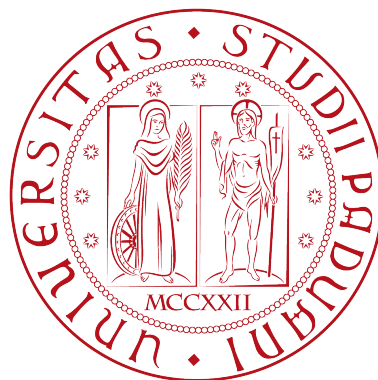


TESI DI LAUREA IN INFORMATICA

NICOLA MORETTO



Analisi e progettazione di un'interfaccia grafica per la  
consultazione dei contenuti informativi in una piattaforma  
web tematica

30 novembre 2012

*Tesi di Laurea in Informatica*

Analisi e progettazione di un'interfaccia grafica per la consultazione  
dei contenuti informativi in una piattaforma web tematica

© 30 novembre 2012

STUDENTE:

Nicola Moretto (matr. 578258)

RELATORE:

Prof. Tullio Vardanega

LUOGO:

Padova

DATA:

30 novembre 2012

*Ad Alessandro, Christian, Ester, Marta e Romeo con cui tutto ha avuto  
inizio.*

*Al dott. Bovo per la passione con cui mi ha seguito.*

*Al prof. Vardanega per la disponibilità con cui mi ha seguito.*

*Alla mia famiglia che mi ha accompagnato sin qui.*

## SOMMARIO

---

L'attività di stage si è svolta presso l'azienda *Sintesi Sas*, che opera nel settore ICT (*Information and Communication Technology*) realizzando software ERP e piattaforme Web per aziende, in particolare attive nel settore turistico, e fornendo servizi di consulenza e di formazione di imprenditori nell'ambito del marketing strategico, operativo e del controllo di gestione.

Il prodotto di punta dell'azienda - *Planet Hotel* - costituisce un sistema software per la gestione alberghiera tra i più flessibili, ampi e completi presenti nel panorama italiano, in grado di coprire la maggior parte delle necessità aziendali: oltre alla gestione delle prenotazioni e dei conti, esso offre un insieme di moduli integrati per supportare il controllo di gestione e degli interventi di marketing.

Si tratta di una realtà imprenditoriale a clientela nazionale con sede unica a Mestre (VE), la cui direzione e amministrazione è affidata al solo fondatore, che ha assunto il ruolo di tutor esterno e referente aziendale per l'intera durata dello stage.

Le attività svolte si inseriscono nell'ambito di un progetto esterno rispetto al business dell'azienda, finalizzato alla realizzazione di una piattaforma web tematica per la condivisione di informazioni e la vendita diretta di prodotti alla clientela e affidato ad un team costituito da differenti figure professionali (sociologi, informatici, ingegneri, ...).

## CONTENUTI

Il presente documento costituisce una relazione dettagliata in merito all'attività di stage svolta dallo studente Nicola Moretto presso l'azienda *Sintesi Sas*. I contenuti sono organizzati nei seguenti capitoli:

### PROGETTO

Il primo capitolo illustra le strategie dell'azienda e gli obiettivi, i requisiti e i vincoli del progetto in cui si inseriscono le attività di stage.

### STAGE

Il secondo capitolo illustra gli obiettivi, i requisiti e l'organizzazione (piano e norme di lavoro) delle attività di stage. A seguire vengono presentate le scelte più rilevanti effettuate e i risultati conseguiti.

## VALUTAZIONI FINALI

Il terzo capitolo presenta un'analisi critica a posteriori dell'attività di stage: raggiungimento degli obiettivi prefissati, competenze professionali acquisite, ... .

## CONVENZIONI TIPOGRAFICHE

Al fine di agevolare la consultazione del documento sono state adottate alcune convenzioni tipografiche illustrate di seguito.

**GLOSSARIO** Gli acronimi, le abbreviazioni, i nomi propri e i termini specialistici contenuti nel presente documento sono illustrati nel **GLOSSARIO**, consultabile in appendice, al fine di agevolare la lettura e la comprensione degli argomenti trattati. La prima occorrenza di ciascun termine o espressione presente nel glossario appare sottolineata.

**TERMINOLOGIA** I termini propri o di provenienza straniera divenuti di uso corrente nella lingua italiana sono evidenziati in *corsivo*, mentre la prima occorrenza di parole o espressioni che assumono particolare significato nel presente contesto è riportata in **MAIUSCOLETTO**.

**CODICE E FORMULE** I nomi di tabelle, classi, package, ... impiegano uno stile di carattere sans serif, mentre i frammenti di codice impiegano un carattere a spaziatura fissa.

## INDICE

---

1	PROGETTO	8
1.1	Genesi	8
1.2	Reti sociali	8
1.3	Architettura	9
1.3.1	Contenuti informativi	10
1.3.2	Sistema di classificazione	13
1.4	Requisiti e vincoli	14
1.4.1	Criterio di classificazione	14
1.4.2	Interfaccia grafica	15
2	STAGE	16
2.1	Piano di stage	16
2.1.1	Obiettivi e requisiti	16
2.1.2	Pianificazione	18
2.2	Norme di stage	18
2.3	Criterio di classificazione	21
2.3.1	Enciclopedia del sapere	22
2.3.2	Ambiguità sintattica	23
2.3.3	Ambiguità semantica	26
2.3.4	Modello relazionale	28
3	VALUTAZIONI FINALI	31
A	GLOSSARIO	32

## ELENCO DELLE FIGURE

---

Figura 1	Gerarchia degli utenti nelle piattaforme web tradizionali	10
Figura 2	Struttura di un contenuto informativo	12
Figura 3	Diagramma di Gantt	19
Figura 4	Valore informativo di un contenuto	21
Figura 5	Relazioni tra le entità del dominio	23
Figura 6	Ambiguità sintattica di un'entità	24
Figura 7	Etichetta primaria e secondarie	26
Figura 8	Accezioni di un'etichetta (+ modello relazionale)	26
Figura 9	Pagina di disambiguazione [2]	26
Figura 10	Accezioni chiave e sinonimiche	27
Figura 11	Modello relazione del criterio di classificazione	28
Figura 12	Modello ad oggetti delle accezioni	29

## ELENCO DELLE TABELLE

---

Tabella 1	Pianificazione settimanale delle attività	18
Tabella 2	Configurazione dell'ambiente di lavoro	19

## PROGETTO

---

### 1.1 GENESI

L'idea della piattaforma *Social (Life) Shuttle* nasce nel 2010 da un progetto concepito per dar vita ad una comunità virtuale destinata agli artisti sconosciuti e accessibile in mobilità mediante un'applicazione dedicata, *ArtYR*.

