28/11/2015

# OpenIdeas

Documento architetturale



Ciro Amati, Stefania Cardamone, Amedeo Leo, Simone Romano PROGETTO WEB SEMANTICO

# Contents

Introduzione	2
Analisi dei requisiti	3
Gestione Accessi	3
Gestione Utenti	3
Gestione Idee	3
Gestione notifiche e email	3
Requisiti funzionali	4
Requisiti non funzionali	5
Casi d'uso	6
Gestione accessi	6
Gestione utente	9
Gestione idea	13
Manuale utente	18
Architettura e tecnologie	25
Architettura del sistema	25
Gestione dati	26
Tecnologie utilizzate	27
Semantic Web	29
Sentiment analysis	29
Linked data	29
RDFa	29
Conclusioni	32
References	33

## Introduzione

Lo scopo del portale *OpenIdeas* è dare la possibilità ad utenti che hanno voglia di realizzare un progetto (di qualsiasi tipologia) di

- Renderlo pubblico
- Avere un feedback sul giudizio degli utenti
- Trovare eventuali finanziatori

Le sezioni seguenti sono così strutturate:

- La sezione Analisi dei requisiti spiega i possibili utilizzi del sistema tramite una serie di scenari reali;
- La sezione **Manuale utente** guida l'utente nell'utilizzo del sistema (dal login all'inserimento di una idea):
- La sezione **Architetture e tecnologie** illustra con una visione di alto livello la architettura del sistema focalizzandosi sulla parte di gestione dati;
- La sezione **Semantic Web** raccoglie i concetti principali della applicazione, ovvero quelli legati alle tecnologie del semantic web utilizzate;
- Infine, le opportune conclusioni.

# Analisi dei requisiti

## Gestione Accessi

Comprende le attività di gestione degli account per l'accesso al sistema. Gli utenti utilizzano la funzione di log-in per autenticarsi e la funzione log-out per uscire dal sistema. In particolare, la funzione di login può essere effettuata sia tramite le credenziali inseriti al momento della registrazione che tramite le API fornite da Facebook. La registrazione è possibile sia attraverso la compilazione dei campi richiesti sia con Facebook.

#### Gestione Utenti

Comprende le possibili attività degli utenti. Il sistema permette le seguenti azioni:

- Visualizzazione delle proprie informazioni.
- Visualizzazione delle idee.
- Visualizzazione delle migliori idee.
- Visualizzazione delle notifiche.
- Registrazione al sistema.

#### Gestione Idee

È il fulcro del sistema. Il sistema permette di gestire i seguenti aspetti:

- Inserimento di un'idea, con relative informazioni.
- Inserimento di un commento ad un'idea non proprietaria; il sistema ricaverà il punteggio tramite sentiment analysis.
- Possibilità di seguire un'idea per ricevere gli aggiornamenti.
- Possibilità di finanziare un'idea.
- Visualizzazione di tutte le idee.
- Visualizzazione delle informazioni di un'idea.
- Visualizzazione delle idee tramite query SPARQL.

#### Gestione notifiche e email

Il sistema permette all'utente di essere in continuo aggiornamento sui cambiamenti delle idee; in particolare:

- Se si ha commentato, si tiene traccia di commenti successivi al proprio commento.
- Se si ha scelto di seguire un'idea, si tiene traccia di nuovi commenti e finanziatore.
- Se si ha inserito un'idea, si tiene traccia di commenti, finanziatore e followers.
- All'atto della registrazione, all'utente viene inviata una email per conferma autenticazione.

## Requisiti funzionali

Di seguito verranno elencati i requisiti funzionali in dettaglio, suddivisi per attori, specificando tra parentesi la relativa priorità (*ALTA*: la funzione verrà implementata; *MEDIA*: la funzione verrà documentata e implementata a seconda del tempo a disposizione; *BASSA*: la funzione non verrà implementata).

#### Gestione accessi:

- o Tutti gli utenti
  - Login Facebook(ALTA): Ogni utente può effettuare l'accesso alla piattaforma tramite le proprie credenziali di Facebook.
  - Login (ALTA): Ogni utente può effettuare l'accesso alla piattaforma con le proprie credenziali.
  - Logout (ALTA): Ogni utente può effettuare il logout dalla piattaforma.

