



Heimdall pre-alpha
Visual document enhancement for eJustice

HEIMDALL

Un tool prototipale di *document enhancement* per il
magistrato
e l'Ufficio Per il Processo*

Sviluppato dall' Università degli Studi del Sannio
nell'ambito del progetto:

“Modelli organizzativi e innovazione digitale. Il nuovo Ufficio per il Processo
per l'efficienza del sistema-justizia”

*Alfonso Guarino, Ciro Fusco, Nicola Lettieri, Eugenio Zimeo

1. Introduzione	3
a. Scopo del documento	3
2. Descrizione Generale	3
a. Scopo del prodotto	3
b. Classi utente e caratteristiche	3
c. Ambiente operativo	4
d. Documentazione utente	4
3. Requisiti funzionali	4
3.1 Obiettivi e funzionalità	4
3.2 Estrazione dei riferimenti normativi dai documenti del fascicolo	5
3.3 Annotazione dei documenti per evidenziare informazioni importanti	5
3.4 Ricerca documenti simili	7
3.5 Ricerca contenuti nei documenti	7
3.6 Casi d'uso	7
4. Requisiti interfaccia esterna	15
a. Requisiti GUI	15
b. Altri tipi di interfaccia	16
5. Requisiti non funzionali	16
a. Requisiti di sicurezza	16
b. Requisiti prestazionali e scalabilità	16
6. Architettura del software	16
a. Paradigma di riferimento	16
b. Architettura del software	16
c. Assunzioni e Dipendenze	17
d. Ambiente operativo	17
7. Descrizione delle funzionalità erogate	17
8. Descrizione delle fasi di sviluppo ed integrazione nei sistemi IT della giustizia	21
Manuale d'uso del software	22
8. Installazione	22
a. Requisiti di sistema	22
b. Installazione del software	23
c. Configurazione del software	24
d. Aggiornamento del software	24
e. Disinstallazione del software	24
9. Guida all'uso	26
a. Accesso al software	26
b. Interfaccia utente	26
c. Operazioni di base	30
d. Funzionalità avanzate	31
e. Gestione dei dati	32

f. Personalizzazione delle impostazioni

33

1. Introduzione

a. Scopo del documento

L'obiettivo di questo documento è presentare l'applicativo prototipale Heimdall sviluppato per Magistrati e UPP nell'ambito del progetto "Modelli organizzativi e innovazione digitale. Il nuovo Ufficio per il Processo per l'efficienza del sistema-giustizia" dall'Università degli Studi del Sannio, Benevento.

Si offre una descrizione generale del prodotto realizzato soffermandosi su scopo, utenti, ambiente operativo (vedi Sezione 2), si spiegano in dettaglio i casi d'uso del sistema (Sezione 3), l'organizzazione dell'interfaccia (Sezione 4), l'architettura (Sezione 5) e tutte le sue funzionalità, sia quelle già sviluppate, sia quelle in fase di integrazione / completamento (Sezione 7).

Infine, il documento include il manuale utente (Sezioni 8-9) con modalità di installazione e disinstallazione passo-passo, versione illustrata delle modalità di interazione con il prodotto e le sue funzionalità, informazioni sui dati e la loro gestione.

2. Descrizione Generale

a. Scopo del prodotto

Il sistema Heimdall, sviluppato in forma prototipale, è pensato per supportare l'attività dei magistrati e dell'UPP abilitando una serie di funzionalità di *document enhancement* ed *information retrieval* utili per gestire la grande mole di documenti con cui quotidianamente hanno a che fare gli operatori della giustizia.

In particolare, lo strumento è pensato per: i) estrarre riferimenti normativi e giurisprudenziali contenuti nel fascicolo di causa per accesso rapido al testo della normativa citata; ii) annotare semanticamente i documenti per la concettualizzazione (al fine di individuare argomentazioni delle parti per un raffronto immediato) e per la massimazione delle sentenze (evidenziare le partizioni contenenti informazioni rilevanti ai fini della scrittura della massima); iii) ricercare documenti simili (in base ai riferimenti normativi e giurisprudenziali estratti, o in base al contenuto) per trovare relazioni e connessioni tra loro; iv) ricercare contenuti nei documenti (es. a testo libero) per recuperare gli stessi o informazioni chiave.

b. Classi utente e caratteristiche

Magistrato: annotare atti processuali per scopi di information retrieval e/o massimazione delle sentenze, consultazione immediata della normativa collegata ai documenti (estratta in modo automatico).

UPP: Un discorso analogo può esser fatto per l'UPP, che tramite questo software può raffinare le informazioni di interesse del magistrato per supportare le sue decisioni.

c. Ambiente operativo

Heimdall è progettato come una Web application modulare. Può essere utilizzato su tutti i sistemi operativi desktop (Linux, Windows, MacOS, ecc) che siano dotati di un applicativo Web browser. Qui di seguito una lista dei browser e relativa versione su cui è stato testato Heimdall:

- Google Chrome (ver 114.0.5735.199)
- Safari (ver 16.0)

d. Documentazione utente

Il presente documento è unica documentazione per gli utenti del software e per gli sviluppatori. Tutte le indicazioni utili sono disponibili nella sezione "Manuale d'uso del Software".

3. Requisiti funzionali

Sulla base delle analisi condotte, si ritiene opportuno proporre e progettare un set di funzionalità, eventualmente integrabili nella Consolle del magistrato, o eseguibili in un tool separato, per l'estensione dei documenti con metadati impiegabili per una più facile gestione degli stessi nelle diverse fasi di trattazione delle cause. La ricerca mira a supportare l'attività di magistrati e addetti all'Ufficio Per il Processo (UPP) elaborando soluzioni software innovative di document enhancement. L'idea proposta è quella di sfruttare le più recenti tecniche di estrazione della conoscenza e di visualizzazione dei dati per favorire l'accesso, la condivisione e la comprensione dell'informazione giuridica necessaria all'esercizio della funzione giurisdizionale e di tutte le attività ad essa strumentali. I paragrafi che seguono descrivono gli aspetti che saranno considerati nelle attività che saranno svolte.

3.1 Obiettivi e funzionalità

Le funzionalità del prodotto riguardano:

- Estrazione e visualizzazione dei riferimenti normativi contenuti nei documenti del fascicolo di causa;
- Annotazione dei documenti: tagging di informazioni rilevanti all'implementazione di funzioni di information retrieval avanzato e per attività di massimazione delle sentenze;
- Ricerca di documenti simili;
- Ricerca di contenuti nei documenti;

- Visualizzazioni che consentano a magistrati e ad addetti all'UPP di interagire in maniera intuitiva con i contenuti dei documenti e di metterli in relazione fra loro.

3.2 Estrazione dei riferimenti normativi dai documenti del fascicolo

L'estrazione automatica di riferimenti normativi rappresenta il punto di partenza per l'implementazione delle seguenti funzionalità:

- accesso rapido al testo vigente della normativa citata negli atti di causa;
- ricerca per riferimento normativo dei documenti del fascicolo;
- identificazione e integrazione di ulteriori dati, documenti e conoscenze (giurisprudenza, dottrina, letteratura grigia, contenuti Web) a partire dalla normativa citata nel documento e dalle caratteristiche di quest'ultima (connessione del documento in esame con altri documenti attraverso i riferimenti normativi)

3.3 Annotazione dei documenti per evidenziare informazioni importanti

L'arricchimento dei documenti dei procedimenti con meta informazioni legate al contenuto dei testi (es. informazioni/tag di tipo semantico, strutturale, normativo) genera diverse opportunità per i magistrati e gli addetti all'UPP. In primis, consente il recupero più agevole dei documenti presenti nel fascicolo, ma fornisce anche la possibilità di implementare soluzioni più efficaci ed intuitive (es. visuali) di information retrieval e di effettuare una ricerca facilitata di informazioni e dati di rilievo presenti nel fascicolo.

Le interlocuzioni avute con i magistrati del tribunale di Benevento hanno suggerito l'individuazione di due diverse tipologie di annotazioni e, conseguentemente, di tag possibilmente utilizzabili all'interno dei documenti. La prima tipologia di annotazione è quella finalizzata alla concettualizzazione del documento. La seconda, invece, risulta strumentale all'attività di massimazione. Ad ognuna di queste categorie di annotazioni corrispondono alcuni tag apponibili all'interno dei documenti. Questi tag possono variare in relazione allo scopo per il quale vengono inseriti nel documento oggetto di annotazione.

Tag per la concettualizzazione

L'attività di annotazione risulta strumentale alla individuazione delle argomentazioni addotte dalle parti a sostegno della propria tesi consentendone un raffronto agevole e immediato. Tale annotazione attiene alle asserzioni relative alle circostanze di fatto, alle conclusioni e alle istanze di merito, istruttorie o relative alle spese di lite presenti all'interno del documento, ma anche alle norme invocate all'interno dello stesso.