Nello stesso periodo una consulenza nell'ambito dei sistemi informativi territoriali ad un'azienda di Bolzano conduce allo sviluppo di un'innovativa piattaforma software: un sistema informativo territoriale in cui l'erogazione di informazioni turistiche è integrata con la vendita di servizi collegati.

Il progetto evolve - grazie alla partecipazione di Comuni, Province e Regioni - in una rete tematica di agenzie di viaggio con un'identità comune e finalizzata alla fusione dei sistemi informativi distrettuali e di vendita.

L'architettura di *Social (Life) Shuttle* trae profonda ispirazione, integrando tre componenti differenti:

#### BUSINESS

Vendita diretta di prodotti alla clientela.

#### SOCIALE

Creazione e sviluppo delle relazioni sociali attraverso la condivisione di informazioni e conoscenza.

#### TERRITORIO

Sistema di erogazione di informazioni turistiche e territoriali.

### 1.2 RETI SOCIALI

Il modello sociologico di rete sociale non ha attualmente riscontro presso le piattaforme web di condivisione dei contenuti (*blog, forum, ...*) o i *social network* (*Facebook, Twitter, ...*), che si limitano a considerare e concretizzarne singoli aspetti.

Nelle moderne reti sociali è infatti assente l'incentivo alla condivisione e distribuzione della conoscenza, fattore cruciale per l'aggregazione fisica dei membri delle comunità, da intendersi a sua volta



come aggregazioni formatesi intorno ed attraverso la manifestazione di interesse nei confronti di uno specifico tema di dialogo o discussione, che attraversa la sfera individuale, intima e personale dei suoi membri.

Il progetto *Social (Life) Shuttle* rappresenta una nuova generazione di piattaforma di socializzazione, in cui il web diventa solamente un canale di condivisione e un serbatoio della conoscenza generata dalla dialettica tra persone e dove vengono integrati i canoni classici di *blog*, *forum*, *social network* e *media*.

Una relazione sociale nata e costruita su un interesse comune stravolge l'attuale paradigma delle reti sociali virtuali, in cui il legame nasce a prescindere dalla presenza di interessi comuni o informazioni da condividere, e favorisce l'incontro tra persone aventi esperienze simili frutto di tali interessi condivisi.

Ove l'esperienza riguarda anche beni o prodotti, la componente *business* intende offrire ai membri la possibilità di interagire con i produttori, anch'essi attori della comunità.

L'architettura di *Social (Life) Shuttle* consente di declinare la piattaforma in innumerevoli varianti, applicabili ai temi più svariati: al momento sono in fase di sperimentazione per il mondo del vino, il cibo biologico, l'arte commercializzabile e l'attività di ricerca e progettazione collaborativa.

### 1.3 ARCHITETTURA

Tale piattaforma presenta numerose aspetti che la differenziano dalla concorrenza attuale:

- profonda integrazione degli aspetti *social* e *business*;
- nessuna distinzione tra creatori e fruitori dei contenuti (ciascun membro può condividere le proprie esperienze, segnalare eventi, pubblicare articoli critici, ...);
- l'autorevolezza di ciascun membro della comunità si rafforza o si indebolisce a seconda della qualità dei contenuti pubblicati, dei giudizi degli altri membri e di altri criteri di valutazione;
- lo sfruttamento di tecnologie e dispositivi mobili per favorire la crescita di relazioni al di fuori dell'ambito virtuale della piattaforma (partecipazione ad eventi, raccolta e condivisione di informazioni geolocalizzate, ...).

Per quanto concerne le attività di stage, due aspetti della piattaforma assumono particolare rilevanza: i contenuti informativi e i relativi criteri di classificazione.

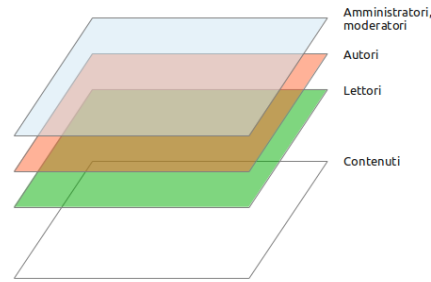


Figura 1: Gerarchia degli utenti nelle piattaforme web tradizionali

### 1.3.1 *Contenuti informativi*

I contenuti informativi rappresentano lo strumento essenziale per la condivisione delle esperienze e della conoscenza intorno al tema specifico della piattaforma.

Per individuare le classi di contenuti adatte a esprimere in una forma strutturata le informazioni si è tratta ispirazione dalle forme espressive e comunicative tipiche della dialettica quotidiana, poiché immediatamente e intuitivamente comprensibili agli utenti, a prescindere dal loro livello di esperienza.

In particolare, si distinguono la natura della comunicazione, connessa allo scopo e al tono con cui ci esprimiamo, e il formato delle informazioni, che dipendono strettamente dai sensi e dai canali di comunicazione a disposizione per scambiare informazioni con l'interlocutore, sia esso un individuo singolo o un gruppo.

**CLASSI** I tipi di contenuto pubblicabili nella piattaforma dovrebbero essere in numero adeguato a coprire il maggior numero possibile di esigenze comunicative pur rimanendo facilmente e intuitivamente distinguibili, ossia l'utente non dovrebbe nutrire dubbi circa il più adatto a formalizzare di volta in volta l'informazione che desidera condividere.

#### **DOMANDA**

La domanda classica rende particolarmente esplicito lo scopo della comunicazione, ossia la richiesta di informazioni di varia natura agli altri utenti della piattaforma. Si distingue in pubblica o privata, a seconda che l'utente desideri rivolgerla ad un particolare sottoinsieme di utenti.

#### **RISPOSTA**

Duale della domanda, la risposta è anch'essa in forma pubblica o privata per consentire all'utente di renderla accessibile e consultabile solo a certi utenti, spesso l'autore della domanda a cui risponde.

**PENSIERINO**

Il pensierino rappresenta una forma di comunicazione adatta ad esprimere un contenuto di lunghezza breve e prevalentemente superficiale (considerazioni, stati d'animo, freddure, ...).

**EVENTO**

L'evento aiuta a promuovere qualsiasi iniziativa che rientri nell'ambito tematico della piattaforma e cui possano prender parte altre persone (incontro pubblico, concerto, fiera, ...).

**DISCORSO**

Il discorso identifica un contenuto articolato, sia nella forma sia nei contenuti, destinato alla condivisione di informazioni dettagliate e approfondite.

**RECENSIONE**

La recensione esprime un giudizio critico nei confronti di un prodotto specifico.

**COMUNICAZIONE PRIVATA**

La comunicazione privata è l'unica forma di contatto diretto e riservato tra due utenti.