#### • Gestione utente:

- Tutti gli utenti
  - <u>Visualizzazione dati personali (ALTA):</u> Qualsiasi utente può visualizzare le proprie informazioni.
  - Registrazione al sistema (ALTA): Qualsiasi utente esterno può registrarsi al sistema.
  - <u>Visualizzazione delle idee (ALTA):</u> Qualsiasi utente può visualizzare ogni idea.
  - Modifica informazioni (BASSA): Qualsiasi utente può modificare le proprie informazioni personali.

#### • Gestione idee:

- o Tutti gli utenti:
  - Inserimento di un'idea (ALTA): Qualsiasi utente può inserire un'idea con relative informazioni.
  - <u>Inserimento di un commento (ALTA):</u> Qualsiasi utente può inserire un commento ad un'idea non proprietaria. Il sistema ricaverà il punteggio tramite sentiment analysis.
  - Seguire un'idea (ALTA): Qualsiasi utente può seguire un'idea non proprietaria.
  - Finanziare un'idea (ALTA): Qualsiasi utente può finanziare un'idea non proprietaria.
  - Visualizzazione di un'idea (ALTA): Qualsiasi utente può visualizzare ogni idea.
  - Modifica di un'idea (BASSA): L'autore può modificare la propria idea.

## Requisiti non funzionali

Il sistema vedrà come fondamentali i requisiti non funzionali di usabilità, performance, affidabilità e sicurezza:

- Usabilità: il sistema dovrà essere utilizzabile da un gruppo molto vasto di utenti.
- **Performance**: il sistema dovrà soddisfare le esigenze di ogni utente nel minor tempo possibile.
- Affidabilità: il sistema dovrà essere affidabile e mantenere inalterati tutti i dati contenuti in esso in caso di un funzionamento anomalo o di un errato utilizzo.
- **Sicurezza**: il sistema utilizzerà una funzione di autenticazione in modo da negare l'accesso da parte di utenti non registrati rispettando la tutela della privacy. È prevista una cifratura della password inserita all'atto della registrazione.

# Casi d'uso

## Gestione accessi

ID	UC_AC_01	
Preambolo	Preambolo	
L'utente si trova nella pagina	L'utente si trova nella pagina relativa al login con l'intenzione di accedere al sistema.	
Caso D'uso	Login	
Attori	Utente	
Pre-Condizioni	L'utente si trova nella schermata di login	
Punti di estensione		
Generalizzazione di		
Scenario Principale		
1	"OpenIdeas" mostra la pagina dove inserire i dati	
	necessari al login.	
2	L'utente inserisce i dati necessari:	
	• Email	
	Password	
	E clicca sul bottone per la conferma.	
3	"OpenIdeas" legge i dati dal database e mostra la pagina	
	iniziale.	
Scenario Secondario		
2.a	L'utente clicca sul bottone per accedere tramite	
	Facebook.	
3.a	"OpenIdeas" legge le informazioni dal database relative	
	alla mail di Facebook e mostra la pagina iniziale.	

ID	UC_AC_02
Preambolo	
L'utente si trova in qualsiasi pagina di "OpenIdeas" con l'intenzione di effettuare il	
logout.	
Caso D'uso	Logout
Attori	Utente
Pre-Condizioni	L'utente ha effettuato l'accesso al sistema.
Scenario Principale	
1	L'utente clicca sul bottone per effettuare il logout.
2	"OpenIdeas" procede al logout e visualizza la pagina di
	login.

ID	UC_AC_03	
Preambolo		
L'utente si trova nella pagin	L'utente si trova nella pagina di login di "OpenIdeas" con l'intenzione di registrarsi.	
Caso D'uso	Registrazione	
Attori	Utente	
Pre-Condizioni		
Scenario Principale		
1	L'utente inserisci i campi:	
	• Nome	
	• Cognome	
	• Email	
	• Password	
	Immagine profilo	
	Pagina web	
	Data di nascita	
	Quindi clicca su "Registrati".	
2	"OpenIdeas" procede con l'inserimento nel database	
	delle informazioni e visualizza la pagina iniziale.	
Scenario Secondario		
1.a	L'utente clicca sul bottone per effettuare il login con	
	Facebook.	
2.a	"OpenIdeas" procede con l'inserimento nel database	
	delle informazioni ottenute tramite le API di Facebook e	
	visualizza la pagina iniziale.	

