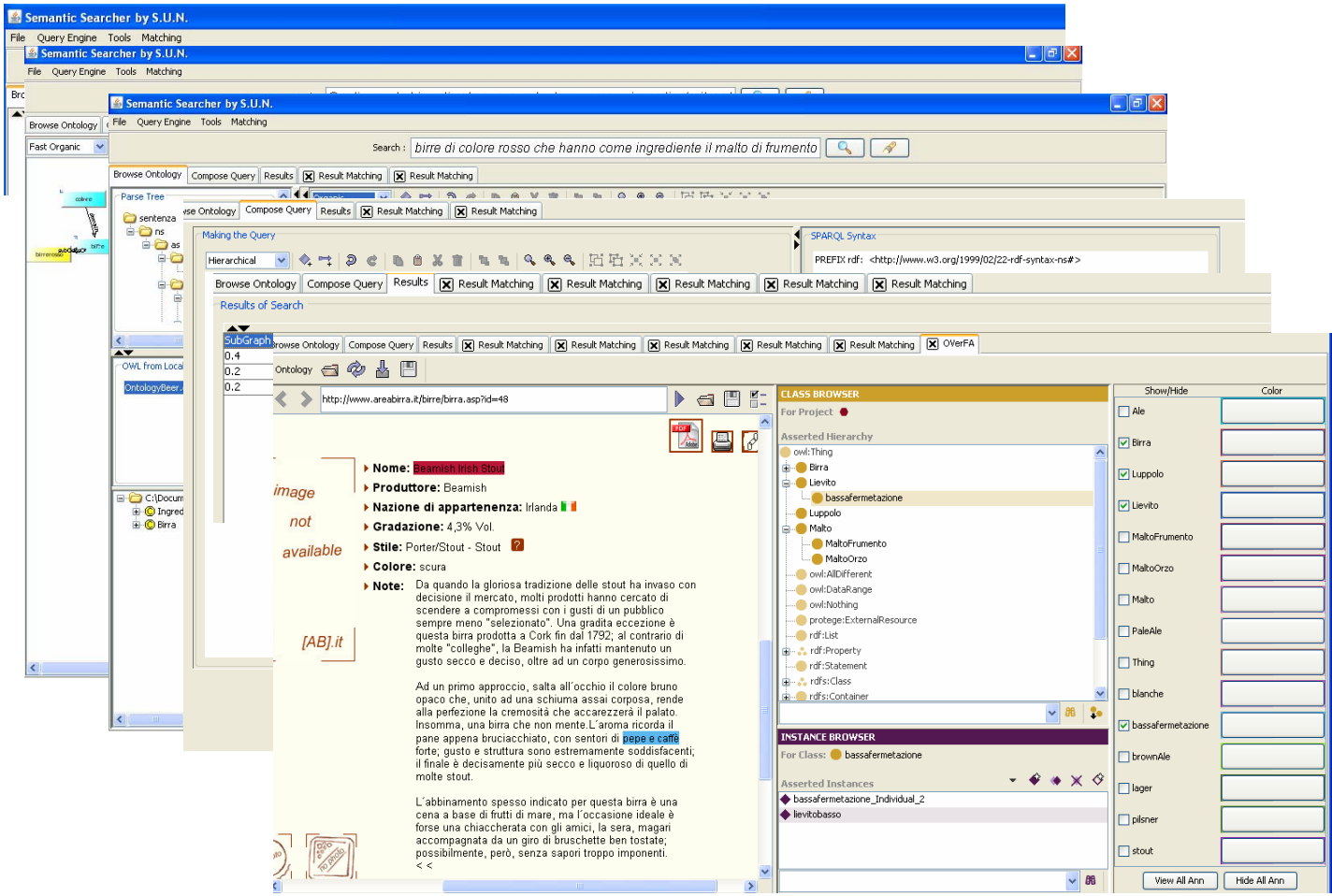
La tecnica utilizzata per effettuare il matching strutturale si basa sulla ricerca del massimo comune sottografo (mcs). A tale scopo viene utilizzato l'algoritmo di McGregor

Metrica di Similarità:

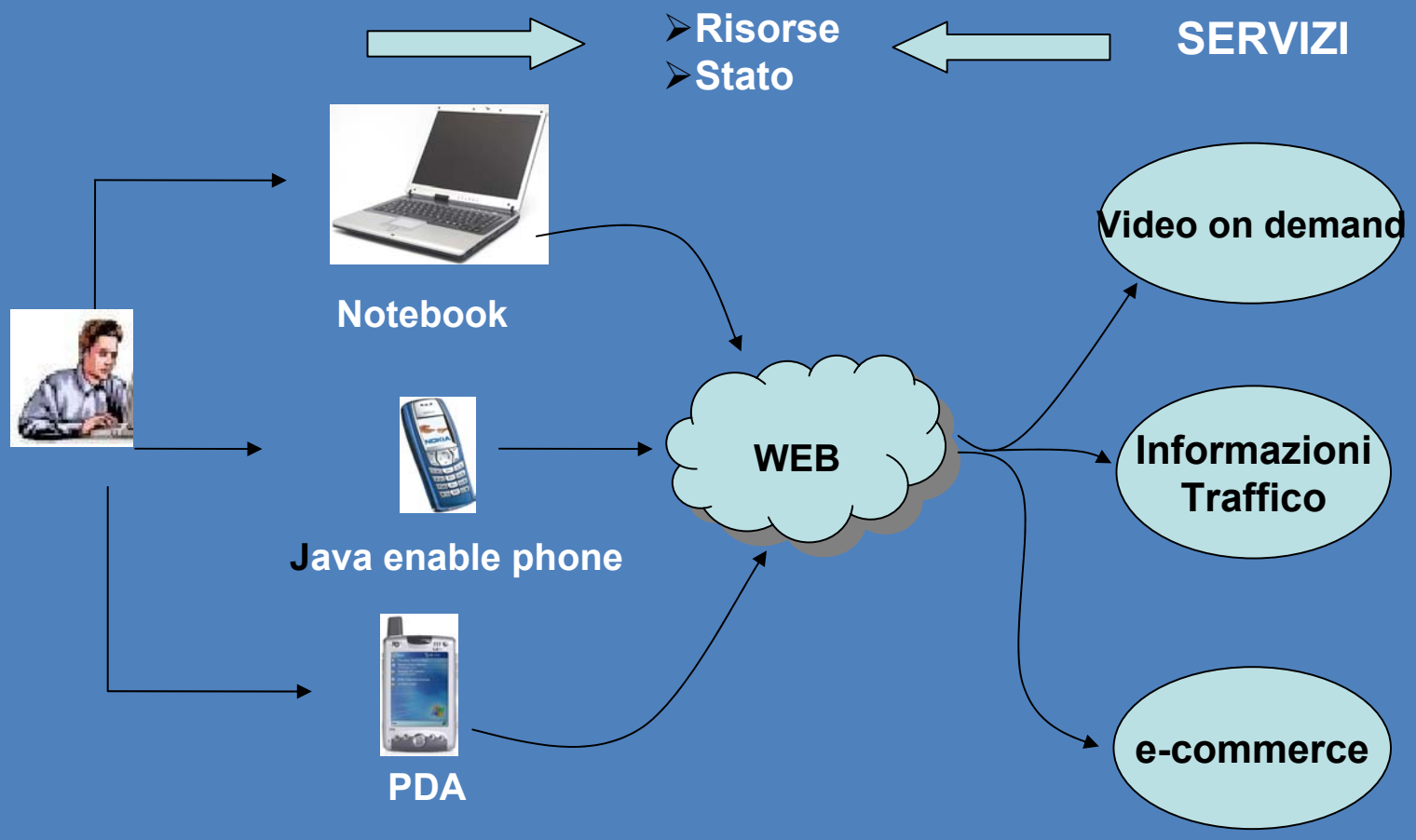
$$\mathrm{Sim}(G_Q, G_A) = \frac{|mcs(G_Q, G_A)|}{|G_Q|}$$

```
procedure McGregor_MCS(s)
begin
   while (NextPair(s,n1,n2))
      if (IsFeasiblePair(s,n1,n2)) then
         s' = AddPair(s,n1,n2);
         if (size(s') > CurrentMCSSize) then
            SaveCurrentMCS(s')
            CurrentMCSSize = size(s');
         end if
         if(!LeafOfSearchTree(s') and !PruningCondition(s')) then
            McGregor_MCS(s');
         end if
          BackTrack(s');
       end if
   end while
end procedure
```