**ELEMENTI** Ciascun tipo di contenuto esprime un intento comunicativo ben preciso, ma non è vincolato ad una struttura e ad un formato predefiniti: la classe, che esprime l'intento della comunicazione, si colloca in un piano distinto rispetto al formato, ossia la struttura e le caratteristiche specifiche del contenuto informativo condiviso.

Ove la dialettica quotidiana dispone infatti di cinque sensi e può esprimersi in forma non solo verbale, nel web gli utenti sperimentano differenti forme di comunicazione: contenuti testuali e grafici, flussi audio e video, documenti elettronici, messaggistica istantanea, ...

I contenuti informativi non presentano dunque una struttura fissa a seconda della classe, ma possono essere liberamente redatti a partire da una serie di elementi predefiniti, frutto di una ricerca tra le principali e più diffuse piattaforme web disponibili (*blog, forum, social network, chat, ...*) e di una successiva analisi e rielaborazione dei risultati ottenuti:

**AUDIO**

Contenuto audio statico o in tempo reale (*live streaming, ...*).

**IMMAGINI**

Contenuto grafico statico.

**VIDEO**

Contenuto video statico o in tempo reale (*live streaming, ...*).

**SONDAGGIO**

Domanda a risposta multipla.

**DOCUMENTO**

File di testo o binario caricato nella piattaforma.

**STRINGA**

Contenuto testuale avanzato (intestazioni, formattazione dei caratteri, collegamenti ipertestuali, ...).

**CITAZIONE**

Citazioni o riferimenti ad altri elementi di un contenuto, ad un contenuto informativo o a prodotti presenti nella piattaforma.



Figura 2: Struttura di un contenuto informativo

La struttura modulare dei contenuti informativi consente di riusare, riferire o citare gli elementi costituenti e di catalogarli con maggior facilità e precisione, riuscendo a classificare ciascun frammento di informazione presente al loro interno.

**PROPRIETÀ** I contenuti informativi - a prescindere dalla classe e dalla struttura - presentano un insieme di proprietà comuni, alcune delle quali assumono particolare rilevanza per l'attività di stage, rappresentando utili criteri aggiuntivi per filtrare i contenuti informativi durante una ricerca: si tratta di AUTORE, DATA DI PUBBLICAZIONE e TIPO del contenuto.

**RELAZIONI** Ove tradizionalmente ci si affida ai commenti per consentire agli utenti di esprimere un'opinione rispetto alle informazioni riportate o alle posizioni espresse in un contenuto, in *Social (Life) Shuttle* si permette di rispondere ad un contenuto pubblicato nella piattaforma direttamente con altri contenuti, in numero arbitrario.

La relazione di dipendenza tra i contenuti prescinde dalla classe specifica, non ponendo vincoli di alcun genere circa la classe ed il formato della risposta ad un contenuto informativo.

Ciò consente maggiore libertà all'utente nello scegliere la forma espressiva più adeguata per condividere il proprio messaggio, ne facilita la

catalogazione e allo stesso tempo rispecchia il principio di uguaglianza tra gli utenti espresso in precedenza ed elemento cardine della piattaforma.

Nel corso del tempo a partire da ciascun contenuto informativo possono così svilupparsi e ramificarsi diverse **DISCUSSIONI**, senza limiti di ampiezza o profondità.

### 1.3.2 *Sistema di classificazione*

Il sistema di classificazione consiste in un insieme di criteri che associano a ciascun contenuto alcuni metadati, in grado di fornire agli utenti della piattaforma informazioni utili a contestualizzarlo, ad interpretarlo e a valutarne l'interesse soggettivo:

#### ARGOMENTO

L'argomento di un contenuto rappresenta la branca del sapere - agnostica rispetto al tema specifico della piattaforma - cui appartiene.

#### EMOZIONI

Le emozioni indicano lo stato d'animo con cui un contenuto sia stato pubblicato dall'autore.

#### GIUDIZI

I giudizi forniscono una valutazione qualitativa sul contenuto e sono espressi dagli utenti.

#### INTENZIONI

Le intenzioni indicano lo spirito con cui l'autore redige il contenuto (opinione, critica, ...).

#### INTERESSI

Gli interessi rappresentano temi specifici della piattaforma nei confronti dei quali ciascun utente registrato dichiara di nutrire passione.

Queste meta-informazioni assumono particolare rilevanza nel processo di ricerca di informazioni all'interno della piattaforma, poiché consentono di escludere o meno determinati contenuti dai risultati.

Tra quelli evidenziati, tuttavia, spicca l'assenza di un criterio in grado di catalogare ordinatamente le informazioni presenti nei contenuti per facilitarne la ricerca, il reperimento e la consultazione: il primo obiettivo dell'attività di stage consiste nell'individuare un meccanismo efficiente per rendere più agevole la consultazione della conoscenza custodita nella piattaforma, fornendo un livello di astrazione rispetto alla semplice enumerazione dei contenuti.

## 1.4 REQUISITI E VINCOLI

Durante gli incontri preliminari all'attività di stage sono stati fissati gli obiettivi, i requisiti ed i vincoli concernenti le attività previste ed i prodotti attesi.

### 1.4.1 *Criterio di classificazione*

La progettazione del criterio di classificazione deve tenere conto di alcuni vincoli e requisiti riguardanti l'architettura della piattaforma e l'integrazione con il sistema di classificazione:

#### INDIPENDENZA DAI CRITERI ESISTENTI

Il criterio deve minimizzare il grado di accoppiamento per risultare facilmente mantenibile e aggiornabile senza intaccare lo stato, l'integrità e le funzionalità dei rimanenti e deve tenere conto di possibili evoluzioni della piattaforma, che comportino l'aggiornamento o la rimozione dei criteri esistenti o l'aggiunta di nuovi.

#### INDIPENDENZA DALLE CLASSI DI CONTENUTI

Il criterio non deve distinguere tra contenuti di classi diverse, ma deve considerare esclusivamente le proprietà e le relazioni definite o definibili sul contenuto generico.

#### INDIPENDENZA DALLE COMPONENTI DEL SISTEMA

Il criterio deve minimizzare le dipendenze e l'accoppiamento con le altre componenti del sistema, che possono essere soggette ad aggiornamenti sostanziali (in particolare quelle di terze parti) o interventi di manutenzione evolutiva.

#### MODULARITÀ

Il criterio dev'essere progettato in modo tale da potersi avvantaggiare - in futuro - di soluzioni tecniche o tecnologiche in grado di automatizzare (in parte o del tutto) le operazioni di classificazione dei contenuti.

Le potenziali criticità tecniche, legate all'implementazione del criterio di classificazione, devono essere raccolte e condivise con il team di progetto, che provvederà a valutarle, a fornire eventuali indicazioni e ad individuare le soluzioni ritenute appropriate e compatibili con l'architettura e le specifiche della piattaforma.