Alcuni dei contenuti su cui potrebbe concentrarsi tale annotazione sono:

- 1) Nome del giudice autore della sentenza e quello delle parti coinvolte;
- 2) Conclusioni e richieste avanzate dalle parti;
- 3) Esposizione di circostanze di fatto;
- 4) Questioni di diritto sostanziale;
- 5) Questioni di rito.

La scelta di tali contenuti, che devono scaturire dai dialoghi con i magistrati del Tribunale di Benevento, risulta sicuramente preliminare rispetto ad un'analisi più ampia e raffinata da

sviluppare nelle fasi successive previste dal Progetto (fine tuning durante la sperimentazione).

Tag per la massimazione

L'annotazione si rivela, inoltre, di particolare importanza per la redazione delle massime, principi di diritto estratti da sentenze o ordinanze che il massimatore, dopo un'attenta analisi del provvedimento giudiziario, mette in risalto senza ripercorrere l'iter argomentativo del provvedimento e senza soffermarsi su digressioni e obiter dicta¹. La massimazione è pensata infatti per fornire notizie essenziali e sintetiche circa gli elementi di fatto, necessari a comprendere l'esatta portata del decisum ed il corretto ambito di applicazione del principio massimato².

In questa prospettiva, l'annotazione assume un ruolo strategico consentendo di mettere in evidenza attraverso dei tag quelle partizioni interne alla sentenza che contengono informazioni rilevanti ai fini della scrittura della massima e che possono variare in base alla materia oggetto della sentenza, all'argomento in esame e alla tipologia della controversia trattata. In termini generali, l'annotazione potrebbe concentrarsi su contenuti diversi, come:

- 1) Argomento della sentenza;
- 2) Tipologia di procedimento;
- 3) Motivazione fornita dal giudice;
- 4) Riferimenti giurisprudenziali e normativi richiamati a sostegno di tale motivazione;
- 5) Altri punti nodali della sentenza - "highlights" - utili a individuare il principio di diritto espresso nel provvedimento.

Vale la pena sottolineare che i contenuti in parola non costituiscono tutti elementi necessariamente presenti all'interno della sentenza. La struttura e le informazioni indicate all'interno di quest'ultima possono variare in base a più fattori come, ad esempio, lo stile redazionale dell'autore, la complessità della causa trattata e le modalità di svolgimento del processo. È in ogni caso opportuno per una buona riuscita del lavoro di massimazione, in linea con il criterio di omogeneità cui quest'ultima si ispira³, che le sentenze relative ad un medesimo dominio (es., diritto bancario, diritto di famiglia ecc.) vengano annotate utilizzando la stessa classe di tag.

3.4 Ricerca documenti simili

Il sistema software Heimdall ha l'obiettivo di offrire la possibilità di recuperare documenti simili su diversi livelli. La similarità è potenzialmente espressa sulla base dei seguenti contenuti:

¹ Sul punto di v. Giovanni Amoroso, *Nomofilachia e Massimario, intervento all'incontro L'ufficio del Massimario e del Ruolo della Corte di cassazione: il presente che guarda al passato per pensare al futuro*, Ufficio della Formazione decentrata della Corte di Cassazione, Roma 12 aprile 2017.

² DI IASI, C. (2017). La fata ignorante (a proposito di Ufficio del Massimario e funzione di nomofiliachia). *Questione Giustizia*, 3, 82.

³ Sul punto si v. le linee guida per il massimario civile proposte dall'Ufficio del massimario della Corte Suprema di Cassazione.

- riferimenti normativi contenuti nei documenti (sulla base del principio che documenti che citano leggi simili sono simili)
- intero contenuto in formato testuale (utilizzando tecniche di Text Mining e Natural Language Processing)
- contenuto annotato dall'utente (utilizzando tecniche di Text Mining e Natural Language Processing).

3.5 Ricerca contenuti nei documenti

La ricerca nei contenuti dei documenti è utile al giudice per ritrovare elementi importanti ai fini delle sue attività. La ricerca può avere diversi oggetti:

- contenuto annotato, cioè stralci di testo, frasi o parole annotati con un determinato tag. Tale ricerca può eventualmente essere raffinata utilizzando keyword filtro (es. ricerca di tutte le occorrenze di "my_keyword" annotato con tag "tag1")
- intero contenuto dei documenti, sulla base della classica ricerca full-text

3.6 Casi d'uso

Qui di seguito, si illustrano i casi d'uso del sistema sviluppato. Per ciascuno di essi si identifica un nome (del tipo UC<X>: <nome del caso d'uso>), Attore, Flusso (si distingue qui U per Utente e S per Sistema) e le eventuali precondizioni.

UC1: NUOVO FASCICOLO	
Attore	Magistrato / UPP
Flusso	<p>S: il sistema mostra una pagina con al centro la lista dei fascicoli disponibili organizzati in una griglia.</p> <p>S: il sistema mostra il tasto "Aggiungi fascicolo"</p> <p>U: l'utente clicca (singolo click) sul pulsante "Aggiungi fascicolo"</p> <p>U: l'utente, utilizzando la tastiera, assegna un nome al fascicolo, le parti in causa (tramite appositi campi input di tipo testo).</p> <p>U: l'utente preme il tasto "Invio" sulla tastiera.</p>
Precondizione	L'utente ha effettuato l'accesso ad Heimdall (l'utente apre il browser Web; l'utente inserisce l'URL di Heimdall)

UC2: ACCESSO FASCICOLO	
Attore	Magistrato / UPP
Flusso	<p>S: il sistema mostra una pagina con al centro la lista dei fascicoli disponibili organizzati in una griglia.</p> <p>U: l'utente clicca (singolo click) sull'icona di un fascicolo di interesse dalla griglia.</p> <p>S: il sistema mostra i documenti appartenenti al fascicolo in una griglia.</p>
Precondizione	L'utente ha effettuato l'accesso ad Heimdall

UC3: CARICAMENTO DOCUMENTO	
Attore	Magistrato / UPP
Flusso	<p>S: il sistema mostra i documenti appartenenti al fascicolo in una griglia.</p> <p>S: il sistema mostra il tasto in alto a sinistra "Aggiungi Documento"</p> <p>U: l'utente clicca (singolo click) sul pulsante "Aggiungi Documento"</p> <p>U: l'utente seleziona dal proprio File System il documento da caricare</p> <p>U: L'utente clicca o il tasto "Invio" oppure il tasto "OK"</p> <p>S: il sistema mostra il documento appena aggiunto nel fascicolo all'interno della griglia.</p>
Precondizione	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso ai documenti contenuti in un fascicolo. 2. L'utente ha effettuato l'accesso ad HeimTag.

UC4: VISUALIZZAZIONE RIFERIMENTI GIURIDICI E NORMATIVI	
Attore	Magistrato / UPP
Flusso	<p>S: il sistema mostra i documenti appartenenti al fascicolo in una griglia.</p> <p>S: il sistema mostra, al di sotto delle icone di ciascun documento, un'icona a forma di "libro".</p> <p>U: l'utente clicca sull'icona a forma di "libro"</p>

	<p>S: il sistema, in una nuova pagina, mostra il contenuto del documento (testo) e l'insieme dei riferimenti normativi e giurisprudenziali estratti e raggruppati per tipologia (legge, decreto legislativo, ecc) tramite dei pulsanti.</p> <p>U: l'utente clicca su uno dei raggruppamenti di suo interesse e ne visualizza il contenuto.</p>
Trigger	Un nuovo documento viene caricato in Heimdall. Al caricamento del documento sono estratti i riferimenti giuridici e normativi in esso contenuti.
Precondizione	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso ai documenti contenuti in un fascicolo. 2. Almeno un documento è presente nel fascicolo.

UC5: NAVIGAZIONE / ACCESSO RIFERIMENTI GIURIDICI E NORMATIVI	
Attore	Magistrato / UPP
Flusso	<p>S: il sistema, in una nuova pagina, mostra il contenuto del documento (testo) e l'insieme dei riferimenti normativi e giurisprudenziali estratti e raggruppati per tipologia (legge, decreto legislativo, ecc) tramite dei pulsanti.</p> <p>U: l'utente clicca su uno dei raggruppamenti di suo interesse e ne visualizza il contenuto.</p> <p>U: l'utente clicca (singolo click) su uno dei riferimenti normativi e giurisprudenziali estratti</p> <p>S: il sistema apre, in una nuova pagina Web in una nuova scheda, il testo/la fonte del riferimento cliccato (es. una legge su normattiva.it) .</p>
Precondizione	UC4: VISUALIZZAZIONE RIFERIMENTI GIURIDICI E NORMATIVI

UC6: RICERCA A TESTO LIBERO	
Attore	Magistrato / UPP

Flusso	<p>S: il sistema mostra i documenti appartenenti al fascicolo in una griglia.</p> <p>S: il sistema mostra, al di sopra della griglia, una barra di ricerca di tipo testo ed accanto (a destra) un pulsante "Cerca"</p> <p>U: l'utente digita tramite tastiera una stringa che vuole cercare all'interno dei documenti del fascicolo sfruttando l'apposita barra di ricerca</p> <p>U: l'utente clicca sul tasto "Cerca" (click singolo)</p> <p>S: il sistema mostra solo i documenti che contengono la stringa in input all'interno della griglia.</p>
Precondizione	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso ai documenti contenuti in un fascicolo. 2. Almeno un documento è presente nel fascicolo.