Alcune Snapshot del Prototipo

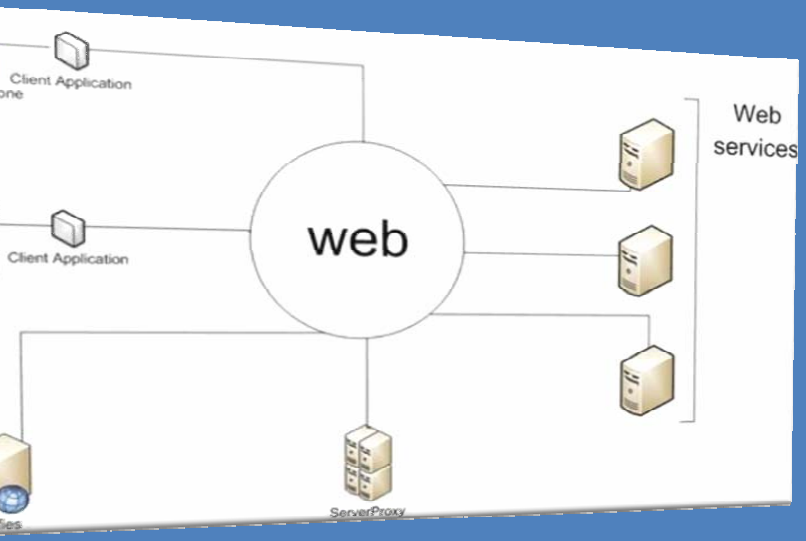


Adattamento dellla fruizione rispetto al terminale



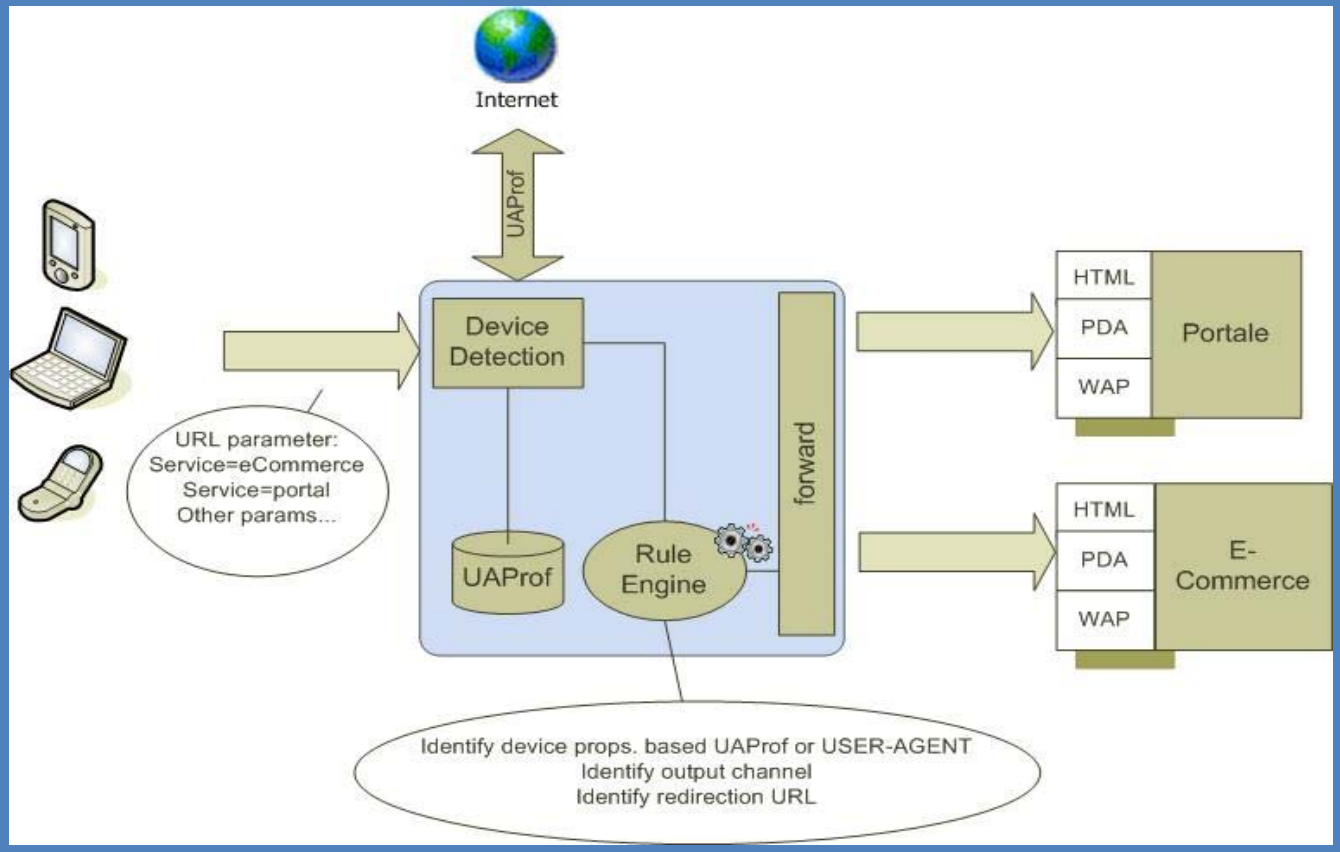
Obiettivi

Ottimizzare la QoS e l'utilizzo delle risorse attraverso lo sviluppo di un middleware che supporta:



- **Profilazione utente/device/sessione**
- **Personalizzazione dell'accesso ai servizi**
- **Adattamenti dei contenuti erogati**

Esempio di sistema multicanale



Funzionalità del dimostratore

- **Profilazione**
  - Profilazione classica standard
  - Profilazione avanzata con agenti
- **Sistema di adattamento**
  - Gestione della sessione e del profilo standard
  - Tunnelling basato su proxy
- **Interfacciamento a una teca digitale**
  - Adattamento MAG
  - Adattamento Immagini