## Gestione utente

ID	UC_GU_01
Preambolo	
L'utente si trova in qualsiasi pagina e vuole visualizzare le sue informazioni	
Caso D'uso	Visualizza informazioni
Attori	Utente
Pre-Condizioni	
Scenario Principale	
1	L'utente clicca sul proprio username. Quindi clicca su
	"Profilo".
2	"OpenIdeas" mostra le informazioni dell'utente.

ID	UC_GU_02
Preambolo	
L'utente si trova in qualsiasi pagina e vuole visualizzare la pagina di about del sistema.	
Caso D'uso	Visualizza informazioni sistema
Attori	Utente
Pre-Condizioni	
Scenario Principale	
1	L'utente clicca sul bottone per visualizzare le
	informazioni del sistema.
2	"OpenIdeas" mostra le informazioni del sistema.

ID	UC_GU_03
Preambolo	
L'utente si trova nella pagina iniziale e vuole visualizzare solo le idee di una	
determinata categoria	
Caso D'uso	Visualizza idee di una categoria
Attori	Utente
Pre-Condizioni	
Scenario Principale	
1	L'utente clicca sulla categoria "Software"
2	"OpenIdeas" mostra le idee afferenti alla categoria
	Software.

ID	UC_GU_04
Preambolo	
L'utente si trova in qualsiasi e vuole visualizzare tutte le idee	
Caso D'uso	Visualizza idee.
Attori	Utente
Pre-Condizioni	
Scenario Principale	
1	L'utente clicca su "OpenIdeas".
2	"OpenIdeas" mostra le idee inserite.

# Gestione idea

ID	UC_GI_01	
Preambolo	Preambolo	
L'utente si trova nella sezior	ne "Profilo" è vuole inserire un'idea.	
Caso D'uso	Inserimento idea	
Attori	Utente	
Pre-Condizioni		
Scenario Principale		
1	L'utente inserisce il nome della descrizione, quindi	
	clicca su "Step 2" e inserisce un'immagine e l'url del	
	video contenente un riepilogo della propria idea; infine	
	clicca su "Summary" e conferma.	
2	"OpenIdeas" inserisce l'idea nel database.	

ID	UC_GI_02
Preambolo	
L'utente si trova nell'idea "SmartAPI" e vuole inserire un commento.	
Caso D'uso	Inserimento commento
Attori	Utente
Pre-Condizioni	L'utente non è l'autore dell'idea
Scenario Principale	
1	L'utente inserisce il testo del commento, quindi clicca
	su "Inserisci Commento".
2	"OpenIdeas" inserisce il commento dell'utente nel
	database; quindi aggiorna i punteggi dell'idea, ottenuti
	tramite sentiment analysis.

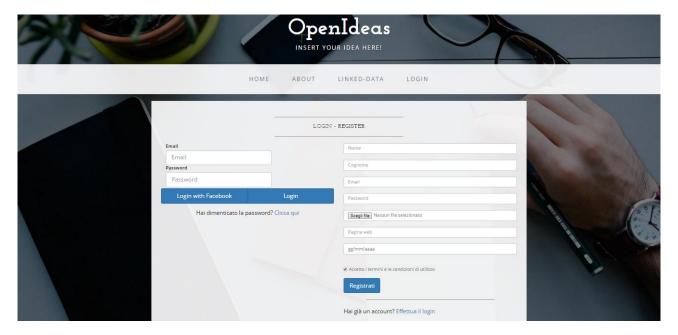
ID	UC_GI_03
Preambolo	
L'utente si trova nell'idea "SmartAPI" e vuole seguirla.	
Caso D'uso	Follow di idea
Attori	Utente
Pre-Condizioni	L'utente non è l'autore dell'idea
Scenario Principale	
1	L'utente clicca su "Segui".
2	"OpenIdeas" inserisce l'utente tra i follower dell'idea
	"SmartAPI" e lo informerà ad ogni successivo
	aggiornamento.