#### 1.4.2 *Interfaccia grafica*

L'interfaccia grafica per la consultazione dei risultati di una ricerca sui contenuti informativi deve soddisfare alcuni requisiti essenziali:

- deve consentire all'utente di inserire dei termini di ricerca e selezionare un ambito;
- deve potersi interfacciare a componenti terze per ottenere i risultati di ricerca;
- deve mostrare i contenuti con forme geometriche elementari, che permettano di distinguerne intuitivamente la classe di appartenenza;
- deve permettere la consultazione della DISCUSSIONE associata ad un contenuto, evidenziando il flusso informativo (le sequenze di risposte ad un contenuto);
- dovrebbe gestire dei filtri basati sugli interessi ed il livello di esperienza di un utente registrato;
- dev'essere in grado di visualizzare ordinatamente un numero elevato di risultati di ricerca, evitando un sovraccarico cognitivo;
- dev'essere utilizzabile agevolmente da utenti con differenti livelli di esperienza;
- dev'essere adeguatamente fruibile su dispositivi mobili.

## STAGE

---

### 2.1 PIANO DI STAGE

#### 2.1.1 *Obiettivi e requisiti*

L'attività di stage si colloca nell'ambito del progetto presentato nel capitolo 1 e persegue due obiettivi distinti ma correlati, focalizzandosi sulla componente *social* della piattaforma.

**CRITERIO DI CLASSIFICAZIONE** Il primo obiettivo consiste nell'estendere l'attuale sistema di classificazione (v. sezione 1.3.2), integrandovi un criterio aggiuntivo per la catalogazione dei contenuti pubblicati dagli utenti e la costruzione di un'enciclopedia della conoscenza, al fine di rendere il reperimento e la consultazione delle informazioni desiderate il più efficiente ed agevole possibile.

L'ideazione e concezione del suddetto criterio deve tener conto della natura tematica della piattaforma, conciliando due esigenze distinte:

- dev'essere sufficientemente astratto e flessibile per adattarsi alla molteplicità di varianti tematiche in cui la piattaforma stessa può essere declinata;
- deve cercare di avvantaggiarsi delle peculiarità di una piattaforma tematica, ad esempio la maggior correlazione degli argomenti trattati.

La soluzione individuata deve inoltre prescindere da assunzioni legate alla tecnologia utilizzata.

Alla luce di possibili evoluzioni nello sviluppo della piattaforma, si desidera infine che la classificazione di un contenuto informativo (assegnazione di metadati, individuazione di correlazioni, ...) possa essere - in futuro - demandata a componenti software integrate nella piattaforma.

**INTERFACCIA GRAFICA** Il secondo obiettivo consiste nel progettare un'interfaccia grafica per la consultazione dei contenuti informativi, che sfrutti il criterio di classificazione aggiuntivo per facilitare la ricerca ed il reperimento delle informazioni di interesse per l'utente all'interno del patrimonio enciclopedico della piattaforma.



La sfida principale consiste nel progettare un'interfaccia in grado di visualizzare in maniera chiara e ordinata un numero arbitrario di contenuti, a prescindere dalla classe del dispositivo impiegato (*smartphone, tablet, notebook, ...*).

Il primo passo consiste nell'individuare le informazioni essenziali per una rapida e precisa identificazione dei contenuti (titolo, autore, data di pubblicazione, ...) e nel valutare la notazione (grafica o testuale) più adatta per esprimerle, al fine di renderle accessibili al maggior numero possibile di utenti; le informazioni aggiuntive devono risultare comunque accessibili, ma solo su esplicita richiesta dell'utente.

In questo frangente si inseriscono alcune analisi e valutazioni di carattere sociologico, svolte da altri membri del team di progetto, per individuare le soluzioni più idonee a comunicare tali informazioni in modo da renderne la comprensione chiara e intuitiva a qualsiasi utente.

Il secondo passo richiede di definire le specifiche per un'interfaccia facilmente navigabile, che sia in grado di mostrare in modo ordinato e intuitivo i contenuti e le reciproche relazioni. Occorre individuare opportuni criteri di raggruppamento, ordinamento e collocazione dei contenuti visualizzati per favorirne la consultazione, evitando un sovraccarico cognitivo e garantendo un livello adeguato di leggibilità.

Con il terzo ed ultimo passo si mira ad aggiungere la possibilità per l'utente di filtrare i contenuti visualizzati in accordo a proprietà (argomento, autore, data di pubblicazione, tipo) o metadati associati (emozioni, giudizi, intenzioni, ...).

Per gli utenti autenticati si desidera offrire un livello aggiuntivo di personalizzazione, che consenta di filtrare automaticamente i contenuti secondo le preferenze associate al profilo (interessi, livello di esperienza).

Per individuare i requisiti essenziali si prendono in considerazione alcuni casi d'uso:

1. l'utente naviga liberamente tra i contenuti (più recenti, più letti, più discussi, ...);
2. l'utente consulta la discussione generata da un singolo contenuto;
3. l'utente cerca le informazioni riguardanti un certo tema;
4. l'utente esplora gli argomenti trattati e le reciproche relazioni.

### 2.1.2 Pianificazione

L'attività di stage viene suddivisa in due fasi distinte per semplificarne la pianificazione:

1. l'estensione del sistema di classificazione;
2. l'analisi e la progettazione dell'interfaccia grafica.

Per ciascuna fase sono fissati gli obiettivi generali, sono individuate e organizzate su base settimanale le attività da svolgere, cercando di garantire un carico di lavoro equilibrato, e sono indicati i prodotti attesi.

La durata complessiva dello stage si attesta su 8 settimane a tempo pieno, corrispondenti a 320 ore di lavoro.

ATTIVITÀ	ORE
<i>Fase 1: estensione del sistema di classificazione</i>	
Analisi delle specifiche del sistema di classificazione	40
Analisi comparativa dei principali sistemi di classificazione della conoscenza	40
Progettazione del sistema di classificazione	40
Implementazione del sistema di classificazione nel modello relazionale	40
<i>Fase 2: analisi e progettazione dell'interfaccia grafica</i>	
Analisi dei requisiti dell'interfaccia grafica	40
Progettazione dell'interfaccia grafica: visualizzazione dei contenuti	40
Progettazione dell'interfaccia grafica: filtraggio dei contenuti	40
Progettazione dell'interfaccia grafica: navigazione dei contenuti	40

Tabella 1: Pianificazione settimanale delle attività

## 2.2 NORME DI STAGE

Nel corso dello stage sono stati impiegati diversi strumenti per gestire le attività di progetto e produrre la documentazione prevista.



in lingua italiana e contengono solo caratteri alfabetici in formato CamelCase, ove la lettera iniziale è sempre in minuscolo.