Tipi di profilazione

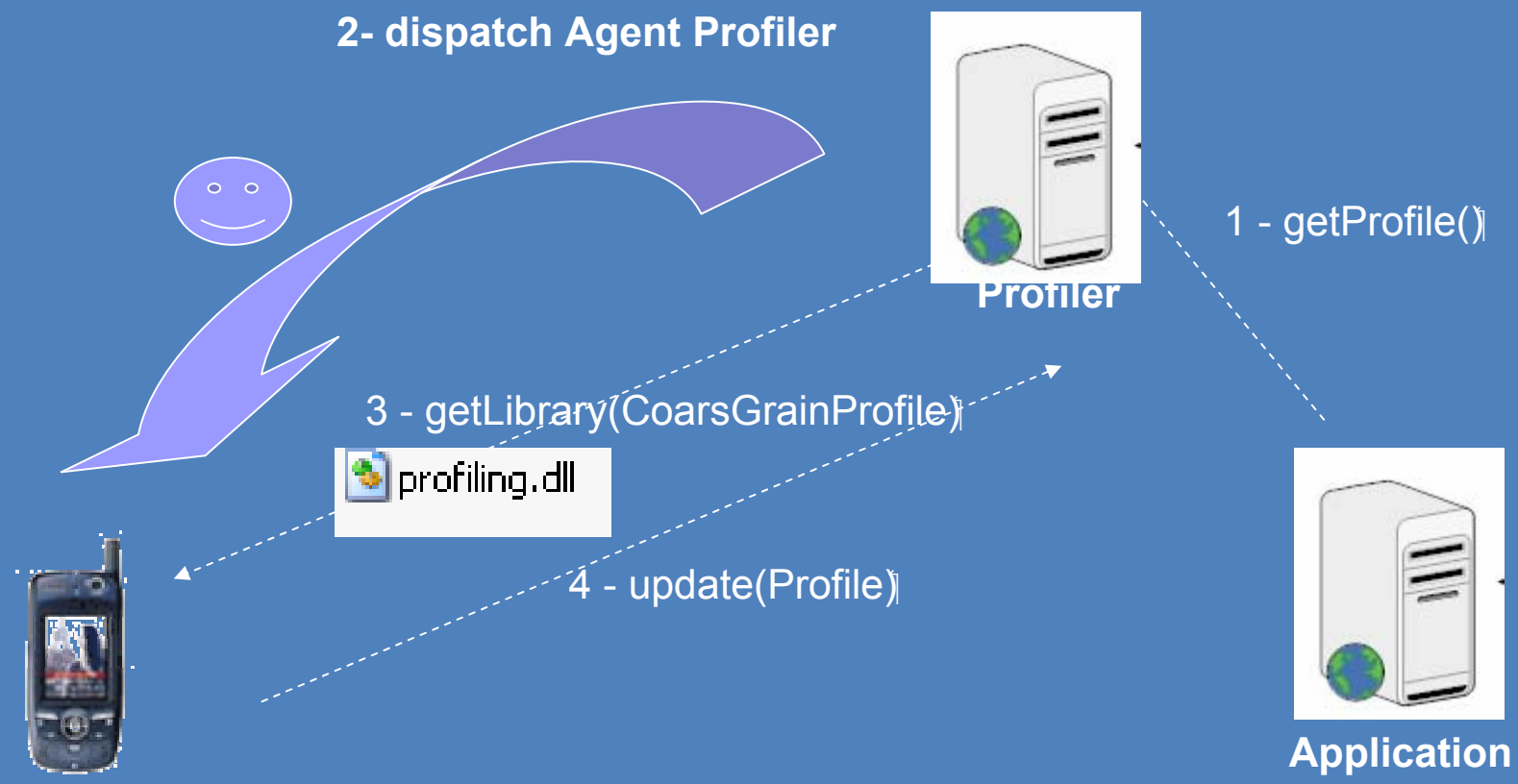
- Profilazione classica standard:

Il sistema riconosce il dispositivo utilizzando le informazioni presenti nell'header della richiesta http, o informazioni statiche registrate dall'utente

- Profilazione avanzata con agenti

Un agente software viene spostato sul dispositivo per misurare parametri dinamici in locale (banda, batteria, posizione gps, dimensioni schermo, memoria disponibile)

Profilazione ad agenti



Rappresentazione del Profilo

UAPROFF:

Uno standard che descrivein XML le capacità del dispositivo, come:

- la piattaforma Hardware,
- il sistema operativo,
- le caratteristiche del display,
- della memoria,
- dello stato della batteria,