ID	UC_GI_04
Preambolo	
L'utente si trova nell'idea "SmartAPI" e vuole finanziarla.	
Caso D'uso	Finanziamento idea
Attori	Utente
Pre-Condizioni	L'utente non è l'autore dell'idea
Scenario Principale	
1	L'utente clicca su "Finanzia".
2	"OpenIdeas" assegna l'idea all'utente.

ID	UC_GI_05
Preambolo	
L'utente si trova in qualsiasi pagina e clicca su "Linked data".	
Caso D'uso	Query SPARQL
Attori	Utente
Pre-Condizioni	
Scenario Principale	
1	L'utente inserisce il testo della query. Quindi sceglie il
	formato in output desiderato e clicca su query.
2	"OpenIdeas" mostra all'utente il risultato della query e
	scarica il relativo file creato.

# Manuale utente

L'accesso al portale prevede una fase iniziale di login. Esso presenta una parte di registrazione ma anche un accesso immediato tramite Facebook.



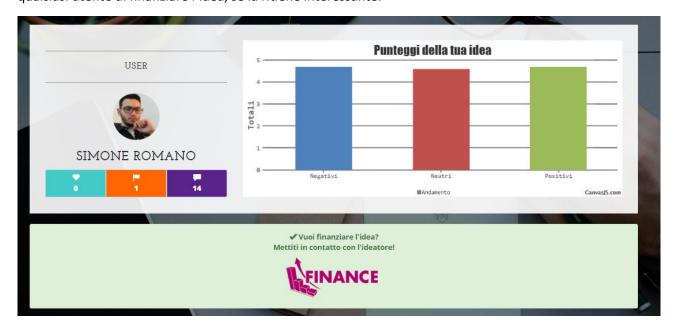
Nella Home, sarà possibile visualizzare tutte le idee inserite, ordinate per categoria. Sarà visibile la valutazione degli utenti e il numero di follower.

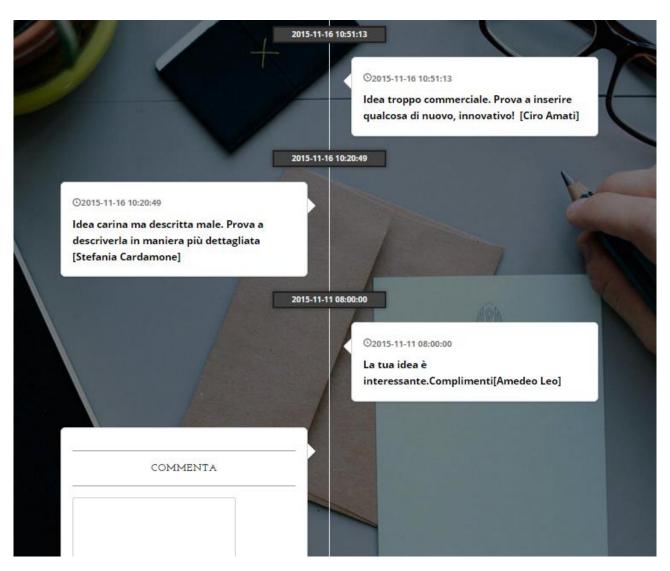


Vediamo in dettaglio la pagina dell'idea. Essa presenta una prima parte descrittiva dell'idea stessa. È possibile "seguirla" effettuando un click sull'apposito tasto "follow".

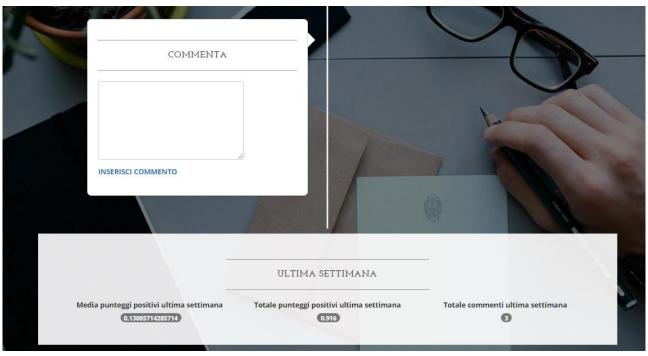


Grazie alla sentiment analysis è possibile visualizzare il punteggio dell'idea analizzando i commenti inseriti dagli utenti. Essi saranno classificati come "negativi", "neutri" e "positivi". Il portale, inoltre, permette ad un qualsiasi utente di finanziare l'idea, se la ritiene interessante.