UC7: ANNOTAZIONE	
Attore	Magistrato / UPP
Flusso	<p>S: il sistema mostra i documenti appartenenti al fascicolo in una griglia.</p> <p>S: il sistema mostra, al di sotto delle icone di ciascun documento, un'icona a forma di "martello".</p> <p>U: l'utente clicca sull'icona a forma di "martello"</p> <p>S: il sistema, in una nuova pagina, mostra la dashboard di annotazione</p> <p>U: l'utente può selezionare un tag (dalla sinistra della dashboard) per annotare il documento oggetto tramite un clic singolo</p> <p>U: una volta scelto il tag, l'utente può selezionare il testo da annotare (dalla destra della dashboard) con il tasto sinistro del mouse - tenendo premuto - e rilasciare al termine dell'operazione.</p> <p>S: il sistema mostra il contenuto selezionato evidenziandolo con il colore associato al tag e aggiungendo in sovrapposizione il nome del tag ad inizio e fine della zona evidenziata.</p>
Precondizione	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso ai documenti contenuti in un fascicolo.

	<ol style="list-style-type: none"> 2. Almeno un documento è presente nel fascicolo. 3. L'utente ha effettuato accesso ad HeimTag
--	--

UC8: CARICAMENTO SCHEMA ANNOTAZIONE

Attore	Magistrato / UPP
Flusso	<p>U: L'utente seleziona il menu "Editor" in alto a destra dello schermo e clicca una volta</p> <p>S: il sistema mostra un menu a tendina</p> <p>U: L'utente clicca su "Manage Schemas" dal menu a tendina (click singolo).</p> <p>S: Il sistema mostra una schermata in cui c'è il tasto "Scegli File" e il tasto "Submit".</p> <p>U: l'utente clicca (singolo click) su "Scegli File".</p> <p>U: L'utente seleziona un file di tipo xsd dal File System e poi preme il tasto "Invio" sulla tastiera.</p> <p>U: l'utente clicca (singolo click) su "Submit" per completare l'operazione.</p>
Precondizione	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso ai documenti contenuti in un fascicolo. 2. Almeno un documento è presente nel fascicolo. 3. L'utente clicca sull'icona a forma di "martello" in corrispondenza del documento. 4. L'utente ha effettuato accesso ad HeimTag

UC9: RICERCA DOCUMENTI SIMILI (NORM-BASED)

Attore	Magistrato / UPP
Flusso	<p>S: il sistema mostra i documenti appartenenti al fascicolo in una griglia.</p> <p>S: il sistema mostra, al di sotto delle icone di ciascun documento, un'icona a forma di "doppio cerchio intersecato".</p> <p>U: l'utente clicca sull'icona a forma di "doppio cerchio intersecato"</p> <p>S: il sistema, nella stessa pagina, in una</p>

	colonna a destra a comparsa, mostra la lista dei documenti (icona e nome) simili a quello selezionato con il click al punto precedente. Accanto a ciascun documento, il sistema mostra il numero di riferimenti normativi e giurisprudenziali in comune con il documento scelto.
Precondizione	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso ai documenti contenuti in un fascicolo. 2. Almeno due documenti sono presenti nel database del sistema.

UC10: RICERCA DOCUMENTI SIMILI (TEXT-BASED SEMANTICO)	
Attore	Magistrato / UPP
Flusso	<p>S: il sistema mostra i documenti appartenenti al fascicolo in una griglia.</p> <p>S: il sistema mostra, al di sotto delle icone di ciascun documento, un'icona con l'aspetto di un cervello con l'acronimo NLP.</p> <p>U: l'utente clicca sull'icona con con l'aspetto di un cervello con l'acronimo NLP</p> <p>S: il sistema, nella stessa pagina, in una colonna a destra a comparsa, mostra la lista dei documenti (icona e nome) simili a quello selezionato con il click al punto precedente. Accanto a ciascun documento, il sistema mostra un punteggio di similarità calcolato sulla base della somiglianza dell'intero contenuto testuale (embedding BERT + cosine similarity).</p>
Precondizione	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso ai documenti contenuti in un fascicolo. 2. Almeno due documenti sono presenti nel database del sistema.

Inoltre, sono stati ipotizzati una serie di casi d'uso riguardanti potenziali funzionalità del sistema e modalità di interazione ulteriori rispetto a quelle viste precedentemente.

[SVILUPPI FUTURI] UC11: RICERCA DOCUMENTI SIMILI (TAG-BASED SEMANTICO)	
Attore	Magistrato / UPP
Flusso	<p>S: il sistema mostra i documenti appartenenti al fascicolo in una griglia.</p> <p>S: il sistema mostra, al di sotto delle icone</p>

	<p>di ciascun documento, un'icona con l'aspetto di un cervello con l'acronimo NLP.</p> <p>U: l'utente clicca sull'icona con l'aspetto di un cervello con l'acronimo NLP.</p> <p>S: il sistema, nella stessa pagina, in una colonna a destra a comparsa, mostra la lista dei documenti (icona e nome) simili a quello selezionato con il click al punto precedente. Accanto a ciascun documento, il sistema mostra degli score di similarità calcolati sulla base della somiglianza tra gli stralci di testo annotati con medesimo tag.</p>
Precondizione	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso ai documenti contenuti in un fascicolo. 2. Almeno due documenti sono presenti nel database del sistema. 3. Almeno due documenti sono stati annotati con l'apposita funzionalità (vedi UC7)

[SVILUPPI FUTURI] UC12: RICERCA DOCUMENTI SIMILI (TEXT-BASED LESSICALE)

Attore	Magistrato / UPP
Flusso	<p>S: il sistema mostra i documenti appartenenti al fascicolo in una griglia.</p> <p>S: il sistema mostra, al di sotto delle icone di ciascun documento, un'icona con l'acronimo TF-IDF.</p> <p>U: l'utente clicca sull'icona con l'acronimo TF-IDF</p> <p>S: il sistema, nella stessa pagina, in una colonna a destra a comparsa, mostra la lista dei documenti (icona e nome) simili a quello selezionato con il click al punto precedente. Accanto a ciascun documento, il sistema mostra un punteggio di similarità calcolato sulla base della somiglianza dell'intero contenuto testuale (Term Frequency-Inverse Document Frequency + cosine similarity).</p>
Precondizione	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso ai documenti contenuti in un fascicolo. 2. Almeno due documenti sono presenti nel database del sistema.

[SVILUPPI FUTURI] UC13: RICERCA DOCUMENTI SIMILI LESSICALE/SEMANTICO ON-THE-FLY

Attore	Magistrato / UPP
Flusso	<p><i>U: l'utente clicca sull'icona con l'acronimo TF-IDF (rispettivamente l'icona a forma di cervello con acronimo NLP) dall'editor della consolle del Magistrato.</i></p> <p><i>S: il sistema, in una colonna a destra a comparsa, mostra la lista dei documenti (icona e nome) simili a quello che l'utente sta redigendo. Accanto a ciascun documento, il sistema mostra degli score di similarità calcolati sulla base della somiglianza tra i contenuti testuali dei documenti. La similarità è calcolata con cosine similarity. Può essere effettuato il calcolo sia sulla rappresentazione lessicale del testo tramite TF-IDF (vedi UC12), sia su quella semantica tramite BERT (vedi UC10).</i></p>
Precondizione	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso alla Consolle del Magistrato. 2. L'utente sta scrivendo un documento con l'editor della Consolle del Magistrato 3. Almeno due documenti sono presenti nel database del sistema.