**DIGRAMMI UML** Durante l'attività di stage sono stati redatti e inclusi nella documentazione diversi diagrammi UML dei casi d'uso, dei package e delle classi, di cui sono presenti svariati esempi nel prosieguo del documento.

I nomi delle sottoclassi riportano - per esteso o in forma abbreviata - l'identificatore della superclasse diretta: nel secondo caso sono presenti - come prefisso - le sole lettere maiuscole, nel medesimo ordine di apparizione.

**CASI D'USO** La notazione utilizzata per identificare un caso d'uso è così definita:

$$UC.x.y$$

ove:

- UC è l'abbreviazione di *Use Case* (Caso d'uso);
- $x \in \{1, 2, \dots\}$  è il numero identificativo del diagramma cui appartiene il caso d'uso;
- $y \in \{1, 2, \dots\}$  è il numero associato al caso d'uso.

**REQUISITI FUNZIONALI** I requisiti del sistema software sono univocamente identificati mediante la seguente notazione:

$$Rf.x.y$$

ove:

- Rf è l'abbreviazione di *requisito funzionale*;
- $x \in \{ob, de\}$  rappresenta il tipo di requisito funzionale (*ob* per obbligatorio, *de* per desiderabile);
- $y \in \{1, 2, \dots\}$  è il numero associato ad un requisito.

**TRACCIAMENTO DEI CASI D'USO** Il tracciamento delle dipendenze tra casi d'uso e requisiti software è realizzato mediante un foglio elettronico, ove:

- ciascuna riga rappresenta un requisito del sistema software;
- ciascuna colonna rappresenta un caso d'uso;
- ciascuna cella contiene il carattere 'X' se esiste una relazione di dipendenza tra il caso d'uso e il requisito, altrimenti è vuota.

Per ciascuna riga e colonna viene impiegata una semplice formula per asserire la completezza e la necessità della matrice dei requisiti:

CONTA.SE(A:Z;"X")

ove:

- A:Z corrisponde all'intervallo di celle di una singola riga o colonna;
- X rappresenta la stringa da cercare;
- CONTA.SE è una funzione che accetta due argomenti (l'intervallo di celle ed il *pattern*) e restituisce il numero di celle appartenenti all'intervallo contenenti una o più occorrenze del *pattern* specificato.

**COMPLETEZZA** Per ogni colonna, se la formula restituisce un valore pari a o (zero) sta ad indicare che il requisito utente non è soddisfatto da alcun requisito software.

**NECESSITÀ** Per ogni riga, se la formula restituisce un valore pari a o (zero) sta ad indicare che il requisito software corrispondente è superfluo.

## 2.3 CRITERIO DI CLASSIFICAZIONE

Il patrimonio di conoscenza della piattaforma è garantito essenzialmente dai contenuti pubblicati dagli utenti ed arricchito dal loro valore informativo: ciascuno di essi, a prescindere dalla forma (testo, immagini, audio, video, ...) o dalla classe (domanda, discorso, evento, recensione, ...), condivide delle informazioni inerenti uno o più elementi del dominio tematico della piattaforma.

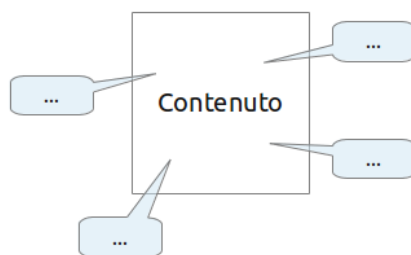


Figura 4: Valore informativo di un contenuto

### 2.3.1 *Enciclopedia del sapere*

Attualmente, il limite della piattaforma consiste nell'essere un serbatoio di CONTENUTI INFORMATIVI disaggregati e priva degli strumenti per classificare e catalogare il sapere custodito conferendovi una struttura ordinata, una sorta di indice in grado di facilitarne la ricerca, il reperimento e la consultazione.

L'obiettivo del criterio di classificazione consiste essenzialmente nel costruire un'enciclopedia del sapere a partire dalle informazioni presenti nei contenuti informativi.

Un'enciclopedia è un'opera letteraria che raccoglie e ordina la sintesi della conoscenza umana in tutti i campi o in un determinato settore. Le enciclopedie sono divise in voci, o lemmi, cui si accede di solito in ordine alfabetico.  
[1]

Il modello concettuale dell'enciclopedia diventa un riferimento utile ad identificare:

1. i casi d'uso essenziali concernenti la consultazione del sapere custodito nella piattaforma (ricerca per lemma, argomento o affinità);
2. gli elementi strutturali del sistema di classificazione (lemmi, accezioni, ...);
3. i requisiti e le specifiche più rilevanti per il criterio di classificazione;
4. le principali criticità rispetto alla coerenza e consistenza dell'enciclopedia.

**CONTENUTO INFORMATIVO** Ciascun contenuto può essere considerato una collezione di frammenti di informazioni (v. figura 4), che contribuiscono ad arricchire la conoscenza relativa ad uno o più lemmi dell'enciclopedia. Il criterio di classificazione deve quindi tenere traccia della relazione tra contenuti informativi e lemmi per essere in grado di catalogare e ricostruire l'intero sapere disponibile riguardo un certo tema.

**ENTITÀ** Nell'ambito della piattaforma, i lemmi vengono definiti entità ( $d_i \in D$ ) e rappresentano elementi concreti (luoghi, persone, eventi, ...) o astratti (concetti, temi, ...) a cui afferiscono i contenuti.

L'insieme delle entità definite - in un certo istante - all'interno della piattaforma rappresenta il **DOMINIO DELLA CONOSCENZA** (di seguito per brevità **DOMINIO**).

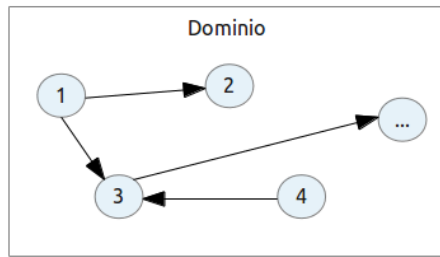


Figura 5: Relazioni tra le entità del dominio

Così come ciascun lemma di un'enciclopedia contiene spesso riferimenti ad altre voci, che trattano temi specifici o attinenti, nella piattaforma ciascuna entità può riferire o essere riferita da un numero arbitrario di entità distinte e possono esistere riferimenti incrociati, ossia coppie di entità che si citano a vicenda.

Ciò si traduce nel modello relazionale con un vincolo referenziale di tipo molti-a-molti tra le entità, che distingue referenti e riferite e permette di interpretare il dominio  $D$  mediante una struttura a grafo orientato ove:

- ciascun nodo rappresenta un'entità;
- ciascun arco uscente identifica un'ENTITÀ RIFERITA;
- ciascun arco uscente identifica un'ENTITÀ REFERENTE.

