



POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Un Q&A Systems per ATS

TESI DI LAUREA MAGISTRALE IN
INGEGNERIA INFORMATICA

Autore: Marco Livrieri

ID studente: 639461

Tutor: Alessandro Campi

Anno accademico: 2022-23

Riassunto

L'ATS di Monza ha commissionato al Politecnico di Milano la realizzazione di un KMS (Knowledge Management System) che permetta alla propria utenza di fruire di un servizio di Q&A (Question & Answer) in modo simile al sito <https://stackoverflow.com/>

L'attività, oggetto della tesi, è stata quella di confrontare alcune soluzioni open source di KMS presenti sul web per determinare la più completa dal punto di vista delle funzionalità messe a disposizione e contemporaneamente la più intuitiva e facile da utilizzare.

Il prodotto è stato poi calato nel contesto dell'ATS di Monza personalizzando il layout secondo lo stile e i colori dell'azienda.

Parole chiave: KMS, Q&A, Scoold, stackoverflow

Contenuti

Riassunto	3
Contenuti.....	4
1 Knowledge Management System	5
1.1. Tipi di conoscenza.....	5
1.2. Knowledge Management	6
1.3. Strumenti	6
1.4. Benefici.....	7
2 Question&Answer System	8
3 Soluzioni open source	9
4 Scoold	10
4.1. Pagina principale.....	13
4.2. Ricerca.....	15
4.3. Nuova domanda.....	16
5 Bibliografia.....	19
6 Elenco delle figure	20

1 Knowledge Management System

Un Knowledge Management System (KMS) è un sistema che sfrutta la conoscenza collettiva dell'organizzazione, portando a migliori efficienze operative. Questi sistemi sono supportati dall'uso di una Knowledge Base (KB). Di solito sono fondamentali per una gestione della conoscenza di successo, fornendo un luogo centralizzato per archiviare le informazioni e accedervi facilmente.

Il KMS è un sistema che permette di attuare il Knowledge Management (KM), cioè il processo di identificazione, organizzazione, archiviazione e diffusione delle informazioni all'interno di un'organizzazione.

1.1. Tipi di conoscenza

Quali sono i tipi di conoscenza che un'organizzazione può gestire? Si possono identificare tre tipi di conoscenza:

- **Tacita:** questo tipo di conoscenza è tipicamente acquisita attraverso l'esperienza ed è intuitivamente compresa. Di conseguenza, è difficile da articolare e da codificare, rendendo difficile il trasferimento di queste informazioni ad altri individui. Esempi di conoscenza tacita possono includere il linguaggio, il riconoscimento facciale o le capacità di leadership.
- **Implicita:** a differenza della tacita, questo tipo di conoscenza può essere documentata formando così una Knowledge Base.
- **Esplicita:** viene acquisita all'interno di vari tipi di documenti come manuali, report e guide, consentendo alle organizzazioni di condividere facilmente le conoscenze tra i vari gruppi. Questo tipo di conoscenza è forse il più noto e gli esempi includono risorse di conoscenza come database, white paper e casi di studio. Questa forma di conoscenza è importante per mantenere il capitale intellettuale all'interno di un'organizzazione e facilitare il trasferimento di conoscenze di successo ai nuovi dipendenti.

1.2. Knowledge Management

Il processo di KM può essere sintetizzato in tre passaggi principali:

- **Creazione:** durante questa fase, le organizzazioni identificano e documentano qualsiasi conoscenza esistente, o nuova, che desiderano far circolare in tutta l'azienda.
- **Archiviazione:** durante questa fase, un sistema informatico viene in genere utilizzato per ospitare le conoscenze organizzative per la distribuzione. Potrebbe essere necessario formattare le informazioni in un modo particolare per soddisfare i requisiti di tale repository.
- **Condivisione:** in questa fase finale, i processi per condividere le conoscenze vengono comunicati ampiamente in tutta l'organizzazione. La velocità con cui le informazioni si diffondono varia a seconda della cultura organizzativa. Le aziende che incoraggiano e premiano questo comportamento avranno sicuramente un vantaggio competitivo rispetto ad altre nel loro settore.