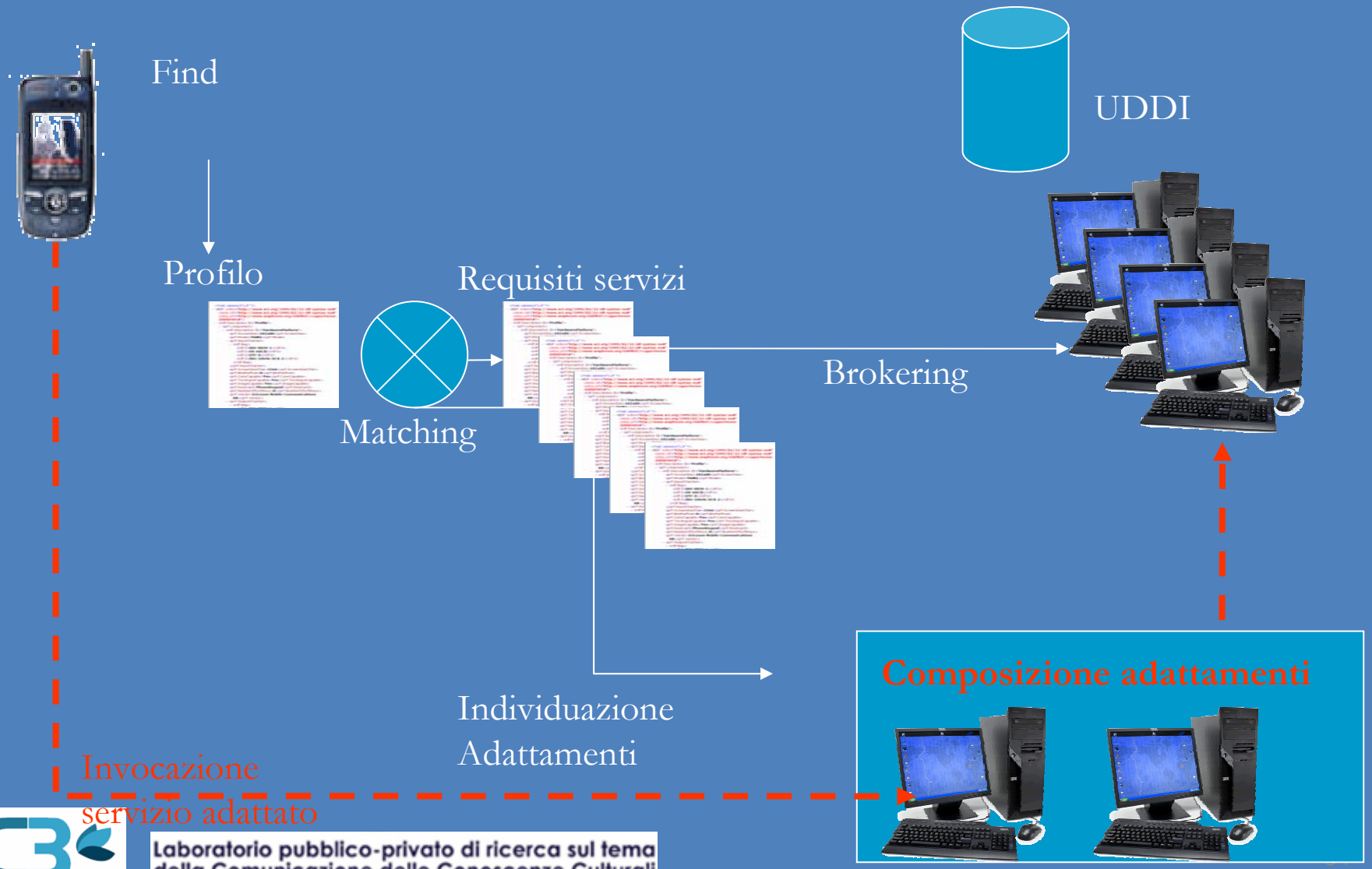
etc.

```
<?xml version="1.0" encoding="utf-8" ?>
<!--  * * (c) 2008 Research In Motion Ltd., * 295 P
10:41:44 EDT 2008 *  * Notes: *        Elements that
- <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rd
  xmlns:prf="http://www.openmobilealliance.org/tech/
  xmlns:mms="http://www.openmobilealliance.org/tech
  - <rdf:Description rdf:ID="DeviceProfile">
    <!--  Hardware Platform Description  -->
    + <prf:component>
    <!--  Software Platform Description  -->
    + <prf:component>
    <!--  Browser UA Description  -->
    + <prf:component>
    <!--  Network Characteristics Description  -->
    + <prf:component>
    <!--  WAP Characteristics Description  -->
    + <prf:component>
  </rdf:Description>
</rdf:RDF>
```