**ETICHETTE** Ciascuna entità ha un valore semantico preciso ed univoco, ma dev'essere identificata anche sul piano sintattico mediante un'etichetta ( $e_j \in E$ ), ossia una stringa di lunghezza variabile che consenta agli utenti di riferirla all'interno di ciascun contenuto.

L'insieme di etichette definite in un certo istante costituisce il **DIZIONARIO** della piattaforma.

Il modello illustrato presenta tuttavia due notevoli inconvenienti, che devono essere risolti per soddisfare i requisiti e gli obiettivi fissati: l'ambiguità sintattica e quella semantica.

### 2.3.2 Ambiguità sintattica

Nel linguaggio comune, è possibile riferirsi ad una certa entità ( $d_i$ ) con termini o espressioni differenti ( $e_{i,j} \in E_i$ ): la presenza di sinonimi, aventi il medesimo valore semantico ma diversa sintassi, rappresenta un fattore di ambiguità intrinseco e non trascurabile, che trasforma la relazione uno-a-uno tra entità ed etichette in una di tipo uno-a-molti.

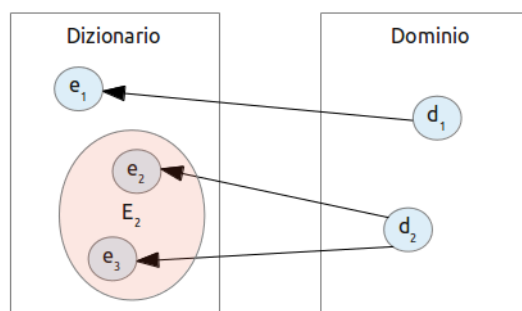


Figura 6: Ambiguità sintattica di un'entità

Il criterio di classificazione è dunque chiamato a conciliare la possibilità per l'utente di riferirsi ad una certa entità con etichette differenti e l'esigenza di identificarla univocamente nei contenuti informativi.

**SINONIMI** La gestione dei sinonimi di un'entità rappresenta un aspetto cruciale, poiché scegliere tra i possibili sinonimi un'etichetta arbitraria con cui identificare un'entità ed ignorare i rimanenti, pur semplificando il modello, costringerebbe gli utenti a conoscere e ad utilizzare solo quell'etichetta, imponendo una scelta del tutto arbitraria e - in quanto tale - fortemente soggettiva.

Per facilitare la ricerca di un'entità è opportuno includere e conservare nel dizionario tutte le etichette note o utilizzate dagli utenti, così da rendere il sistema sempre più accurato nell'individuare e riconoscere l'entità cui un utente fa riferimento quando utilizza una certa etichetta. Con il passare del tempo, il numero dei sinonimi di ciascuna entità è destinato a crescere grazie al contributo degli utenti, garantendo così una migliore copertura sintattica.

La proliferazione di sinonimi, ossia etichette equivalenti sul piano semantico ma sintatticamente differenti, rischia tuttavia di avere pesanti ripercussioni sull'efficacia del criterio di classificazione e particolarmente sull'efficienza della ricerca: se ai contenuti vengono assegnate delle etichette, reperire tutte e sole le informazioni inerenti una certa entità richiederebbe infatti di cercare riscontri nei contenuti pubblicati per ogni etichetta con cui possa venir riferita.

**CONTENUTI** A tale inconveniente si pone rimedio facilmente definendo una relazione molti-a-molti tra le entità e i contenuti, che ignori le etichette utilizzate per indicarle.

Così facendo si preserva l'identificazione univoca di ciascuna entità nei contenuti informativi, almeno sul piano semantico, e si guadagna in termini di efficienza.



Si considerino il dizionario  $E$  ed il dominio  $D$ . Il numero medio  $\alpha$  di etichette associate a ciascuna entità risulta pari a:<sup>1</sup>

$$\alpha = \frac{|E_0| + |E_1| + \dots}{|D|} = \frac{\sum |E_i|}{|D|} \quad (1)$$

ove

$$\sum |E_i| = |E| \quad (2)$$

Ciò significa intuitivamente che assegnare ai contenuti un'etichetta arbitraria aumenterebbe di una costante moltiplicativa  $\alpha$  la complessità dell'operazione di ricerca, che dovrebbe essere ripetuta per ciascuna delle  $\alpha$  etichette anziché per la sola entità corrispondente.<sup>2</sup>

**ETICHETTE** Tuttavia, per riferire un'entità nella piattaforma web occorre caratterizzarla anche sul piano sintattico. Sebbene non vi siano particolari controindicazioni nel permettere all'utente di scegliere l'etichetta da utilizzare nei contenuti pubblicati, si rischia di:

- generare ambiguità e confusione tra gli utenti stessi, rendendo poco intuitivo stabilire se due entità citate in due contenuti diversi e con etichette differenti rappresentino effettivamente la stessa;
- dover aggiungere alla relazione tra entità e contenuti un'informazione aggiuntiva, ossia l'etichetta scelta per indicarla.

Per tale ragione, si individua arbitrariamente - per ciascuna entità  $d_i$  - un'ETICHETTA PRIMARIA ( $e_{i,0} \in E_i$ ), che la identifica univocamente nell'ambito della piattaforma, mentre le restanti (ETICHETTE SECONDARIE) ne vengono considerate sinonimi.

Così facendo si introducono tuttavia alcuni vincoli aggiuntivi sulla relazione, che risulta scissa in due distinte:

1. una relazione di tipo uno-a-uno tra l'entità e la corrispondente etichetta primaria;
2. una relazione uno-a-molti tra l'entità e le etichette secondarie.

In conclusione, il nodo cruciale della soluzione individuata consiste nel mantenere separato l'aspetto semantico (le entità) da quello sintattico (le etichette): ogni qualvolta l'utente ricerca o assegna un'etichetta ad un contenuto, il sistema traduce il suo ingresso sintattico (l'etichetta) in un'uscita semantica (l'entità corrispondente).

<sup>1</sup> Poiché la relazione tra etichette ed entità è di tipo uno-a-molti, ciascuna etichetta è associata ad una ed una sola entità. Ne consegue che  $P = \{E_0, E_1, \dots\}$  è una partizione di  $E$ , ossia vale  $\bigcup E_i = E$  e  $\forall A \in P, B \in P : A \neq B \implies A \cap B = \emptyset$ .

<sup>2</sup> Si assuma per semplicità che la ricerca verifichi - per ciascun contenuto - quali entità o etichette cercate siano presenti, una per volta, e che la complessità computazionale sia equivalente in entrambi i casi.

