



Dipartimento di Informatica Corso di Laurea in Informatica



Tesi di Laurea in Metodi per
il Ritrovamento dell'Informazione

Sistemi di Supporto alle Decisioni basati sull'elaborazione di dati testuali e metadati per la gestione delle politiche di innovazione

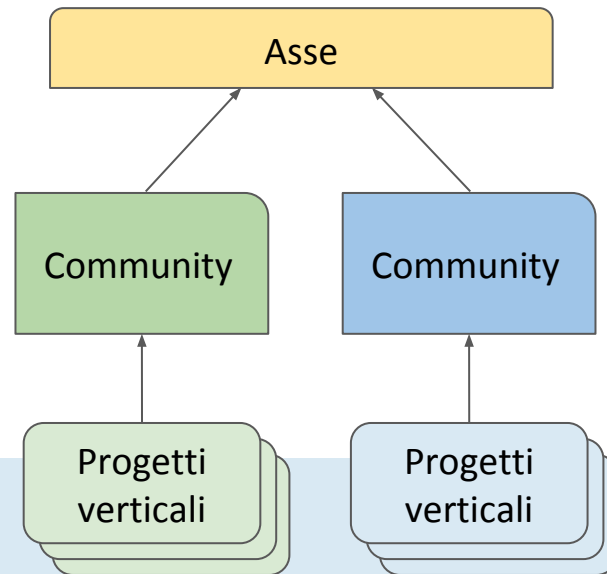
Relatore:
prof. Pasquale Lops
dott. Pierpaolo Basile

Laureando:
Flavio Petruzzellis

Il Programma Interreg-Med

Interreg-Med è un programma di finanziamento europeo finalizzato allo sviluppo sostenibile dell'area del Mediterraneo tramite il supporto di progetti innovativi operanti in diversi ambiti.

I progetti sono raggruppati in nove Community, a loro volta suddivise in tre Assi tematici principali.



Il Progetto TALIA

Il progetto TALIA, finanziato nell'ambito del Programma Interreg-Med, è un progetto trasversale che coinvolge tutti i progetti appartenenti alla Community **Social and Creative**.

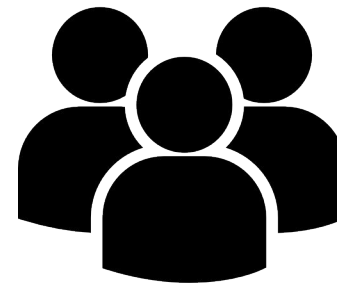
TALIA ha l'obiettivo di favorire lo sviluppo di questa Community, aumentando la capacità di cooperazione tra gli attori coinvolti nei suoi progetti.



Il Progetto TALIA

Per raggiungere l'obiettivo del progetto si è scelto di sviluppare un Sistema di Supporto alle Decisioni basato sull'analisi intelligente dei documenti scritti in ogni progetto del programma Interreg-Med.

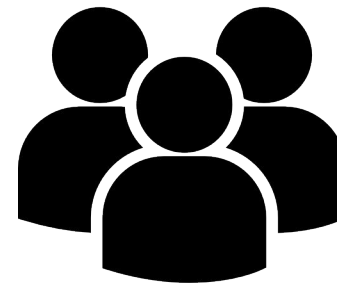
Il Sistema di Supporto Decisionale è basato sul **Semantic Framework**, una piattaforma per l'indicizzazione e l'analisi semantica dei documenti.



Il Sistema di Supporto alle Decisioni

Il Sistema di Supporto alle Decisioni mette a disposizione degli utenti due servizi principali: un **motore di ricerca** semantico e una **matrice di correlazione** tra concetti.