1.3. Strumenti

Quali strumenti mette a disposizione un KMS? Di seguito si propone un elenco di strumenti normalmente utilizzati dalle organizzazioni:

- **Gestione dei documenti:** fungono da sistema di archiviazione centralizzato per documenti digitali, come PDF, immagini e file di elaborazione testi. Questi sistemi migliorano i flussi di lavoro dei dipendenti consentendo un facile recupero dei documenti.
- **Gestione dei contenuti (CMS):** sono applicazioni che gestiscono contenuti Web in cui gli utenti finali possono modificare e pubblicare contenuti. Questi sono comunemente confusi con i sistemi di gestione dei documenti, ma i CMS possono supportare altri tipi di media, come audio e video.
- **Intranet:** sono reti private che esistono esclusivamente all'interno di un'organizzazione, che consentono la condivisione di abilitazioni, strumenti e processi all'interno degli stakeholder interni. Sebbene possano richiedere molto tempo ed essere costosi da mantenere, forniscono una serie di servizi groupware, come directory interne e ricerca, che facilitano la collaborazione.
- **Wiki:** possono essere uno strumento di gestione della conoscenza popolare data la loro facilità d'uso. Semplificano il caricamento e la modifica delle informazioni, ma questa facilità può portare a preoccupazioni sulla disinformazione in quanto i lavoratori possono aggiornarle con informazioni errate o obsolete.

- **Data Warehouse:** aggregano i dati provenienti da fonti diverse in un unico archivio dati centrale e coerente per supportare l'analisi dei dati, il data mining, l'intelligenza artificiale (AI) e l'apprendimento automatico. I dati vengono estratti da questi repository in modo che le aziende possano ricavare informazioni, consentendo ai dipendenti di prendere decisioni basate sui dati.
- **Question&Answer:** sistema di recupero automatico delle informazioni, destinato a rispondere alle domande che vengono poste in un linguaggio naturale.

1.4. Benefici

Sono diverse le direttrici attraverso le quali una piattaforma di Knowledge Management System consente di efficientare fin dalla sua adozione un'azienda, generando essenzialmente **cinque benefici**:

- **Facilità d'uso:** ricerche in linguaggio naturale, multilingua e cross-lingua, con la navigazione di concetti, argomenti ed entità
- **Maggiore precisione:** aumenta la qualità della risposta a fronte di una ricerca
- **Facilità di navigazione** su dati strutturati e non strutturati: si riducono i tempi di ricerca delle informazioni e si valorizza il patrimonio di competenze ed esperienze aziendali
- **Potenziamento del Document management:** attraverso specifiche funzionalità end-to-end diventa possibile gestire, consultare e aggiornare i contenuti in modalità omni-channel.
- **Miglioramento continuo:** la reportistica realizzata contestualmente all'uso del Knowledge Management System permette infine di monitorare tutte le fasi del processo editoriale e di accesso alla base di conoscenza, con la possibilità di farlo progredire costantemente.

2 Question&Answer System

“Question answering” è una disciplina informatica nei campi del recupero delle informazioni e dell'elaborazione del linguaggio naturale (NLP), che si occupa della costruzione di sistemi che rispondono automaticamente alle domande poste dagli esseri umani in un linguaggio naturale. Esistono 2 tipi di sistemi:

- **QA a dominio chiuso:** riguarda la creazione di sistemi che rispondono alle domande di un dominio specifico e le domande sono solitamente limitate per essere di natura "descrittiva". Quindi, ad esempio, supponiamo di aver usato questo documento come input per il nostro sistema di controllo qualità a dominio chiuso, potremmo chiedergli "Quali sono i benefici di un MKS?", ma non potremmo chiedergli qualcosa come "Perché si piange quando si tagliano le cipolle?". L'idea è che i sistemi di QA a dominio chiuso sono esposti a set di dati molto più piccoli da cui estrarre la risposta, il che naturalmente limita la gamma di domande che possono essere poste (dato che le risposte devono "provenire da qualche parte").
- **QA a dominio aperto:** si occupa di domande su quasi tutto e può fare affidamento su ontologie generali e conoscenza del mondo. Tuttavia, questi sistemi di solito hanno molti più dati disponibili da cui estrarre la risposta. Possono rispondere a domande come "Qual è il significato della vita?", "Mi sostituirai nel mio lavoro" (e altri tipi esistenziali) e sono i tipi di sistemi utilizzati dagli assistenti virtuali.