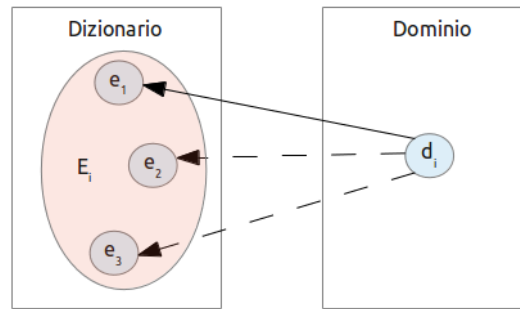


Figura 7: Etichetta primaria e secondarie

Tale meccanismo è tuttavia soggetto ad alcune complicazioni, dovute all'ambiguità semantica di ciascuna etichetta.

### 2.3.3 Ambiguità semantica

Riprendendo il modello dell'enciclopedia, si osserva come ciascuna etichetta  $e_j$  possa avere svariate ACCEZIONI ( $a_{j,k} \in A_j$ ), ciascuna delle quali ne individua un significato differente e si riferisce ad un'entità distinta.




Figura 8: Accezioni di un'etichetta (+ modello relazionale)

Sebbene la natura tematica della piattaforma tenda a mitigare la presenza di etichette aventi accezioni multiple, tale eventualità non può essere esclusa e dev'essere dunque opportunamente gestita.

## The Newsroom

Da Wikipedia, l'enciclopedia libera.

**A questo titolo corrispondono più voci, di seguito elencate.**

 Questa è una [pagina di disambiguazione](#); se sei giunto qui cliccando un [collegamento](#), puoi tornare indietro e correggerlo, indirizzandolo direttamente alla voce giusta.

- **The Newsroom** - serie televisiva canadese, andata in onda dal 1996 al 2005
- **The Newsroom** - serie televisiva statunitense del 2012

Figura 9: Pagina di disambiguazione [2]

Occorre innanzi tutto rivedere la relazione uno-a-molti tra entità ed etichette, che diventa di tipo molti-a-molti e per cui vale la proprietà illustrata di seguito.

Siano  $\alpha_{j,k}$  e  $\alpha_{j,k'}$  accezioni distinte di un'etichetta  $e_j$  e sia definita la funzione

$$f(\alpha_{j,k}) : A_j \rightarrow D \quad (3)$$

che restituisce l'entità  $d_i$  riferita dall'accezione  $\alpha_{j,k}$ . Allora

$$\alpha_{j,k} \neq \alpha_{j,k'} \implies f(\alpha_{j,k}) \neq f(\alpha_{j,k'}) \quad (4)$$

L'equazione 1 conserva la propria validità, mentre la 2 dev'essere modificata per riflettere la differente molteplicità della suddetta relazione, ossia il fatto che ciascuna etichetta possa riferire una o più entità:

$$\sum |E_i| = \sum |B_i| = |A| = \sum |A_j| \geq |E| \quad (5)$$

Si registrano inoltre rilevanti ripercussioni in merito alla distinzione tra etichette primarie e secondarie. A ciascuna entità continuano ad essere associate una e una sola etichetta primaria, oltre ad un numero arbitrario (anche nullo) di etichette secondarie: ove ciascuna etichetta può riferirsi a diverse entità, tale distinzione varia a seconda dell'accezione considerata, poiché la medesima etichetta può essere primaria rispetto ad un'entità e secondaria rispetto ad un'altra.

Per preservare i suddetti vincoli sulle relazioni tra etichette ed entità, si classificano le accezioni in CHIAVE ( $\alpha_{j,0}$ ) o SINONIMICHE ( $\alpha_{j,1}, \dots, \alpha_{j,|A_j|}$ ), a seconda che l'etichetta risulti rispettivamente primaria o secondaria per l'entità corrispondente. In questo modo si riesce a trasferire tale distinzione dalle etichette alle accezioni, preservando la percezione della relazione dal punto di vista dell'entità.



Figura 10: Accezioni chiave e sinonimiche

A risentire maggiormente dell'esistenza delle accezioni è il processo di ricerca di un'entità  $d_i$  a partire da un'etichetta  $e_j$ : se  $|A_j| \geq 2$  l'identificazione dell'entità richiede - da parte dell'utente - la selezione di un'accezione  $\alpha_{j,k} \in A_j$  tra le  $|A_j|$  disponibili per indicare esplicitamente l'entità cui fa riferimento.

## 2.3.4 Modello relazionale

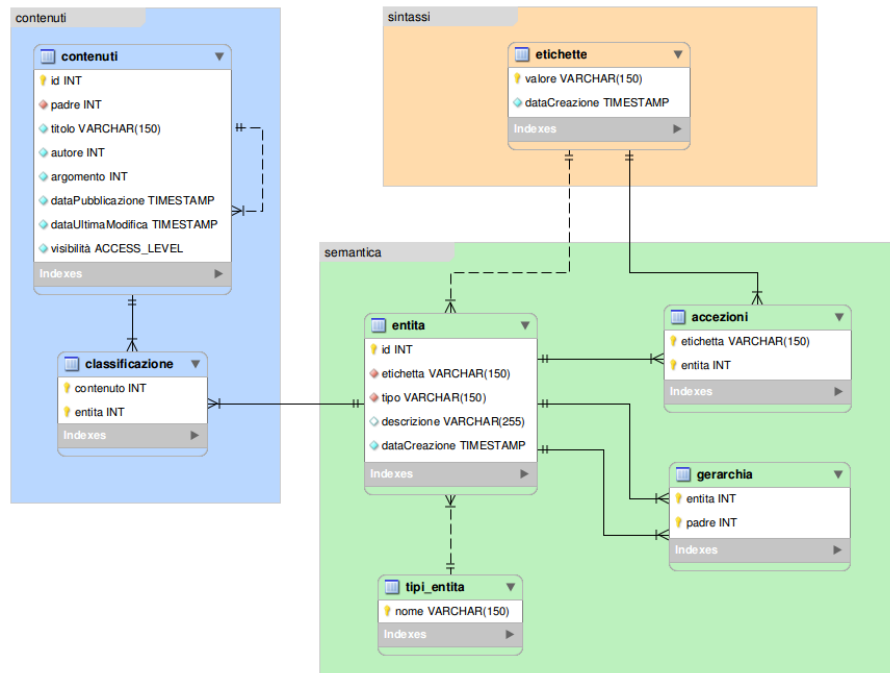


Figura 11: Modello relazione del criterio di classificazione

Al termine della fase progettuale, si rende necessario modificare il modello relazionale della piattaforma per integrare le informazioni aggiuntive legate al nuovo criterio di classificazione.

Rispetto all'immagine 11, la sola tabella contenuti risulta importata dal modello relazionale della piattaforma per evidenziare alcune relazioni fondamentali. Le rimanenti sono organizzate in tre *layer* (contenuti, sintassi e semantica) per chiarirne il ruolo all'interno del sistema di classificazione.