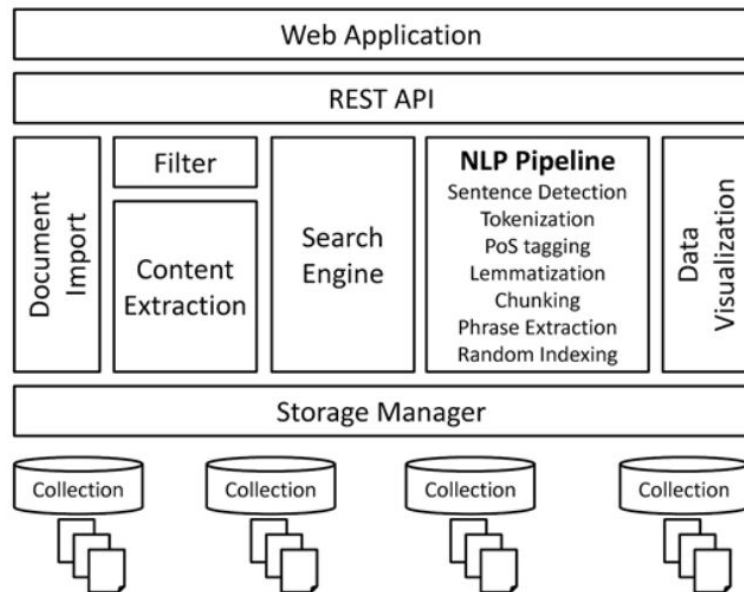
I due servizi hanno l'obiettivo di favorire il coordinamento e la cooperazione tra gli attori della Community Social and Creative facilitando la scoperta di informazioni utili e connessioni esistenti tra le azioni concrete svolte dai diversi attori.



Il Semantic Framework

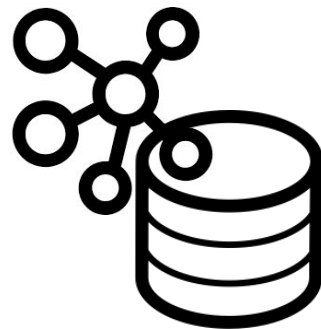
La piattaforma Semantic Framework consente di gestire collezioni di documenti e di applicarvi operazioni di elaborazione del linguaggio naturale e di rappresentazione vettoriale e semantica del testo.

Il sistema è stato sviluppato secondo un'architettura client-server che permette l'accesso ai servizi tramite API REST.



I dati strutturati nel dominio del progetto TALIA

Nel progetto TALIA, oltre a beneficiare della conoscenza estratta dai documenti, è risultato utile sfruttare le informazioni strutturate disponibili relative ad ogni Community e Progetto verticale.



Gestione dei dati strutturati

Inizialmente, le informazioni strutturate relative al dominio e a ciascun documento erano state memorizzate all'interno del Semantic Framework come campi dei documenti, sfruttando una funzionalità offerta dalla libreria Apache Lucene.

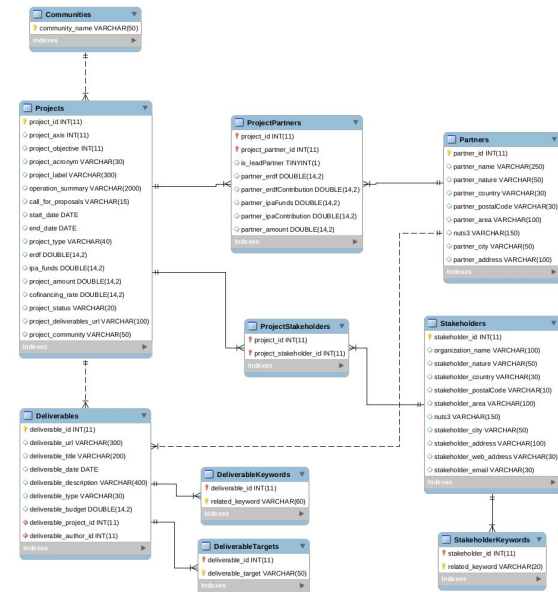


Gestione dei dati strutturati

La gestione tramite Lucene presentava alcuni problemi:

- ✗ Quantità di dati memorizzabile limitata;
- ✗ Dati ridondanti;
- ✗ Vincoli di integrità ingestibili;
- ✗ Impossibile rappresentare le relazioni tra le entità;