3 Soluzioni open source

Navigando sul web è possibile trovare svariate proposte opensource alternative a stackoverflow che presentano caratteristiche simili e scritte in vari linguaggi di programmazione.

Tra le soluzioni che vengono citate più comunemente troviamo:

	Sito web	Download	Linguaggio	Database	Licenza
Question2Answer	Question2Answer	GitHub - Question2Answer	PHP 5.4 o superiore	MySQL 5.0 o superiore	GPL v2+
Codidact	Codidact	GitHub - codidact/qpixel	Ruby on Rails ver 3	MySQL	AGPL
Scoold	Scoold	GitHub - Erudika/scoold	Java	H2	Apache 2.0
AskBot	AskBot	GitHub - ASKBOT	Python 3.7	MySQL / PostgreSQL	GPL 3

Tutte le soluzioni, essendo alternative dello stesso sito, offrono caratteristiche molto simili, come ad esempio la possibilità di votare le domande/risposte, l'utilizzo dei tag, il punteggio di reputazione, ecc...

Tra tutte le alternative, ho scelto Scoold in quanto basato sul linguaggio java. Java è largamente diffuso nel mondo per la sua portabilità, la facilità di reperimento di documentazione e per la presenza di innumerevoli community che permettono di risolvere facilmente qualsiasi problema che si possa riscontrare durante gli sviluppi.

Oltre al linguaggio, Scoold presenta un'interfaccia intuitiva, semplice e facile da utilizzare. Questo aspetto è fondamentale per essere fruibile anche da utenti un po' meno esperti nell'utilizzo degli strumenti informatici.

4 Scoold

Scoold è una piattaforma di domande e risposte e condivisione delle conoscenze per i gruppi di lavoro. Il progetto è stato creato nel 2008, rilasciato nel 2012 come social network per le scuole ispirato a Stack Overflow. Nel 2017 è stato rivisto, riconfezionato e reso open source.

Scoold può essere eseguito ovunque: Heroku, DigitalOcean, AWS, Azure o qualsiasi provider di hosting VPS. È leggero (~ 7000 LOC), il backend è gestito da un servizio separato chiamato Para (<https://paraio.com/>). Scoold non richiede un database e la logica del controller è davvero semplice perché tutto il lavoro è delegato a Para. Questo rende il codice facile da leggere e può essere appreso rapidamente dagli sviluppatori junior.

Tra le caratteristiche di Scoold troviamo:

- **Piattaforma completa di domande e risposte:** Scoold si ispira a Stack Overflow e implementa la maggior parte delle sue funzionalità. Puoi porre e rispondere a domande, votare positivamente o negativamente un messaggio, guadagnare distintivi e reputazione. Ogni modifica crea una nuova revisione, semplificando il monitoraggio della cronologia delle modifiche.
- **Ricerca full-text:** ogni messaggio viene indicizzato e analizzato da Para attraverso Elasticsearch, un potente motore di ricerca. Puoi cercare utenti, domande e risposte dalla barra di ricerca in alto.
- **Ricerca basata sulla posizione e filtraggio dei post "vicino a me":** le domande possono essere associate a una posizione specifica e quindi filtrate in base a tale posizione. Questa funzione consente le query di geolocalizzazione e dà la possibilità di cercare domande intorno a te.
- **I18n con supporto linguistico RTL:** l'intero sito web è facilmente traducibile in altre lingue. Sono supportate anche le lingue da destra a sinistra e il layout viene rispecchiato se viene rilevata una lingua RTL.
- **Reputazione e sistema di voto con distintivi:** Scoold ha un sistema di reputazione semplice ma efficace con distintivi. Alle persone vengono assegnati punti reputazione per buone domande e risposte, oltre a vari distintivi per i loro risultati sul sito. E' possibile anche creare distintivi personalizzati.