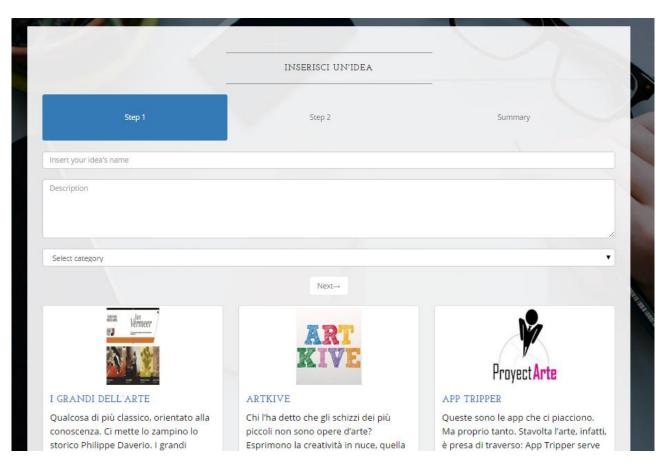
L'utente potrà esprimere una propria opinione commentando l'idea. I commenti, come detto in precedenza, saranno utili per la sentiment analysis.



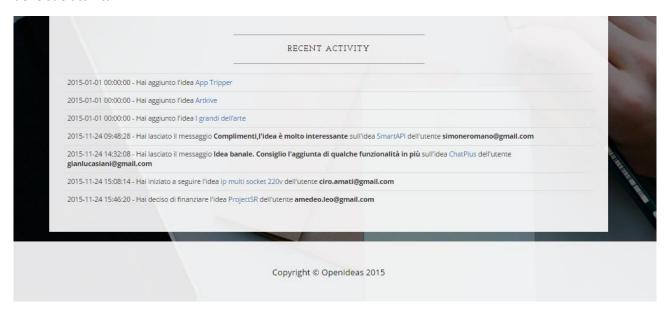
È prevista anche una sezione, all'interno della pagina Idea, in cui l'utente può visualizzare i dati relativi all'ultima settimana.

Profilo dell'utente: I dati relativi alle informazioni sull'idee dell'utente, possono essere visibili nella propria pagina. Prevede appunto una sezione per le notifiche e il conteggio delle attività.

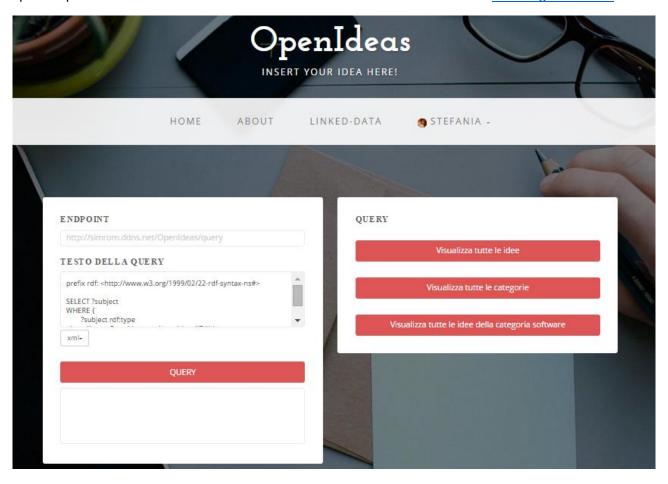




L'utente inserirà l'idea nella propria user page grazie ad una form dedicata. Visualizzerà, inoltre, l'history delle sue attività.