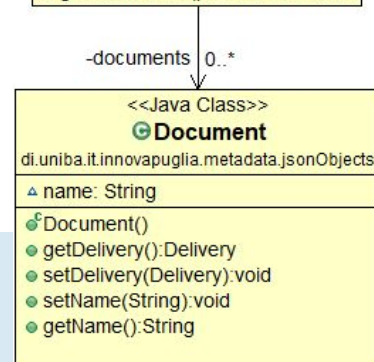
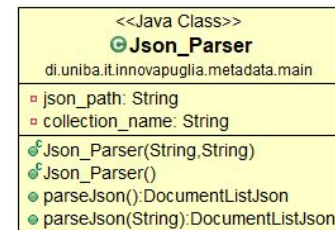
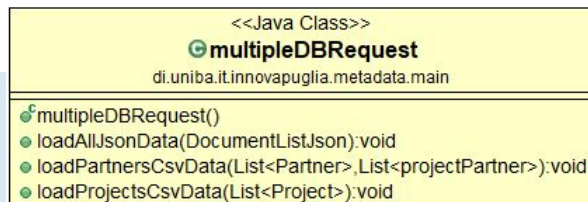
Per far fronte a questi problemi, è stata progettata una base di dati relazionale ed integrato nel Semantic Framework un modulo di interfaccia ad essa.



Gestione dei dati strutturati

I dati sono stati reperiti tramite un processo di crawling e scraping del sito Web ufficiale del programma Interreg-Med, che ha generato file Json e Csv.

Per gestire l'inserimento di questi dati e le operazioni di aggiornamento e interrogazione sul database creato, è stato implementato ed integrato nel Semantic Framework un modulo di gestione dei metadati.

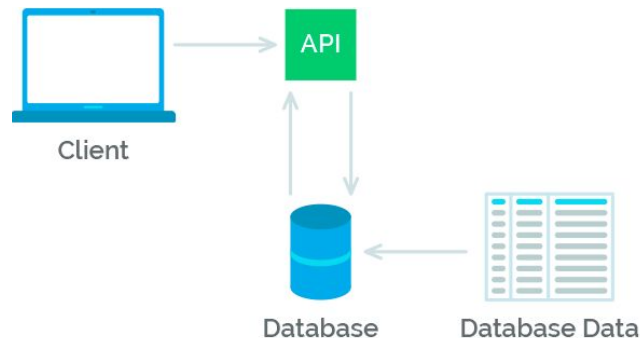


-documents 0..*

REpresentational State Transfer

Le funzionalità di gestione dei metadati sono state esposte tramite servizi REST.

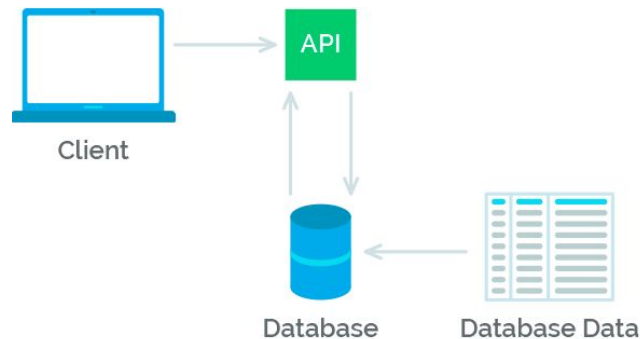
Lo stile architetturale REST impiega il protocollo HTTP per trasferire le informazioni in una comunicazione tra host client e server.



REpresentational State Transfer

Ogni entità coinvolta nella comunicazione è vista come “**risorsa**”, univocamente identificata e localizzata dal suo URL, della quale si può richiedere o modificare lo **stato**.

Ogni servizio viene invocato tramite una chiamata HTTP con un metodo corrispondente al tipo di operazione richiesta (GET per la visualizzazione, POST per la creazione, PUT per la modifica, ecc.).



Integrazione nel servizio di ricerca semantica

I dati strutturati sono stati integrati principalmente nel servizio di ricerca semantica. Ogni documento rilevante viene mostrato assieme alle informazioni relative ad esso ed al progetto al quale appartiene.

D.3.2.1_REGIONAL_ANALYSIS_PP3

Regional analysis on CCIs sector/subsectors and main aspects of innovation supporting system in CCIs sector/subsectors in Slovenia

Summarization

Deliverable details

Partners details

Date: 2017-03-29

Keywords: arts, awareness-raising, clustering, cultural heritage, economic cooperation, entrepreneurship, innovation capacity, sme

Target Audience: Higher education and research

Deliverable budget: €8,000

PROJECT NAME: ChIMERA

PROJECT BUDGET: €2,412,326

PARTNERS:

Name: Autonomous Region Friuli Venezia Giulia - Department for culture, sports and solidarity (**Lead**)