[SVILUPPI FUTURI] UC14: RICERCA DOCUMENTI SIMILI LESSICALE/SEMANTICO ON-THE-FLY (TAG-BASED)

Attore	Magistrato / UPP
Flusso	<p><i>U: l'utente seleziona uno stralcio di testo annotato dall'editor della Consolle del Magistrato</i></p> <p><i>U: l'utente clicca sull'icona con l'acronimo TF-IDF (rispettivamente l'icona a forma di cervello con acronimo NLP) dall'editor della consolle del Magistrato.</i></p> <p><i>S: il sistema, in una colonna a destra a comparsa, mostra la lista dei documenti (icona e nome) che riportano contenuti simili a quello che l'utente ha selezionato/annotato. Accanto a ciascun risultato, il sistema mostra degli score di similarità calcolati sulla base della</i></p>

	<p>somiglianza tra i contenuti annotati (con lo stesso tag) dei documenti. La similarità è calcolata con cosine similarity. Può essere effettuato il calcolo sia sulla rappresentazione lessicale del testo tramite TF-IDF (vedi UC12), sia su quella semantica tramite BERT (vedi UC10).</p>
<p>Precondizione</p>	<ol style="list-style-type: none"> 1. L'utente ha effettuato l'accesso alla Consolle del Magistrato. 2. L'utente sta scrivendo un documento con l'editor della Consolle del Magistrato 3. Almeno due documenti sono presenti nel database del sistema. 4. Almeno due documenti sono stati annotati con l'apposita funzionalità (vedi UC7)

4. Requisiti interfaccia esterna

a. Requisiti GUI

Tutte le funzionalità del sistema sviluppato sono accessibili tramite interfaccia grafica Web-based.

Ogni funzionalità è accessibile tramite il click su un apposito pulsante. L'interazione con il sistema avviene mediante click singoli su pulsanti.

Il sistema è pensato per contesti desktop. Non è possibile usufruire del sistema tramite device mobili.

b. Altri tipi di interfaccia

Il sistema non è dotato di ulteriori interfacce.

5. Requisiti non funzionali

a. Requisiti di sicurezza

Il prototipo sviluppato non gestisce la multi-utenza e l'autenticazione, ma è pensato per un unico utente (implicitamente autenticato) che ha permessi da amministratore, cioè la possibilità di caricare, modificare, eliminare, interagire con tutti i dati ed utilizzare tutte le funzionalità.

Non è possibile tracciare le operazioni di questo utente.

b. Requisiti prestazionali e scalabilità

Il sistema Heimdall ed il micro-servizio di annotazione HeimTag sono stati progettati in forma prototipale per single-user.

6. Architettura del software

a. Paradigma di riferimento

Il sistema è progettato con il paradigma Model View Controller.

b. Architettura del software

Il sistema Heimdall (con micro-servizio HeimTag per l'annotazione dei documenti) è sviluppato secondo l'architettura mostrata in Figura 1:

- Front-end: per quanto concerne Heimdall, il front-end è sviluppato utilizzando Bootstrap 5, JavaScript e CSS per i fogli di stile; per quanto concerne il micro-servizio HeimTag, si utilizza il framework Vue.js;
- Back-end: il back-end di Heimdall è sviluppato in Java 17 con framework Spring Boot e Java Jersey. All'interno si fa uso di librerie quali Apache Tika, Java Tesseract, e Linkoln. Il back-end del micro-servizio HeimTag è sviluppato in Python, con il framework Django.
- Storage: il database è NoSQL document-oriented, MongoDB, ed è condiviso sia da Heimdall che HeimTag. Il micro-servizio di annotazione utilizza inoltre un database SQLite interno per la gestione di file di configurazione e simili.

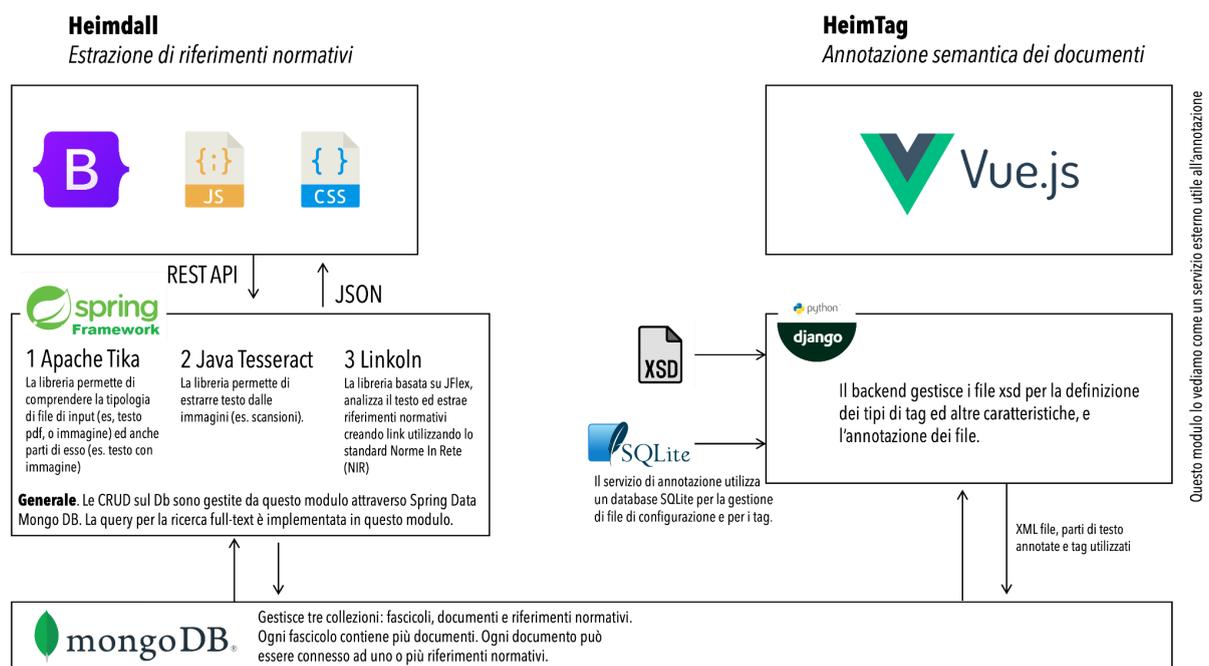


Figura 1. Architettura generale del sistema Heimdall con micro-servizio HeimTag

c. Assunzioni e Dipendenze

Nello sviluppo dell'applicativo Heimdall si assume che il sistema possa accedere ai documenti contenuti nei fascicoli d'ufficio gestiti dall'UPP e dai magistrati, normalmente accessibili mediante Consolle del Magistrato.

Il funzionamento del sistema dipende dalla libreria Linkoln sviluppata dall'Istituto di Informatica Giuridica e Sistemi Giudiziari (IGSG) del Consiglio Nazionale delle Ricerche.

d. Ambiente operativo

Il sistema sviluppato è pensato per agire come plug-in / estensione della Consolle del Magistrato.

7. Descrizione delle funzionalità erogate

Il sistema Heimdall offre le seguenti funzionalità. Per ciascuna di essa, si offre a) una descrizione generale, b) i vantaggi/benefici per gli utenti.

In bianco, vengono mostrate le funzionalità sviluppate all'interno del dimostratore (prototipo di Heimdall); in grigio, vengono evidenziate funzionalità future potenzialmente integrabili nel sistema.

Funzionalità

Importazione documenti

Descrizione. Heimdall offre la possibilità di caricare documenti dal proprio File System sfruttando un apposito modulo e cliccando sul pulsante "Carica documento".

Vantaggi/benefici per gli utenti. L'importazione dei documenti in Heimdall, abilita l'estrazione automatica del contenuto testuale e dei riferimenti normativi. La funzionalità è quindi condizione per l'utilizzo degli strumenti più avanzati del tool.

Accesso ai documenti di un fascicolo

Descrizione. Selezionando un fascicolo di causa è possibile ottenere tutti i documenti al suo interno

Vantaggi/benefici per gli utenti. Tramite la GUI di Heimdall si può accedere in modo intuitivo ai documenti di un fascicolo, con una rappresentazione a griglia simile ad altri software oggi disponibili (es. Google Drive).

Estrazione contenuto testuale del documento

Descrizione. Heimdall, attraverso l'utilizzo combinato di Apache Tika e Java Tesseract riesce ad estrapolare il contenuto testuale dei documenti e lo salva all'interno di un database MongoDB.

Vantaggi/benefici per gli utenti. La funzionalità permette di estrapolare contenuto testuale da documenti doc, docx, pdf (testuali e scansioni), txt ed anche immagini jpeg. Molti documenti sono risultato di scansioni ed accedere al testo è di fondamentale importanza per gli utenti di Heimdall, allo scopo di: a) abilitare l'estrazione dei riferimenti normativi e giurisprudenziali dal documento, b) abilitare le funzionalità di annotazione.

Estrazione dei riferimenti giuridici e normativi nei documenti

Descrizione. mediante la libreria Linkoln, il sistema trova riferimenti giuridici e normativi nei documenti importati e genera URL navigabili ai testi vigenti di tali norme.