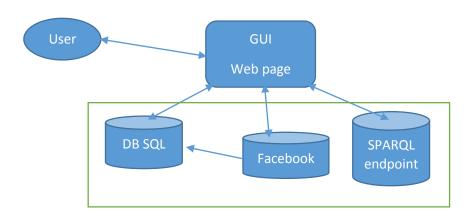
I portale presenta anche una sezione "Linked data" come descritta nella sezione <u>Tecnologie utilizzate</u>:



# Architettura e tecnologie

## Architettura del sistema

L'architettura del sistema è mostrata di seguito:



L'utente del sistema interagisce con una Web Application per visualizzare i contenuti o aggiungerne altri. L'applicazione utilizza più fonti per lo storage dei dati. In particolare è stato realizzato un database SQL per lo storage di tutte le informazioni (vedere sezione seguente). I dati relativi alle informazioni degli utenti possono essere popolati

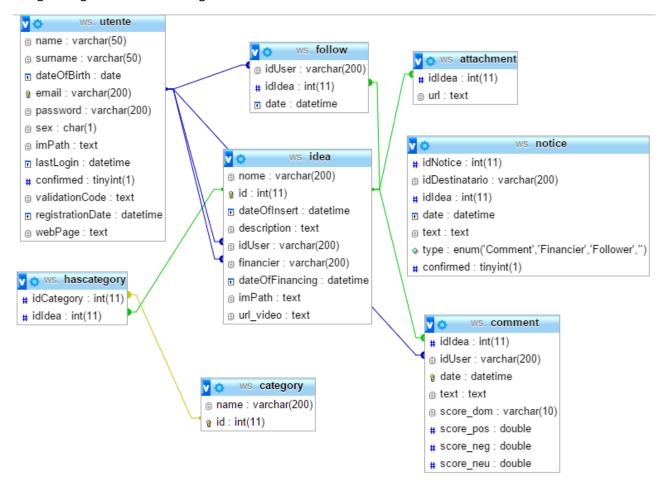
- Dall'utente effettuando una registrazione al sito
- Con le informazioni prelevate dall'account Facebook dell'utente (se l'utente utilizza la funzionalità login with facebook)

Inoltre, i dati relativi alle idee inserite nel sistema sono memorizzati sia nel database SQL che nel triple store appositamente realizzato. Quindi, ogni volta che un utente inserisce una nuova idea nel sistema, questa viene inserita nel database SQL e viene caricata nel grafo del dataset pubblico realizzato (vedieresezione Linked data). Il grafo RDF può essere interrogato tramite query SPARQL per ottenere le informazioni desiderate.

#### Gestione dati

#### Database Sal

La figura seguente mostra il diagramma ER della base di dati realizzata.



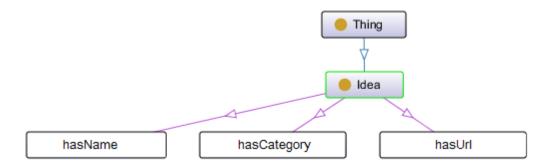
Di seguito è spiegata brevemente ciascuna tabella:

- utente: contiene le informazioni relative all'utente del sistema; tali informazioni potranno essere
  inserite dall'utente in fase di registrazione o potranno essere prelevate da facebook nel caso di login
  con facebook; la password dell'utente non è memorizzata in chiaro
- **follow:** tale tabella esprime la relazione *follow* presente tra un utente ed una idea; l'utente può *seguire* una idea per essere aggiornato sulle news (commenti, finanziatori)
- attachment: ogni idea avrà sicuramente una immagine ed eventualmente altri allegati; tale tabella lega gli allegati all'idea corrispondente
- idea: contiene le informazioni sull'idea
- notice: contiene le informazioni sulle notifiche per ogni utente; in particolare le tipologie di notifiche possono essere
  - o commento su una idea inserita dall'utente o su una idea a cui l'utente è interessato
  - un utente ha deciso di finanziare una tua idea
  - o un utente ha cominciato a seguire una tua idea
- hascategory: ogni idea appartiene ad una categoria; tale tabella lega una idea ad una categoria
- category: contiene le categorie
- **comment:** contiene le informazioni sui commenti inseriti dagli utenti e sull'output della *sentiment* analysis effettuata su di essi

## Triple-store

I dati principali presenti nel sistema sono quelli relativi alle idee inserite dall'utente. Tali dati, come visto nella sezione precedente, sono già raccolti nella tabella *idea* del database SQL. Oggigiorno si sta notevolmente diffondendo il concetto di *Linked-data*, che propone la pubblicazione di dati strutturati e collegati fra loro. È per questo motivo che è stato realizzato un endpoint SPARQL che contiene i dati essenziali dell'idea. Tale dataset è pubblico ed interrogabile tramite la sezione **Linked-data** presente sull'applicazione web con linguaggio **SPARQL**.