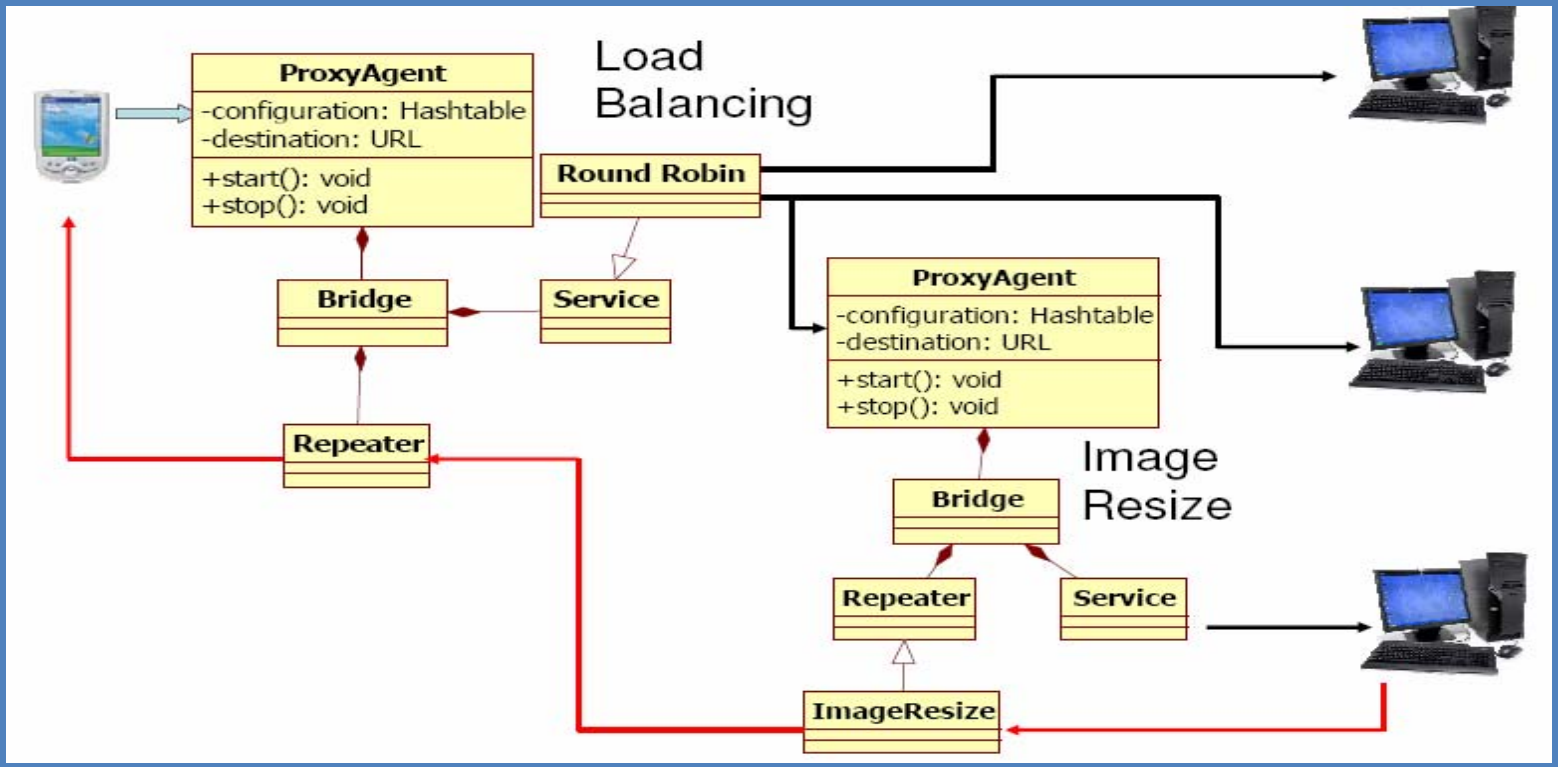
La nostra applicazione consente di misurare parametri dinamici quali:

- Stato della batteria
- Posizione
- Risorse impegnate
- Banda

Ricerca del servizio e adattamento



Adattamento basato su proxy tunneling



L'informazione (richiesta e dato) attraversa un tunnel di agenti proxy

- Lungo il percorso diretto viene specializzata e reindirizzata la richiesta

- Lungo il percorso inverso viene adattata la risposta

Interfacciamento alla teca

- Adattamento MAG

Data la richiesta di un MAG viene applicata per mezzo di un proxy dedicato una trasformazione xsl. Il risultato è la formattazione html del MAG, il recupero delle informazioni di interesse, la selezione delle immagini più vicine alle caratteristiche del display utente.

- Adattamento Immagini

Data la richiesta di un immagine un proxy dedicato effettua, se necessario, un resize e un encoding dell'immagine in funzione delle caratteristiche del terminale.