- **Spazi (gruppi) - gruppi di domande e utenti isolati:** gli spazi sono un modo per organizzare utenti e domande in gruppi isolati. Si può modificare gli spazi degli utenti in blocco e anche spostare una domanda in uno spazio diverso. L'autorizzazione per accedere a uno spazio viene concessa da un amministratore.
- **Webhook con firma:** gli amministratori possono creare webhook a servizi esterni per tutti gli eventi in Para (creazione, aggiornamento, eliminazione). I payload sono firmati con HMAC-SHA256.
- **Integrazione con Zapier:** Zapier semplifica l'integrazione di Scoold con migliaia di altre applicazioni. Grazie alla potenza degli hook REST, è possibile creare automaticamente domande in base ai dati di un'app esterna o configurare Scoold per notificare ad altre app quando viene pubblicata una nuova domanda.
- **Codice JS frontend minimo basato su jQuery:** Scoold ha voluto mantenere le cose semplici sul frontend, ecco perché è stato scelto di utilizzare jQuery, la libreria JavaScript affidabile. Si occupa delle richieste AJAX e attiva gli elementi sullo schermo, niente di più.
- **Layout moderno e reattivo basato su Materialize CSS:** il codice frontend è alimentato da Materialize CSS, un framework basato su Material Design Lite di Google. È reattivo e soddisfacente. Le icone provengono da FontAwesome.
- **Suggerimenti per domande simili e suggerimenti per post duplicati:** Quando si pone una nuova domanda, Scoold fornirà suggerimenti se esiste già una domanda simile. Inoltre, in ogni pagina di domande c'è una barra laterale con suggerimenti per i primi dieci post simili.
- **Notifiche e-mail per le risposte ai post e i commenti:** è possibile iscriversi ad ogni thread di domande facendo click sulla casella di controllo in basso. Scoold invierà una e-mail ogni volta che viene pubblicata una nuova risposta.
- **Backup e ripristino:** gli amministratori possono facilmente esportare tutti i dati da Scoold come archivio zippato. I dati esportati possono essere successivamente importati in un'altra istanza di Scoold.
- **API RESTful definita con OpenAPI 3.0:** Scoold ha una vera e propria API JSON RESTful definita con OpenAPI 3.0. Tutte le operazioni CRUD possono essere eseguite tramite l'API da qualsiasi linguaggio di programmazione.
- **Supporto dell'autenticazione reciproca (mTLS):** E' possibile proteggere tutte le connessioni a Scoold con HTTPS utilizzando il tuo certificato SSL

senza la necessità di un proxy. La sicurezza è rafforzata grazie all'autenticazione reciproca tra proxy, Scoold e Para (mTLS).

- **Supporto per l'autenticazione LDAP:** Scoold è adatto alle aziende e supporta LDAP e Active Directory, quindi, è ottimo per le implementazioni intranet dietro firewall aziendali. Gli utenti Scoold possono autenticarsi senza problemi con il server LDAP o Active Directory della tua azienda per una piacevole esperienza di single sign-on.
- **Login tramite social con supporto Gravatar:** le persone possono accedere rapidamente senza moduli di registrazione o password. Facebook, Google, GitHub, LinkedIn, Microsoft e Twitter sono supportati come provider di identità. Gli accessi e-mail e password sono ancora supportati e le immagini del profilo vengono recuperate da Gravatar.
- **Evidenziazione della sintassi per il codice nei post, supporto del markdown GFM con tabelle, elenchi di attività e barrato:** tutti i frammenti di codice nei messaggi sono evidenziati in base alla sintassi. Sono supportate le tabelle, gli elenchi di attività e barrato. E' possibile utilizzare anche le emoji.
- **Importazione dei dati da Stack Overflow for Teams:** la migrazione da Stack Overflow for Teams a Scoold è semplice e senza problemi. Basta scaricare i dati da Stack Overflow e importarli in Scoold.
- **SEO friendly:** le pagine vengono renderizzate sul server, una tecnica consolidata, rendendole facilmente scansionabili dai motori di ricerca.
- **Temi e modalità scura:** è possibile cambiare rapidamente l'aspetto del tuo sito passando a uno dei temi in bundle oppure creandone uno personalizzato. E' presente anche un tema scuro per coloro che lavorano in una stanza scarsamente illuminata.