**ETICHETTE** Le etichette sono rappresentate mediante la tabella etichette e sono identificate univocamente dalla stringa associata, che rappresenta l'attributo etichette.valore. Tale scelta risponde alla naturale identificazione dell'etichetta nella sequenza di caratteri corrispondente, costituisce una garanzia contro la presenza di duplicati e consente di recuperare il valore dell'etichetta primaria di un'entità senza dover effettuare un'operazione di *join* tra le tabelle entita ed etichette.

**ACCEZIONI** Le accezioni rappresentano un legame univoco tra le etichette e le entità e possono essere di tipo chiave o sinonimico. Esse non presentano attributi propri significativi, ma prevedono tre vincoli referenziali:

1. a ciascuna entità è associata una e una sola ACCEZIONE CHIAVE (relazione uno-a-uno), che identifica l'etichetta primaria;
2. a ciascuna entità sono associate  $0 \dots n$  ACCEZIONI SINONIMICHE (relazione uno-a-molti), che rappresentano i sinonimi dell'etichetta primaria;
3. ciascuna etichetta possiede  $0 \dots n$  ACCEZIONI (relazione uno-a-molti).



Figura 12: Modello ad oggetti delle accezioni

Ne consegue che sia la superclasse (accezioni) sia le sottoclassi (accezioni\_chiave e accezioni\_sinonimiche) presentano dei vincoli referenziali; in particolare, la presenza delle prime due relazioni costringe a distinguere - dal punto di vista dell'entità - tra accezioni chiave e sinonimiche.

Per modellare tale scenario vengono presi in considerazione tre possibili approcci:

#### TABELLA UNICA

La tabella unica ben si adatta a gestire l'assenza di attributi propri per le sottoclassi e ad esprimere i vincoli referenziali che coinvolgono la superclasse, ma non è in grado di esprimere e adeguatamente rappresentare quelli coinvolgenti le sottoclassi.

#### PARTIZIONAMENTO ORIZZONTALE

Il partizionamento orizzontale riesce a modellare i vincoli referenziali delle sottoclassi, ma non quello della superclasse, e genera due classi aventi i medesimi attributi.

#### PARTIZIONAMENTO VERTICALE

Il partizionamento verticale consente di modellare correttamente tutti e tre i vincoli referenziali, relativi sia alla superclasse sia alle sottoclassi. Tuttavia si rende più complesso modificare il tipo di un'accezione e si introduce l'esigenza di un'operazione *join* per recuperare la lista completa delle accezioni, pur non possedendo le sottoclassi attributi propri.

A seguito di alcune osservazioni si decide di adottare la soluzione della tabella unica:

- la distinzione tra etichette chiave e sinonimiche ha rilevanza essenzialmente dal punto di vista della classe entità;

- il vincolo referenziale tra etichette ed accezioni suggerisce che la distinzione di cui al punto precedente sia irrilevante dal punto di vista delle etichette, ragion per cui risulta utile mantenere tutte le accezioni nella medesima tabella.<sup>4</sup>
- le sottoclassi non hanno attributi propri, per cui il partizionamento verticale e orizzontale sono ritenute soluzioni inadeguate.

Il soddisfacimento delle condizioni richieste viene raggiunto eliminando qualsiasi riferimento al tipo dell'accezione nella classe accezioni e modellando direttamente la relazione uno-a-uno tra le entità e le relative etichette primarie mediante un vincolo referenziale di chiave esterna nella classe entita, ossia entita.etichetta, che identifica la corrispondente etichetta primaria nella tabella etichette. Così facendo si riescono ad esprimere tutti i vincoli referenziali senza dover definire le sottoclassi.

---

<sup>4</sup> Si consideri ad esempio il caso d'uso della ricerca di un'entità a partire da un'etichetta, ove occorre recuperare la lista completa delle relative accezioni.

## VALUTAZIONI FINALI

---

GLOSSARIO

---

## B

BITBUCKET - <https://bitbucket.org/>

Piattaforma web per la gestione delle attività di progetto con supporto a strumenti di controllo di versione distribuito.

## C

CAMELCASE

Convenzione per la scrittura di espressioni composte unendo le parole tra loro e mantenendo ciascuna iniziale in maiuscolo.

## G

GEDIT - <http://projects.gnome.org/gedit/>

Editor di testo ufficiale dell'ambiente desktop GNOME.

## L

LATEXILA - <http://projects.gnome.org/latexila/>

Editor LaTeX integrato per l'ambiente desktop GNOME.

LIBREOFFICE CALC - <http://www.libreoffice.org/>

Applicazione per fogli di calcolo della suite di produttività *LibreOffice*.

## M

MERCURIAL - <http://mercurial.selenic.com/>

Strumento multi piattaforma, gratuito ed open source per il controllo di versione distribuito.



## MVC

Design pattern architetturale basato sulla scomposizione del sistema in tre componenti essenziali: *model* (logica di business), *view* (interfaccia grafica) e *controller* (logica di controllo).

MYSQL WORKBENCH - <http://www.mysql.it/products/workbench/>

Applicazione multi piattaforma, gratuita ed open source per la progettazione, lo sviluppo e l'amministrazione di database MySQL.

## P

## PDF (PORTABLE DOCUMENT FORMAT)

Formato di file per la rappresentazione di documenti in maniera indipendente dalla piattaforma hardware e software.

PENCIL - <http://pencil.evolus.vn/>

Applicazione multi piattaforma, gratuita ed open source per la realizzazione di prototipi di interfacce grafiche.

PROJECTLIBRE - <http://www.projectlibre.org/>

Applicazione multi piattaforma, gratuita ed open source per il *project management*, che consente di realizzare diagrammi di Gantt e di PERT, gestire le risorse allocate e le attività pianificate, ... .

## R

## RETE SOCIALE

Insieme di persone, aventi interessi in comune e inclini a collaborare e condividere idee o informazioni, e di relazioni di tipo esperienziale definite tra tali soggetti.

## U

UBUNTU - <http://www.ubuntu.com/>

Distribuzione Linux gratuita derivata da Debian.

UML - <http://www.uml.org/>

Standard internazionale per un linguaggio di modellazione, che definisce un insieme di notazioni grafiche per la rappresentazione visiva di sistemi.

UMLET - <http://www.umlet.com/>

Applicazione multi piattaforma, gratuita ed open source per la realizzazione di diagrammi UML.

## BIBLIOGRAFIA

---

- [1] Wikipedia. Enciclopedia - wikipedia, l'enciclopedia libera, 2012.  
[Online; in data 23-novembre-2012].
- [2] Wikipedia. The newsroom - wikipedia, l'enciclopedia libera, 2012.  
[Online; in data 23-novembre-2012].