Vantaggi/benefici per gli utenti. L'estrazione automatica di riferimenti normativi, oltre che consentire accesso rapido e diretto (tramite URL) al testo vigente della normativa citata negli atti di causa, rappresenta il punto di partenza per l'implementazione di funzioni diverse: (i) ricerca per riferimento normativo dei documenti del fascicolo; (ii) identificazione e integrazione di ulteriori dati, documenti e conoscenze (giurisprudenza, dottrina, letteratura grigia, contenuti Web) a partire dalla normativa citata nel documento

e dalle caratteristiche di quest'ultima (connessione del documento in esame con altri documenti attraverso i riferimenti normativi); (iii) ricerca documenti simili in termini di riferimenti normativi contenuti.

Ricerca a testo libero del contenuto di un documento

Descrizione. L'utente ha la possibilità di inserire del testo (keyword di ricerca) in linguaggio naturale all'interno di un campo di testo appositamente predisposto che fungerà da query / interrogazione al database MongoDB per cercare tutti i documenti che contengono il testo inserito.

Vantaggi/benefici per gli utenti. Attraverso la Consolle del Magistrato non è possibile effettuare ricerche full-text nei testi contenuti nei documenti dei fascicoli. Con Heimdall, questo è possibile. Il vantaggio principale per gli utenti è di poter recuperare informazioni salienti per il processo in modo facile, rapido ed intuitivo.

Ricerca documenti "simili" in termini di riferimenti normativi e giuridici contenuti

Descrizione. Utilizzando un apposito pulsante accanto all'icona di un documento, il sistema restituisce (in ordine decrescente di similarità) la lista dei documenti più simili a quello selezionato. La similarità è espressa con un valore intero. Il set dei documenti in output è mostrato in una sezione-colonna a destra dell'interfaccia.

Vantaggi/benefici per gli utenti. Tale funzionalità permette all'utente di recuperare documenti simili anche in altri casi/fascicoli. Il confronto con altri fascicoli può rendere più fluido il processo decisionale del magistrato (es. si confronta con fascicoli simili del passato su cui ha già preso decisioni).

Annotazione documenti

Descrizione. il sistema permette di selezionare uno schema di annotazione (concettualizzazione, massimazione) e mostra una schermata divisa sostanzialmente in due grandi aree: la prima, a sinistra, utilizzata per la selezione dei tag e delle loro proprietà, la seconda per mostrare il contenuto testuale del documento oggetto di annotazione. Una volta scelto un tag e impostate le sue proprietà, basterà selezionare il testo da annotare per aggiungere il tag al documento. Una volta cliccato sul tasto "Salva", il documento sarà memorizzato in formato xml con tutti i tag inseriti dall'utente.

Vantaggi/benefici per gli utenti. L'arricchimento dei documenti di causa con metainformazione legata al contenuto dei testi (es. informazioni/tag di tipo semantico, strutturale, normativo) genera opportunità diverse per i magistrati e gli addetti all'UPP. In primis consente il recupero più agevole dei documenti presenti nel fascicolo ma fornisce anche la possibilità di implementare soluzioni più efficaci ed intuitive (es. visuali) di information retrieval e di effettuare una ricerca facilitata di informazioni e dati di rilievo

presenti nel fascicolo di causa. Uno schema di tagging per la concettualizzazione è strumentale alla individuazione delle argomentazioni addotte dalle parti a sostegno della propria tesi consentendone un raffronto agevole e immediato. Uno schema di tagging per la massimazione è utile per mettere in evidenza attraverso dei tag quelle partizioni interne alla sentenza che contengono informazioni rilevanti ai fini della scrittura della massima e che possono variare in base alla materia oggetto della sentenza, all'argomento in esame e alla tipologia della controversia trattata

Aggiunta di un nuovo schema di annotazione

Descrizione. il sistema permette di aggiungere un nuovo schema di annotazione utilizzando l'opzione "Manage Schemas" disponibile dal menu a tendina "Editor" del micro-servizio HeimTag. Una volta selezionato, l'utente può scegliere un proprio file xsd dal File System per l'upload. Tale file xsd può essere usato per annotare documenti.

Vantaggi/benefici per gli utenti. L'aggiunta di nuovi schemi di annotazione può favorire diversi vantaggi agli end-users del sistema, in quanto rende la funzionalità di annotazione più personalizzabile. Inoltre, con tale funzionalità il sistema si presta ad essere utilizzato per scopi nuovi e non inizialmente previsti.

Ricerca documenti "simili" in termini di contenuto testuale (semantico)

Descrizione. Utilizzando un apposito pulsante accanto all'icona di un documento, il sistema restituisce (in ordine decrescente di similarità) la lista dei documenti più simili a quello selezionato. La similarità è espressa con un valore nel range $R[-1,1]$ ed è calcolata sulla base dell'intero contenuto testuale dei documenti attraverso la cosine similarity. Il calcolo della cosine similarity viene effettuato sugli incorporamenti (embedding) dei testi dei documenti tramite il modello BERT multilingua. Il set dei documenti in output è mostrato in una sezione-colonna a destra dell'interfaccia.

Vantaggi/benefici per gli utenti. Tale funzionalità permette all'utente di recuperare documenti/sentenze simili anche in altri casi/fascicoli. Il confronto con altri fascicoli può rendere più fluido il processo decisionale del magistrato (es. si confronta con sentenze simili del passato). Inoltre, una volta che il sistema proposto è integrato nell'editor della Consolle del Magistrato, la funzionalità può essere richiamata al bisogno durante la scrittura delle sentenze.

Modalità. Statico; On-the-fly (da editor Consolle del Magistrato, quando si redigono sentenze); On-the-fly contestuale (analizzando stralci di testo annotati in un documento durante la sua stesura)

[SVILUPPI FUTURI] Ricerca documenti "simili" in termini di contenuto testuale annotato

Descrizione. Utilizzando un apposito pulsante accanto all'icona di un documento, il sistema restituisce la lista dei documenti più simili a quello selezionato. Accanto ai documenti restituiti, il sistema mostra dei valori di similarità che variano nel range $R[-1,1]$,

calcolati sulla base della similarità del contenuto testuale annotato nei documenti con un certo tag. Il set dei documenti in output è mostrato in una sezione-colonna a destra dell'interfaccia.

Vantaggi/benefici per gli utenti. Tale funzionalità permette all'utente di recuperare documenti/sentenze simili anche in altri casi/fascicoli. Il confronto con specifiche parti del testo annotato anche in altri fascicoli può essere di aiuto per la decisione finale del giudice (es. ricerca per motivazioni delle parti e fattispecie).

[*SVILUPPI FUTURI*] Ricerca documenti "simili" in termini di contenuto testuale (lessicale)

Descrizione. Utilizzando un apposito pulsante accanto all'icona di un documento, il sistema restituisce la lista dei documenti più simili a quello selezionato. Il calcolo della similarità è effettuato utilizzando la cosine similarity sui vettori risultanti dall'applicazione della tecnica di text mining TF-IDF. Accanto ai documenti restituiti, il sistema mostra dei valori di similarità che variano, quindi, nel range $R[-1,1]$. Il set dei documenti in output è mostrato in una sezione-colonna a destra dell'interfaccia.

Vantaggi/benefici per gli utenti. Tale funzionalità permette all'utente di recuperare documenti/sentenze simili anche in altri casi/fascicoli. Il confronto con specifiche parti del testo annotato anche in altri fascicoli può essere di aiuto per la decisione finale del giudice (es. ricerca per motivazioni delle parti e fattispecie). Inoltre, una volta che il sistema proposto è integrato nell'editor della Consolle del Magistrato, la funzionalità può essere richiamata al bisogno durante la scrittura delle sentenze.