4.1. Pagina principale

La pagina principale presenta il contenuto della sezione “Domande” con l’elenco delle domande in ordine cronologico decrescente (pulsante “Più recenti” selezionato di default). E’ possibile filtrare i risultati utilizzando i pulsanti predefiniti oppure attraverso i filtri avanzati. L’elenco delle domande è visibile a tutti gli utenti che accedono al sito, anche senza autenticazione.

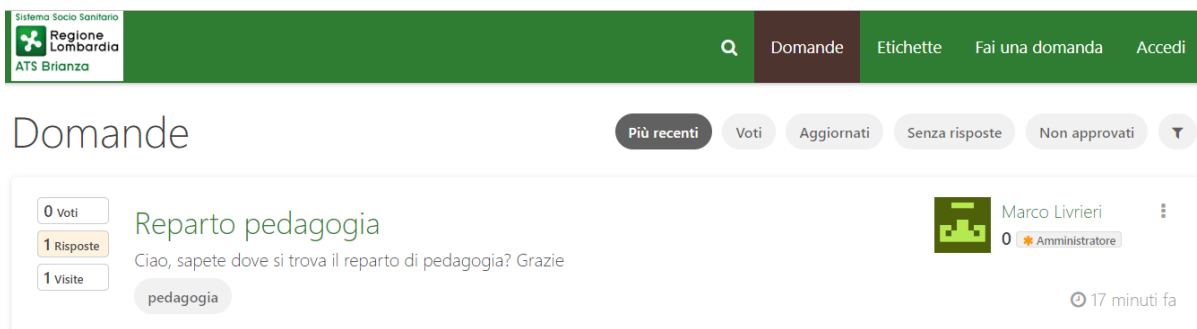


Figura 1: pagina principale

Ogni domanda è caratterizzata da un titolo, da un corpo e da una serie di etichette (parole chiave) che servono per essere visibili attraverso il filtro per etichette. Per ogni domanda vengono visualizzati l’autore, il tempo trascorso dall’ultima modifica, il numero di voti ricevuti, il numero delle risposte e il numero delle visite ricevute.

Cliccando sul titolo della domanda è possibile accedere alla storia delle modifiche / risposte / commenti ricevuti.

The screenshot displays a forum interface with two main sections: a question and an answer. The question, titled "Reparto pedagogia", was posted by Marco Livrieri, an administrator, 42 minutes ago. It has 1 visit, 1 response, and a "pedagogia" tag. The answer, posted by Mario Rossi 37 minutes ago, states "si trova nel padiglione C4." and has 0 votes. A comment by Mario Bianchi, posted 26 minutes ago, suggests indicating the floor. Both the question and answer have a "Segnalazioni" (Report) button. The interface includes navigation arrows, a "Voti" (Votes) button, and a "Più recenti" (More recent) filter.

Con risposta


Reparto pedagogia

Ciao,
sapete dove si trova il reparto di pedagogia?

Grazie

↑ ↓

Pubblicato 42 minuti fa

 Marco Livrieri
0 Amministratore

Segnalazioni

1 Visite

1 Risposta

🕒 42 minuti fa

📝 37 minuti fa

🏷 Etichette

pedagogia


Risposte

Voti Più recenti

Ciao,
si trova nel padiglione C4.

↑ ↓

Pubblicato 37 minuti fa

 Mario Rossi
0

↑ ↓ Avresti potuto indicare anche il piano — Mario Bianchi 26 minuti fa Segnalazioni

Segnalazioni

Figura 2: dettaglio della domanda

Le domande e le risposte possono essere votate positivamente o negativamente in base alla qualità del contenuto.

4.2. Ricerca

La funzione principale di un sistema di Q&A è quella di permettere di reperire l'informazione di cui si ha bisogno nel modo più pratico, facile e veloce possibile. Per fare questo, è possibile utilizzare la lente di ricerca che permette la ricerca di domane/risposte in un linguaggio naturale. Ad esempio, se sto cercando il reparto di pedagogia, potrei chiedere "dove trovo il reparto di pedagogia?" e il sistema mi propone le domande che potrebbero rispondere al mio quesito.