Esempio MAG


```
-<metadigit version="2.0.1" xsi:schemaLocation="http://www.iccu.sbn.it/metaAG1.pdf metadigit.xsd">
 -<gen creation="2008-01-14T13:32:17" last_update="2008-01-14T13:32:17">
    <stprog>http://www.regione.sardegna.it</stprog>
    <agency>Archivio storico Enel "Angelo Omodeo"</agency>
    <access_rights>1</access_rights>
    <completeness>0</completeness>
  -<img_group ID="jpeg300">
   -<image_metrics>
      <niso:samplingfrequencyunit>2</niso:samplingfrequencyunit>
      <niso:samplingfrequencyplane>2</niso:samplingfrequencyplane>
      <niso:xsamplingfrequency>300</niso:xsamplingfrequency>
      <niso:ysamplingfrequency>300</niso:ysamplingfrequency>
      <niso:photometricinterpretation>RGB</niso:photometricinterpretation>
      <niso:bitpersample>8,8,8</niso:bitpersample>
    </image_metrics>
   -<format>
      <niso:name>JPG</niso:name>
      <niso:mime>image/jpeg</niso:mime>
      <niso:compression>JPG</niso:compression>
    </format>
  </img_group>
  -<img_group ID="jpeg72">
   -<image_metrics>
```


Contiene riferimenti a scansioni in diversi formati e risoluzioni

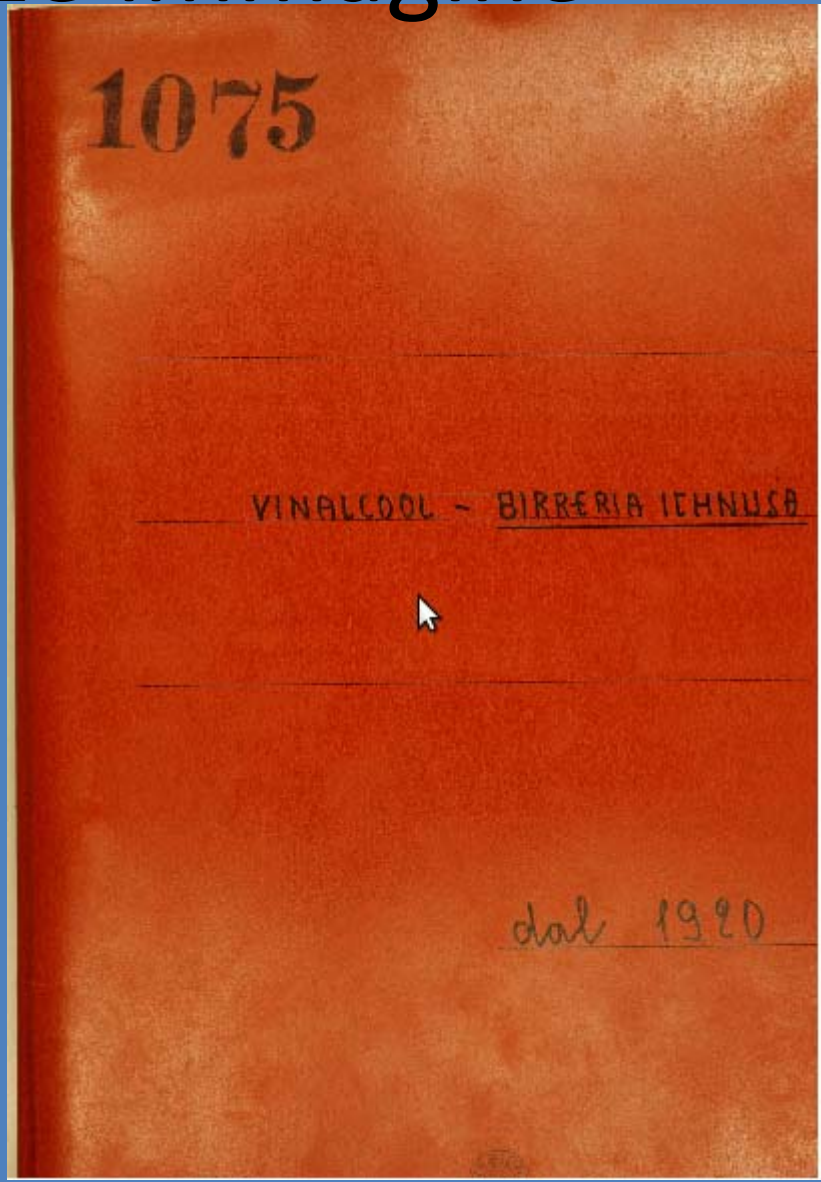
Adattamento MAG

Risultato

http://www.regione.sardegna.it
Archivio storico Enel "Angelo Omodeo"