Modalità. Statico; On-the-fly (da editor Consolle del Magistrato, quando si redigono sentenze); On-the-fly contestuale (analizzando stralci di testo annotati in un documento durante la sua stesura)

8. Descrizione delle fasi di sviluppo ed integrazione nei sistemi IT della giustizia

Heimdall è pensato per agire come plug-in della Consolle del Magistrato. Sono previste le seguenti fasi (F) di sviluppo e integrazione (accanto a ciascuna di esse, è riportato l'effort misurato in ore/uomo):

- F1: studio preliminare della Consolle del Magistrato per identificare i punti e le modalità di integrazione [150].
- F2a: sviluppo API per il dialogo tra Heimdall e i database accessibili tramite Consolle del Magistrato, al fine di poter avere accesso ai fascicoli, documenti, metadati ad essi associati. Sviluppo APIs per il dialogo tra Heimdall e database contenenti riferimenti normativi e giurisprudenziali Italiani ed Europei. [200]
- F2b: definizione schemi di annotazione (tag e proprietà) con gli stakeholders e loro implementazione. [300]
- F3: sviluppo gestione multi-utente (con ruoli e permessi) in Heimdall. [300]

- F4a: integrazione API output di F2a in Heimdall. [100]
- F4b: sviluppo ed integrazione API per la comunicazione Consolle del Magistrato-Heimdall per l'accesso rapido al plugin, garantire consistenza dei dati tra i due sistemi, abilitare l'utilizzo on the fly delle nuove funzionalità dall'editor delle sentenze. [600]
- F5a: refactoring interfaccia grafica. [500]
- F5b: ottimizzazione/parallelizzazione delle funzionalità di Heimdall [200]
- F6a: verifica, testing e convalida delle funzionalità di Heimdall in termini di accuratezza dei risultati, prestazioni, scalabilità in un ambiente rilevante (test pilota con operatori della giustizia). [600]
- F6b: fine-tuning di Heimdall ed APIs. [400]
- F7: distribuzione e dimostrazione del sistema in ambiente di lavoro operativo. [150]

Manuale d'uso del software

8. Installazione

a. Requisiti di sistema

Heimdall deve essere installato su una macchina server dotata delle seguenti librerie (altre specifiche sono indicate nel file pom.xml dell'applicazione):

- Java ver 17
- Spring Boot
- Java Tesseract ver 4 (Tess4J 5.6.0)
- Tesseract Open Source OCR Engine (standalone version)
- Apache Tika (ver 2.6.0)
- Libreria Linkoln, sviluppata da IGSG del CNR (ver 3.2.3)
- Spring Data Mongo DB
- Mongo DB Community Edition
- Sass
- asgiref (ver 3.3.4)
- autopep8 (ver 1.5.7)
- Django (ver 3.2)
- django-appconf (ver 1.0.5)
- django-cors-headers (ver 3.7.0)
- django-crispy-forms (ver 1.11.2)
- django-select2 (ver 8.1.1)
- django-rest-framework (ver 3.12.4)
- elementpath (ver 4.1.0)
- lxml (ver 4.9.2)
- numpy (ver 1.24.2)
- pandas (ver 1.5.3)
- pycodestyle (ver 2.7.0)
- python-dateutil (ver 2.8.2)
- pytz (ver 2021.1)
- simple-dorff (ver 0.0.2)
- six (ver 1.16.0)
- sqlparse (ver 0.4.1)
- toml (ver 0.10.2)
- xmlschema (ver 2.2.2)

b. Installazione del software

Il sistema Heimdall è pensato per una macchina server. Per installare Heimdall sul server, bisogna eseguire passo passo le istruzioni riportate qui di seguito, utilizzando la linea di comando.

1. `sudo apt update`

2. Installare Python e Java
 - a. `sudo apt install openjdk-17-jdk openjdk-17-jre python3`
3. Installare NodeJS 16 sulla macchina
 - a. `cd ~`
 - b. `curl -sL https://deb.nodesource.com/setup_16.x -o nodesource_setup.sh`
 - c. `sudo bash nodesource_setup.sh`
 - d. `sudo apt-get install nodejs`
4. Installare Tesseract OCR sulla macchina
 - a. `sudo add-apt-repository ppa:alex-p/tesseract-ocr-devel`
 - b. `sudo apt install -y tesseract-ocr`
5. Installare Maven sulla macchina
 - a. `wget`
<https://dlcdn.apache.org/maven/maven-3/3.9.3/binaries/apache-maven-3.9.3-bin.tar.gz>
 - b. `tar -xvf apache-maven-3.9.3-bin.tar.gz`
 - c. `mv apache-maven-3.9.3 /opt/`
6. Installare MongoDB sulla macchina
 - a. `sudo apt-get install gnupg curl`
 - b. `curl -fsSL https://pgp.mongodb.com/server-6.0.asc | \`
`sudo gpg -o /usr/share/keyrings/mongodb-server-6.0.gpg \`
`--dearmor`
 - c. `echo "deb [arch=amd64,arm64`
`signed-by=/usr/share/keyrings/mongodb-server-6.0.gpg]`
`https://repo.mongodb.org/apt/ubuntu jammy/mongodb-org/6.0 multiverse"`
`| sudo tee /etc/apt/sources.list.d/mongodb-org-6.0.list`
 - d. `sudo apt-get update`
 - e. `sudo apt-get install -y mongodb-org`
 - f. `sudo systemctl daemon-reload`
 - g. `sudo systemctl start mongod`
 - h. `sudo systemctl enable mongod`
7. Seguire la guida per l'installazione di Heimtag esclusa la parte di creazione dell'utente e del gruppo
 - a. `npm install --global yarn --legacy-peer-deps`
 - b. `npm install --global node-prune --legacy-peer-deps`
 - c. `python3 -m venv env`
 - d. `cd env\Scripts\`
 - e. `.\activate.bat`
 - f. `pip install -r .\Heimtag\req.txt`
 - g. `python3 \Heimtag\manage.py createsuperuser`
 - h. `python3 \Heimtag\manage.py createdefaultgroups <superuser_name>`
 - i. `cd \Heimtag\ui`
 - j. `npm uninstall node-sass`

- k. `npm i -D sass --legacy-peer-deps`
- l. `yarn install`

c. Configurazione del software

Il sistema non prevede ulteriori configurazioni dopo l'installazione. Non è possibile settare preferenze o personalizzare specifiche.

d. Aggiornamento del software

Attualmente, non è prevista una modalità di aggiornamento del sistema.

e. Disinstallazione del software

1. Eliminare le cartelle di HeimTag ed Heimdall
2. Eliminare MongoDB
 - a. `sudo service mongod stop`
 - b. `sudo apt-get purge mongodb-org*`
 - c. `sudo rm -r /var/log/mongodb /var/lib/mongodb`
3. Eliminare Maven
 - a. `sudo rm -rf /opt/apache-maven-3.9.3`
4. Eliminare NodeJS
 - a. `sudo apt-get remove nodejs`
 - b. `cd /etc/apt/sources.list.d`
 - c. `rm -rf nodesource.list nodesource.list.distUpgrade`
 - d. `sudo apt-get update`
 - e. `sudo apt-get purge --auto-remove nodejs`
5. Eliminare Tesseract OCR
 - a. `sudo apt remove --autoremove tesseract-ocr tesseract-ocr-*`

9. Guida all'uso

a. Accesso al software

Il sistema Heimdall è una Web application che, eseguita su un server con che espone un IP pubblico, è raggiungibile tramite un Web browser utilizzando un indirizzo socket IP:<porta>/index.html. L'applicativo fornisce una Graphic User Interface renderizzata tramite il Web browser tramite la quale gli utenti possono interagire con le varie funzionalità. Non c'è altra modalità di interazione, es. linea di comando.

b. Interfaccia utente

L'interfaccia del sistema Heimdall è Web based.

Il sistema permette di importare fascicoli (tasto "Aggiungi fascicolo") o selezionare fascicoli dalla griglia centrale (come riportato in Figura 2).

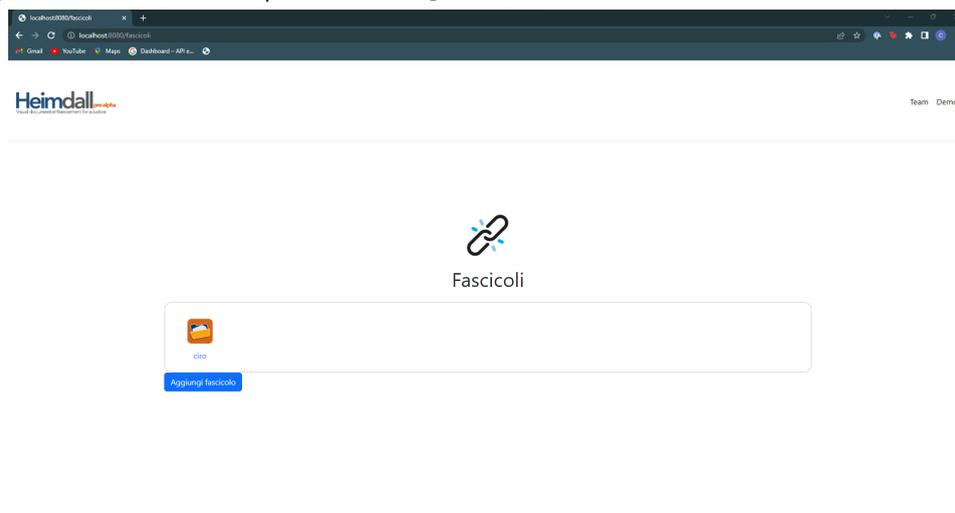


Figura 2: Schermata iniziale di Heimdall

Una volta effettuato l'accesso ad un fascicolo, viene visualizzato il suo contenuto al centro del layout (Figura 3). Ogni documento facente parte del fascicolo viene identificato da un riquadro a sfondo bianco. Ogni documento è disposto secondo una griglia a tre colonne al centro della schermata. Per ciascun documento, si riporta il nome ed un'icona che ne identifica il tipo. Al di sotto del nome del documento, vi sono tre icone (libro, martello, doppio cerchio intersecato) che abilitano altrettante funzionalità (in seguito descritte) e due pulsanti:

- "Visualizza": permette di visualizzare il contenuto testuale del documento
- "Elimina": permette di eliminare il documento dal fascicolo e dalla base di dati MongoDB

Al di sopra della griglia dei documenti, il sistema mostra una barra di ricerca per la funzionalità di ricerca full-text e l'associato pulsante tipo *submit* "Cerca". La ricerca può avvenire anche esclusivamente all'interno di tag di un certo tipo ("cerca <x> all'interno dei

<tag> nei documenti del <fascicolo>), per cui è prevista una sezione (con radio buttons) atta a permettere all'utente la selezione del tag per raffinare la ricerca full-text.