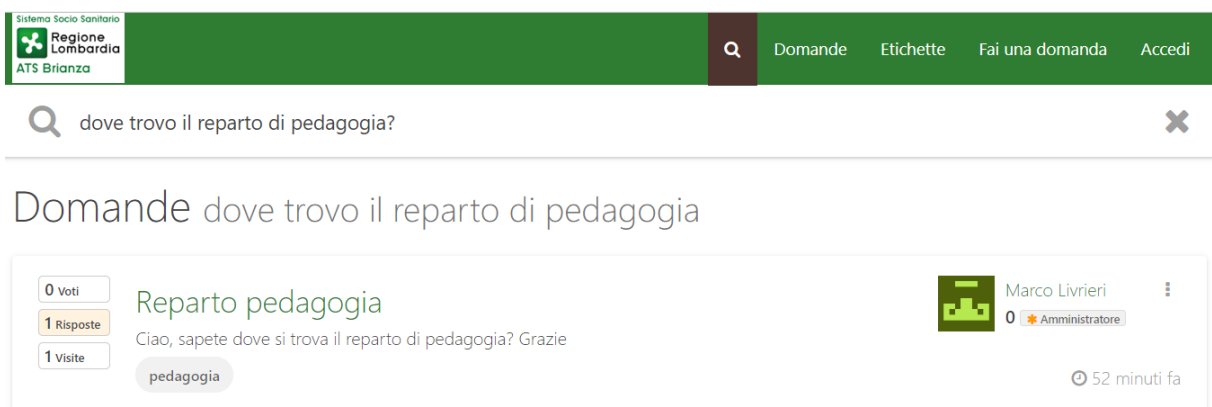
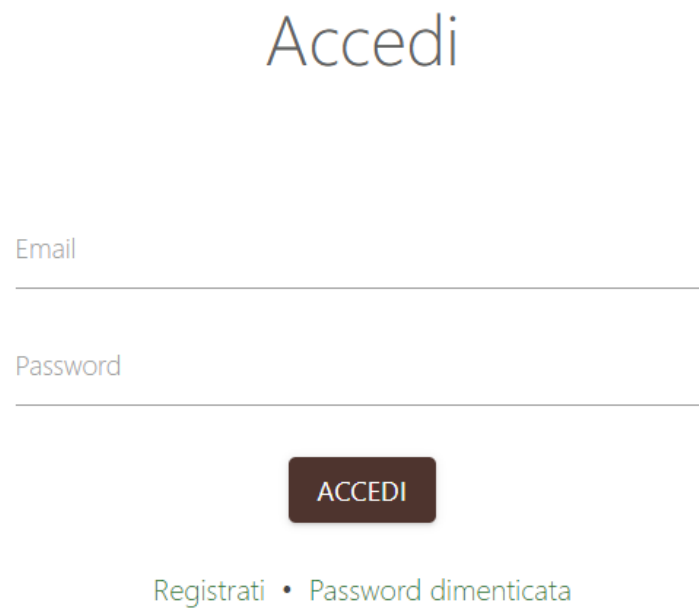


Figura 3: ricerca

4.3. Nuova domanda

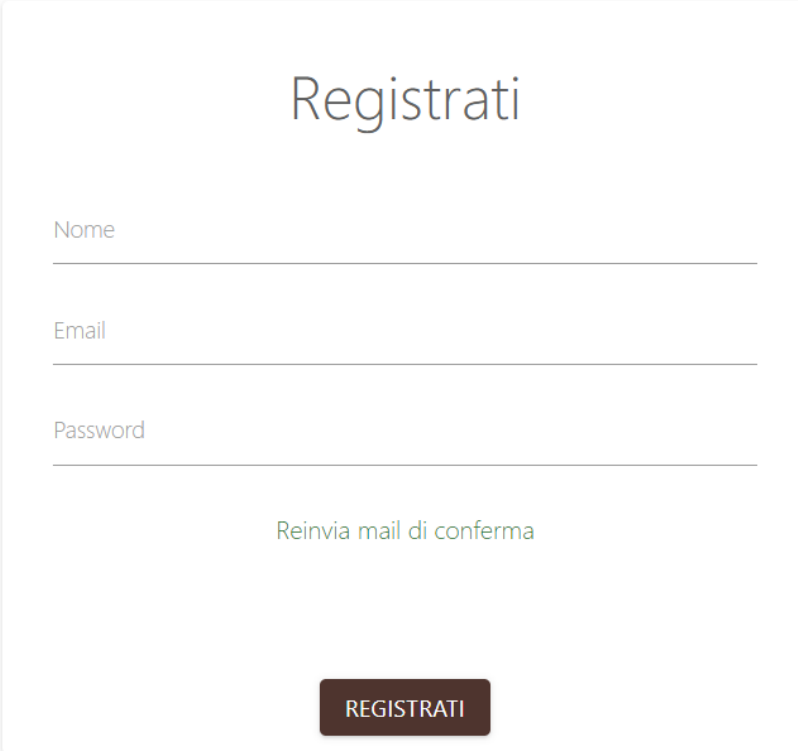
Per poter inserire una domanda è richiesta l'autenticazione. Per potersi autenticare è necessario passare dalla funzione "Accedi" ed inserire le proprie credenziali: indirizzo mail e password.