Di seguito è riportato il grafo della semplice ontologia realizzata



## Tecnologie utilizzate

Come già evidenziato nelle sezioni precedenti, il prodotto descritto consiste in una Web Application. Per lo sviluppo dell'applicazione si è scelto di utilizzare le seguenti teconologie/servizi:

- php
- html
- JavaScript
- jQuery
- SQL
- SPARQL
- Fuseki
- AJAX

Come già evidenziato precedentemente, i commenti inseriti dagli utenti per le varie idee sono utilizzati per abbinare un punteggio all'idea. Sono state utilizzate 2 librerie php per effettuare la sentiment analysis

- 1. google-translate-php
- 2. phpInsight

La seconda libreria permette, dato un testo in lingua inglese, di ottenere 3 punteggi relativi a

- 1. Positività
- 2. Neutralità
- 3. Negatività

del testo inserito. Dal momento che gli utenti possono essere di diversa nazionalità, il testo dei commenti viene prima tradotto con la prima libreria e poi dato in input alla seconda per la sentiment analysis.

È stato inoltre utilizzato l'**sdk di facebook** per consentire agli utenti di effettuare il login utilizzando le credenziali facebook. In questo caso vengono recuperate le informazioni necessarie a riempire la tabella *utente* del database dall'account facebook dell'utente e viene effettuata una registrazione (se l'email dell'utente non è già presente nel database) prima di effettuare il login.

Altra importante funzionalità introdotta nell'applicazione è la sezione *Linked-data*. Tale sezione espone un'interfaccia per interrogare la base ontologica mostrata nel paragrafo precedente tramite linguaggio SPARQL. Per realizzare tale funzionalità è stato utilizzato **Fuseki**, un web-server che consente di gestire dataset e grafi tramite protocollo http. Su tale web-server è stata caricata l'ontologia che viene interrogata tramite richieste http con la libreria curl. In particolare vengono inserite nuove idee ogni volta che l'utente ne aggiunge una dalla pagina personale e vengono interrogate quando l'utente utilizza l'interfaccia della sezione linked-data.

## Semantic Web

Una parte delle tecnologie utilizzate nel progetto puntano ad utilizzare alcune delle tecnologie del Semantic web. In particolare sono stati analizzati i commenti inseriti dagli utenti con tecniche di **sentiment analysis**, è stata inserita una sezione **linked-data** che consente di interrogare un triple-store con informazioni sulle idee inserite tramite SPARQL ed è stato aggiunto codice RDFa alle pagine html.

## Sentiment analysis

Il portale realizzato rappresenta una sorta di *vetrina* dove un utente può inserire delle idee personali. Ci sono almeno 2 motivi per cui dovrebbe rendere pubblica una propria idea:

- 1. Trovare un finanziatore.
- 2. Avere un feedback da parte degli utenti sull'idea.

Capire qual è la valutazione degli utenti è un aspetto molto importante per l'utente che pubblica una propria idea. Esistono diversi modi per conoscere la valutazione dell'utente; un caso classico è quello di inserire una barra numerata (es. da 1 a 10) e chiedere esplicitamente all'utente di votare l'idea.

La cosa più naturale è invece far inserire un commento in linguaggio naturale all'utente ed elaborarlo al fine di ricavare uno score (o una serie di score) che danno una informazione sul contenuto sentimentale del commento. Raccolti questi score è dunque possibile ricavare un punteggio e dunque una valutazione dell'idea. Il paragrafo <u>Tecnologie utilizzate</u> spiega come è stata implementata tale funzionalità.

## Linked data

Il concetto di Linked-data (dati collegati) è un concetto che si sta diffondendo molto sul web ed è attualmente una importante tematica di ricerca. I linked data rappresentano una modalità di pubblicazione di dati strutturati e collegati fra loro.

Nel caso specifico è stata creata una sezione del sito (Linked-data) che espone una interfaccia che consente di effettuare delle query in linguaggio SPARQL per ottenere informazioni sulle idee presenti nel database. Si può richiedere l'output in formato JSON o XML. Per i dettagli implementativi vedi il paragrafo <u>Tecnologie utilizzate</u>.