Immagine	Filesize	URL
B_01_F01_0001	69032	./B_01_F01/jpeg72/B_01_F01_0001.jpg
B_01_F01_0002	72380	./B_01_F01/jpeg72/B_01_F01_0002.jpg
B_01_F01_0003	64016	./B_01_F01/jpeg72/B_01_F01_0003.jpg
B_01_F01_0004	54521	./B_01_F01/jpeg72/B_01_F01_0004.jpg
B_01_F01_0005	91014	./B_01_F01/jpeg72/B_01_F01_0005.jpg
B_01_F01_0006	66256	./B_01_F01/jpeg72/B_01_F01_0006.jpg
B_01_F01_0007	81188	./B_01_F01/jpeg72/B_01_F01_0007.jpg
B_01_F01_0008	51524	./B_01_F01/jpeg72/B_01_F01_0008.jpg
B_01_F01_0009	102280	./B_01_F01/jpeg72/B_01_F01_0009.jpg
B_01_F01_0010	133656	./B_01_F01/jpeg72/B_01_F01_0010.jpg
B_01_F01_0011	58726	./B_01_F01/jpeg72/B_01_F01_0011.jpg
B_01_F01_0012	135776	./B_01_F01/jpeg72/B_01_F01_0012.jpg
B_01_F01_0013	72069	./B_01_F01/jpeg72/B_01_F01_0013.jpg
B_01_F01_0014	45288	./B_01_F01/jpeg72/B_01_F01_0014.jpg
B_01_F01_0015	77965	./B_01_F01/jpeg72/B_01_F01_0015.jpg
B_01_F01_0016	71282	./B_01_F01/jpeg72/B_01_F01_0016.jpg
B_01_F01_0017	41037	./B_01_F01/jpeg72/B_01_F01_0017.jpg
B_01_F01_0018	48701	./B_01_F01/jpeg72/B_01_F01_0018.jpg
B_01_F01_0019	73212	./B_01_F01/jpeg72/B_01_F01_0019.jpg

Re-Format/Resize immagine



Pubblicazioni prodotte

- F.Moscato, B. Di Martino: “Semantic Web and Semantic Information Management”, Int. Journal of Web and Grid Services, Vol. 4, No. 3, 2008.
- F. Moscato, B. Di Martino, S. Venticinque, A. Martone, “OverFA: A collaborative Framework for Semantic Annotation of Documents and Web Sites”, International Journal of Web and Grid Services, Inderscience Press, Vol. 5, n. 1, pp. 30-45, 2009.
- Beniamino Di Martino, Angelo Martone, Francesco Moscato, Salvatore Venticinque, “A versioning based framework for semantic annotation of Web documents: OVerFA” , Proc. of Int. Conf. on Methods, Moels and Information Technologies for Decision Support Systems (MTISD 2008), Lecce, Italy, 18-20 sept. 2008.
- B. Di Martino, “Semantic Web Services Discovery based on Structural Ontology Matching”, International Journal of Web and Grid Services, Inderscience Press, Vol. 5, n. 1, pp. 46-65, 2009.
- Beniamino Di Martino, Pasquale Femia, Laura Manelli, Sergio Muzzupappa, “Metodologie e strumenti informatici per una ontologia sul caso Aldo Moro”, Scritture di Storia, n. 5, pp. 359-380, luglio 2008, Edizioni Scientifiche Italiane.

Pubblicazioni Prodotte

- R. Aversa, B. Di Martino, N. Mazzocca, S. Venticinque, “Proxy Agents for Adaptive Delivery of Multimedia”, Journal of Mobile Multimedia), Rinton Press, Vol.4 No.3-4, pp. 227-240, 2008. ISSN 1550-4646.
- R. Aversa, B. Di Martino, S. Venticinque and N. Mazzocca, “An agents based middleware for personalization of multimedia service delivery in SOAs”, International Journal of Mobile Multimedia (JMM), Rinton Press, Vol. 4 n. 1, pp. 2-18, 2008. ISSN : 1550-4646.
- Rocco Aversa, Beniamino Di Martino and Salvatore Venticinque “Service based Integrated Infrastructure for Adaptive Multimedia Delivery”, Proc. of International Conference on Complex, Intelligent and Software Intensive Systems (CISIS’08), Barcelona (Spain), 4-7 Mar. 2008, pp. 108-114, IEEE CS Press, (ISBN 0-7695-3109-1).
- D. Lorenzoli, D. Tosi, S. Venticinque, R. Micillo, Designing Multi-Layers Self-Adaptive Complex Applications. In Proceedings of the 2007 ACM SIGSOFT Symposium on the Foundations of Software Engineering (ESEC/FSE), SOQUA Workshop. Dubrovnik, Croatia, ISBN:978-1-59593-724-7, pp. 70-77, Sept. 2007.
- Rocco Aversa, Beniamino Di Martino, Salvatore Venticinque and Nicola Mazzocca, “A Mobile Agents based Infrastructure to deliver value added services SOAs”, in Scalable Computing: Practice and Experience, Special Issue: Practical Aspects of Large-Scale Distributed Computing, Edited by Dana Pectu, Vol. 8, N. 3, September 2007, ISSN 1895-1767, pages 313-324