Budget: €346,203

Country: ITALY

Name: Basilicata Region

Budget: €262,665

Country: ITALY

Name: Creative Apulia Cluster Association

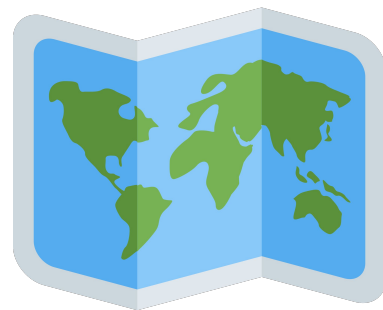
Budget: €250,440

Country: ITALY

Integrazione nel servizio di creazione delle mappe

I servizi di gestione dei metadati sono necessari anche alla realizzazione di un servizio di visualizzazione geografica dei risultati di una ricerca semantica.

Il servizio è utile ad individuare i luoghi in cui operano i partner attivi su un tema e a rappresentare graficamente le relazioni esistenti tra essi, integrando le informazioni rinvenute dall'analisi dei testi e quelle strutturate.



Integrazione nel servizio di creazione delle mappe

← openDOORS - LP - Lazio R... ↗

name
openDOORS - LP - Lazio Region Association of Cities and Municipalities - ANCI LAZIO

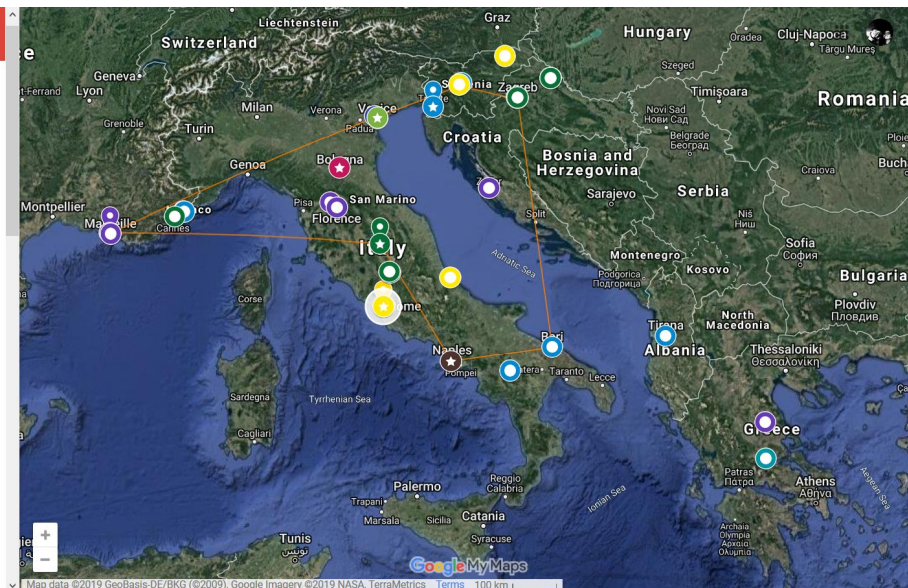
description
Partner:
Lazio Region Association of Cities and Municipalities - ANCI LAZIO

Project(s), type of project and role:
Open DOORS - Study Project - LP

Location:
VIA DEI PREFETTI, 41
00186
ROME
Lazio
Roma
ITALY

Brief description:
ANCI Lazio is a regional branch of the National Association of Italian Municipalities, representing almost all Lazio municipalities. ANCI Lazio is committed in promoting European policies and debates towards its members as well as local authorities' interests at national and regional level. The Association is particularly committed in:

- promoting small municipalities' interests to the



Conclusioni

L'integrazione nel Semantic Framework di un modulo di gestione dei dati strutturati ha reso più efficiente la loro gestione e più flessibile l'accesso ad essi rispetto alla tecnica inizialmente adottata.

Il modulo ha permesso, inoltre, l'arricchimento dei risultati di ricerca semantica e l'implementazione del caso d'uso di visualizzazione geografica dei risultati di ricerca.



Sviluppi futuri



Il sistema potrebbe essere integrato in modo più efficiente con le fonti informative relative ai progetti del programma Med, in modo tale da inserire più facilmente i dati nel Semantic Framework;



il modulo di gestione dei metadati potrebbe essere reso indipendente dal dominio, in modo tale da poter gestire i dati relativi alle collezioni indipendentemente dalla struttura del database nel quale sono memorizzati.

Grazie per l'attenzione!