## **RDFa**

Le pagine Web scritte in XHTML contengono dati intrinsecamente strutturati: eventi di calendario, informazioni anagrafiche, didascalie di foto, titoli di canzoni, informazioni sul copyright, etc. Se gli autori delle pagine sono in grado di esprimere questi dati in maniera precisa, e i tool possono leggerli con affidabilità, si possono immaginare scenari di utilizzo del tutto nuovi. I dati strutturati possono essere trasferiti dal web alle applicazioni, in modo che, per esempio, i dati di calendario, o i diritti della licenza, possono essere comunicati direttamente all'utente, per essere registrati nella sua agenda personale o per informarlo dei suoi diritti. I metadati possono essere inclusi nelle pagine HTML in vari modi (nel nostro caso, tramite RDFa).

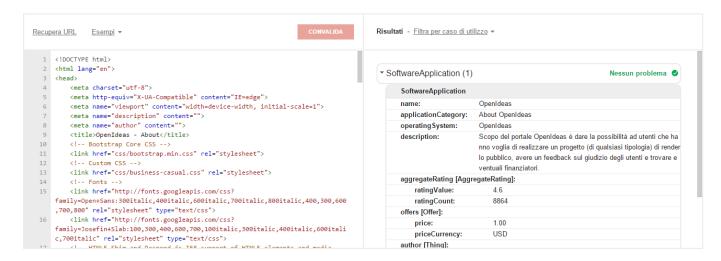
Come è possibile vedere dall'immagine, il prefix = "oi: http://www.OpenIdeas.org/openIdeasKB" specifica il vocabolario OpenIdeas. Questo vocabolario è utilizzato negli attributi "property" per specificare le proprietà del documento o gli elementi contenuti nel documento. Molti di questi vocabolari possono essere inclusi in un unico documento.

Grazie all'attributo "property" di RDFa, è possibile specificare: il nome del progetto, la categoria e l'url in cui risiede la risorsa.

L'output finale sarà il seguente:



Per la sezione "About", è utilizzato il vocabolario standard *vocab=http://schema.org/* specificando il tipo di Rich Snippet *typeof="SoftwareApplication"*. Grazie a un tool di Google, può essere testato il codice HTML.



## Conclusioni

Il sistema proposto raggiunge gli obiettivi fissati in fase di analisi dei requisiti. In particolare consente all'utente di inserire delle idee una volta effettuata la registrazione al portale e consente di visualizzare/commentare/finanziare altre idee presenti sul portale.

Una delle possibili estensioni è quella di consentire agli utenti di interagire tra loro utilizzando il portale (es. tramite chat private).

La sezione linked-data contiene una innovativa interfaccia verso alcuni dei dati e consente di ottenere alcune importanti informazioni sulle idee presenti nel portale. Tali informazioni sono strutturate secondo una semplice ontologia creata ad-hoc per il contesto applicativo. Sviluppi futuri prevedono la estensione di tale knowledge-base per consentire all'utente di ricavare ulteriori informazioni. Eventualmente sarebbe possibile introdurre alcune relazioni che condurrebbero a sistemi di inferenza.

## References

Amati Ciro, C. S. (n.d.).

https://drive.google.com/folderview?id=0ByKArIYxafLqR2N4RU5QSIUtdUE&usp=drive\_web&tid=0ByKArIYxafLqaDl1WkFENGE4cmc.

bootstrap. (n.d.). http://getbootstrap.com/.

Fuseki. (n.d.). https://jena.apache.org/documentation/serving\_data/.

google-translate. (n.d.). https://github.com/Stichoza/google-translate-php.

phpInsight. (n.d.). https://github.com/JWHennessey/phpInsight.

Rdfa. (n.d.). http://www.w3.org/TR/xhtml-rdfa-primer/.

Romano, S., & Leo, A. (n.d.). Fuseki: serving RDF data over HTTP.

https://drive.google.com/folderview?id=0ByKArIYxafLqR2N4RU5QSIUtdUE&usp=drive\_web&tid=0ByKArIYxafLqaDl1WkFENGE4cmc.

Schema.org. (n.d.). https://developers.google.com/structured-data/schema-org.