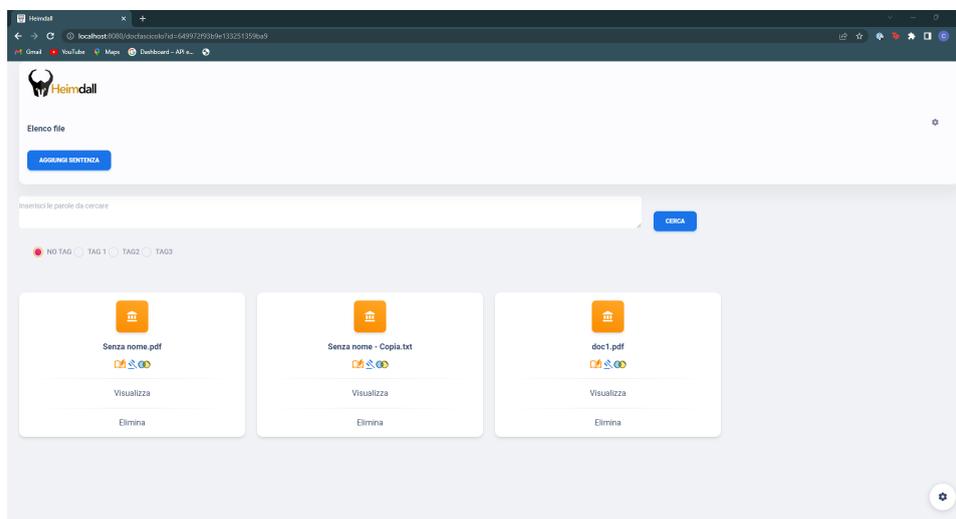


Figura 3: griglia dei documenti in un fascicolo selezionato. Da qui è possibile accedere ad altre funzionalità del sistema.

Tramite il pulsante in alto a sinistra, "Aggiungi Documento" è possibile caricare qualsiasi tipo di documento. Cliccando su tale pulsante, si giunge alla funzionalità di caricamento, riportata in Figura 4. Qui è possibile caricare documenti utilizzando il pulsante "new docs" e rimuovere documenti con il pulsante "remove" accanto a ciascun documento.

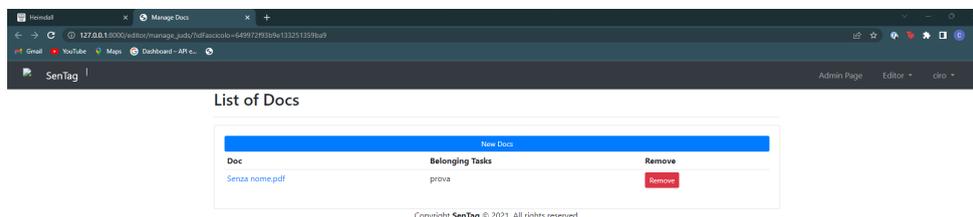


Figura 4: Aggiunta nuovi documenti al fascicolo

Cliccando sull'icona a forma di libro in Figura 3 sotto un documento X, è possibile utilizzare la funzionalità di accesso diretto ai riferimenti normativi e giurisprudenziali all'interno del documento X. Tali riferimenti vengono visualizzati raggruppati in "Decreti legislativi", "leggi", e "Altro" (vedi Figura 5). Cliccando su uno di questi pulsanti si ottiene la lista dei riferimenti direttamente navigabili *onclick*.

In basso, viene mostrato il contenuto testuale del documento.

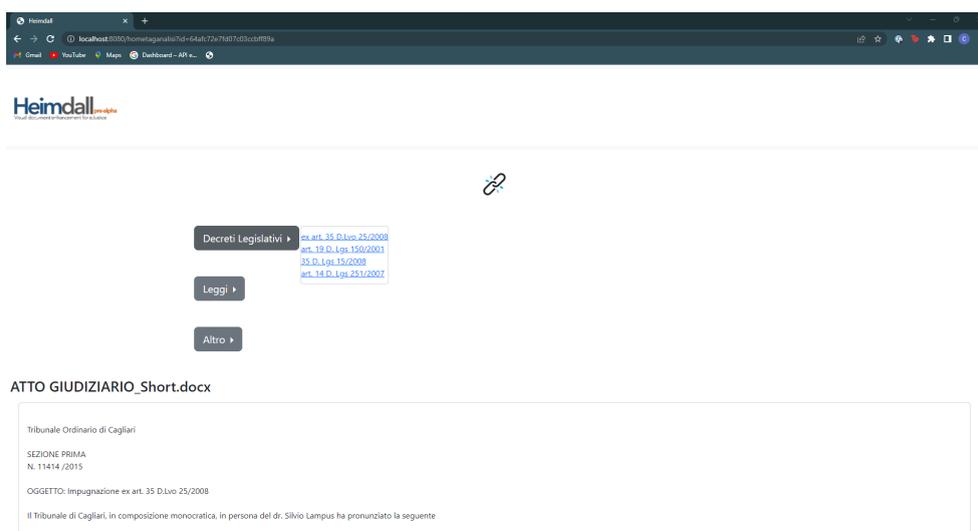


Figura 5: accesso ai riferimenti normativi e giurisprudenziali contenuti nel documento

Utilizzando il pulsante-icona a forma di doppio cerchio intersecato di un documento X è possibile accedere alla funzionalità di Similarità. Tale funzione permette di trovare documenti simili ad X in termini di riferimenti normativi e giurisprudenziali. Il risultato verrà mostrato in una colonna sulla destra della schermata (vedi Figura 6).

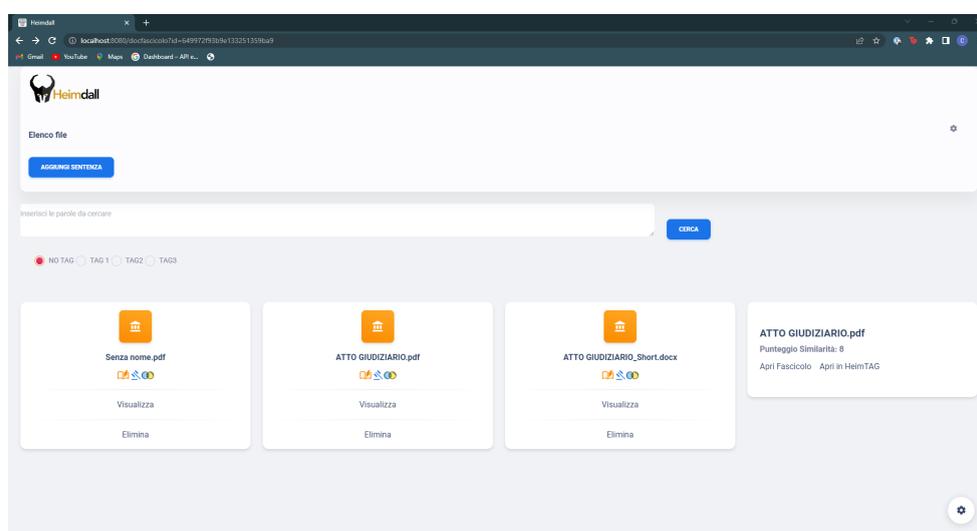


Figura 6: funzione di similarità basata sui riferimenti normativi e giurisprudenziali.

Allo stesso modo, è possibile accedere alla funzionalità di Similarità basata sul testo. Basterà cliccare sull'icona a forma di cervello con l'acronimo "NLP" ed il sistema fornirà in output il set di documenti più simile (cosine similarity) al documento X (vedi Figura 7).