The image shows a login form with the title "Accedi" centered at the top. Below the title, there are two input fields: "Email" and "Password", each with a horizontal line underneath. Below the "Password" field is a dark blue button with the text "ACCEDI" in white. At the bottom of the form, there is a link "Registrati" followed by a bullet point and the text "Password dimenticata".

Figura 4: accesso

Nel caso in cui non si avessero ancora le credenziali per accedere al sito, si deve procedere alla registrazione attraverso il link “Registrati”



Registrati

Nome

Email

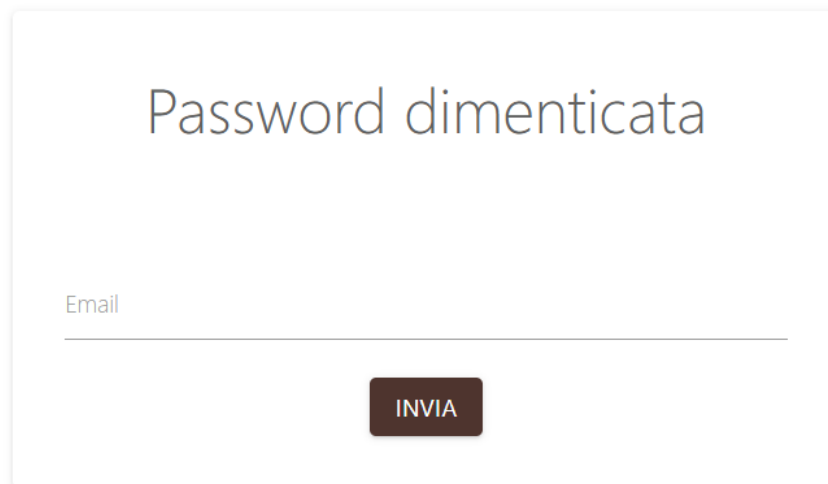
Password

[Reinvia mail di conferma](#)

REGISTRATI

Figura 5: registrazione

La registrazione invia una mail all'indirizzo indicato per accertarsi che sia una mail valida. Si dovrà poi cliccare sul link nella mail per confermare la registrazione ed attivare l'utente. Nel caso in cui si dimenticasse la password, nella sezione "Accedi" è possibile cliccare sul link "Password dimenticata" che attiva il processo per generare una nuova password.



The image shows a web form for password recovery. At the top, the title "Password dimenticata" is displayed in a large, grey font. Below the title, there is a text input field with the placeholder text "Email" in a small, grey font. A horizontal line is positioned below the input field. Centered below the line is a dark brown button with the word "INVIA" written in white, uppercase letters.

Figura 6: recupero password

5 Bibliografia

- [1] KMS ([What is Knowledge Management? | IBM](#))
- [2] SCOOLD ([Scoold | An Open Source Q&A Platform](#))
- [3] PARAIO ([Para - an open-source, general-purpose, backend framework for the cloud. \(paraio.org\)](#))
- [4] Alternative ([Open Source Stack Overflow Alternatives: Top 10 Q&A Services and similar apps | AlternativeTo](#))

6 Elenco delle figure

Figura 1: pagina principale	13
Figura 2: dettaglio della domanda	14
Figura 3: ricerca	15
Figura 4: accesso	16
Figura 5: registrazione	17
Figura 6: recupero password	18