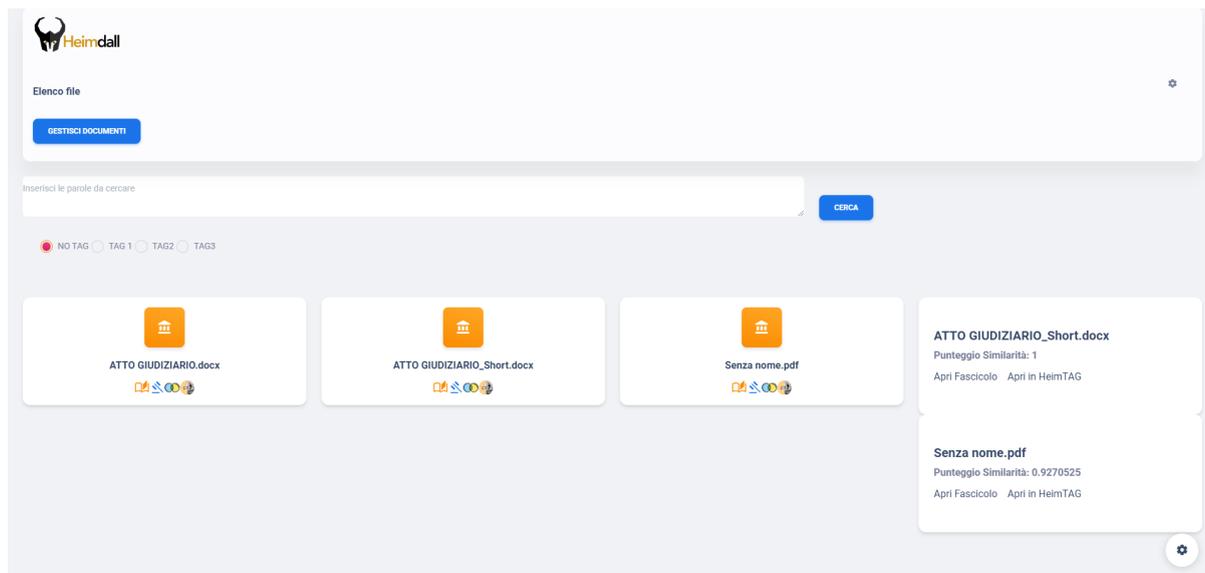


Figura 7: funzione di similarità basata sul testo.

Infine, cliccando sul pulsante-icona a forma di martello in corrispondenza di un documento X è possibile accedere alla funzionalità di annotazione per il documento X (Figura 8). In questo caso, il layout è diviso in due aree. A sinistra, l'utente ha a disposizione il set dei tag che può utilizzare per l'annotazione (con eventuali proprietà da poter assegnare al tag); a destra, l'utente visualizza l'interno contenuto testuale del documento X e può annotarlo nel seguente modo. Si seleziona un tag T dal pannello a sinistra - se necessario si impostano le proprietà -, si seleziona poi il testo da annotare con quel determinato tag T e il risultato è il testo evidenziato con il colore associato a T e l'indicazione di inizio e fine stralcio annotato con il nome di T.

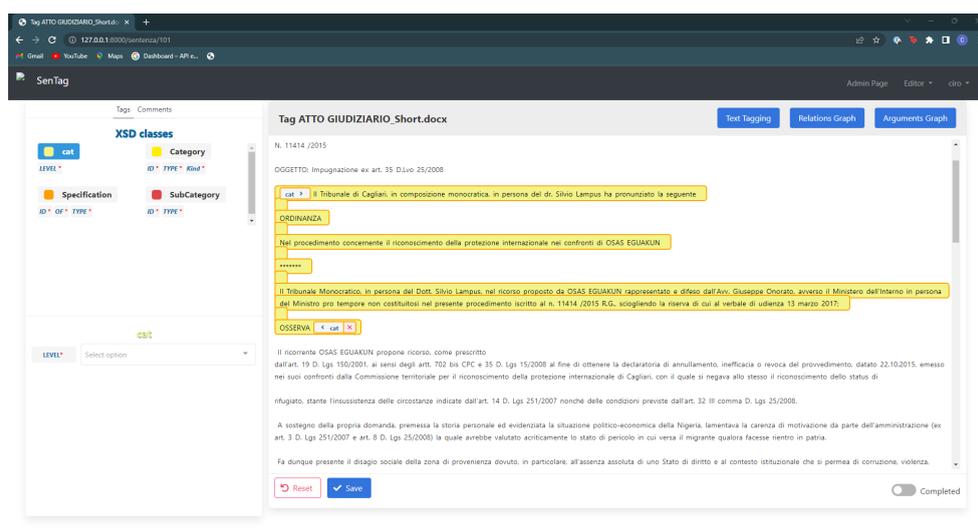


Figura 8: funzionalità di annotazione dei documenti

È inoltre possibile aggiungere nuovi schemi di annotazione utilizzando la funzionalità mostrata in Figura 9. L'utente utilizza il micro-servizio di annotazione. L'utente seleziona il menu "Editor" in alto a dx dello schermo, poi clicca su "Manage Schemas". Qui può aggiungere un nuovo schema xsd di annotazione.

Il sistema permette di selezionare un file xsd dal proprio File System e di importarlo nel micro-servizio di annotazione.



Figura 9: aggiunta nuovo schema di annotazione

c. Operazioni di base

Nuovo fascicolo

1. L'utente clicca sul pulsante "Nuovo Fascicolo"
2. L'utente inserisce un nome per il fascicolo che sta creando nell'apposito campo di testo con sfondo bianco
3. L'utente clicca il tasto "Invio" oppure clicca sul pulsante "OK" visualizzato sotto il campo di testo

Caricamento documento

1. L'utente apre la dashboard di Heimdall che mostra il set di fascicoli a disposizione
2. L'utente clicca sull'icona di un fascicolo, accedendovi
3. L'utente clicca sul pulsante "Carica documento"
4. L'utente seleziona il documento dal proprio File System e con un doppio clic lo seleziona per l'importazione in Heimdall

Ricerca full-text

1. L'utente apre la dashboard di Heimdall che mostra il set di fascicoli a disposizione
2. L'utente clicca sull'icona di un fascicolo, accedendovi
3. L'utente visualizza tutti i documenti del fascicolo con delle icone disposte in una griglia con tre colonne
4. L'utente digita le parole (la stringa) oggetto di ricerca all'interno del database nell'apposito campo di testo visualizzabile nella zona centrale dell'interfaccia.
5. L'utente clicca sul pulsante "Ricerca"
6. L'utente visualizza il risultato della ricerca (il sistema esclude i documenti che non contengono la stringa di ricerca)

Visualizza riferimenti normativi contenuti nel documento

1. L'utente apre la dashboard di Heimdall che mostra il set di fascicoli a disposizione
2. L'utente clicca sull'icona di un fascicolo, accedendovi
3. L'utente visualizza tutti i documenti del fascicolo con delle icone disposte in una griglia con tre colonne
4. L'utente clicca il pulsante a forma di libro accanto all'icona del documento d'interesse (root document)
5. L'utente visualizza i riferimenti normativi e giurisprudenziali contenuti nel documento (raggruppati per tipologia) e il contenuto testuale completo del documento selezionato.

Ricerca documenti simili (basata sui riferimenti normativi e giurisprudenziali comuni)

1. L'utente apre la dashboard di Heimdall che mostra il set di fascicoli a disposizione
2. L'utente clicca sull'icona di un fascicolo, accedendovi
3. L'utente visualizza tutti i documenti del fascicolo con delle icone disposte in una griglia con tre colonne
4. L'utente clicca il pulsante a forma di doppio cerchio intersecato accanto all'icona del documento d'interesse (root document)
5. L'utente visualizza, a destra dell'interfaccia, una colonna con l'insieme dei documenti simili in termini di riferimenti normativi. Sotto ciascuna delle icone che rappresentano i documenti, il sistema riporta un valore numerico che equivale al numero di riferimenti in comune tra il documento selezionato (root document) e il documento simile trovato nel database.

Accesso micro-servizio annotazione

1. L'utente apre la dashboard di Heimdall che mostra il set di fascicoli a disposizione
2. L'utente clicca sull'icona di un fascicolo, accedendovi
3. L'utente visualizza tutti i documenti del fascicolo con delle icone disposte in una griglia con tre colonne
4. L'utente clicca il pulsante a forma di martello accanto all'icona del documento d'interesse oggetto di successiva annotazione con HeimTag
5. Il sistema chiama il micro-servizio HeimTag in una nuova pagina.

d. Funzionalità avanzate

Annotazione dei documenti

1. L'utente ha già utilizzato la funzionalità "Accesso micro-servizio annotazione"
2. L'utente visualizza una pagina dove poter scegliere lo schema di annotazione tra quelli gestiti da HeimTag
3. L'utente seleziona lo schema di annotazione e clicca sul pulsante di conferma
4. L'utente visualizza una schermata con divisa in due aree: (i) a sinistra, si mostrano i tag selezionabili e le loro proprietà, (ii) a destra, si mostra il contenuto testuale del documento
5. L'utente seleziona un tag dalla sinistra dell'interfaccia (può impostare le sue proprietà. Una volta selezionato il tag verranno visualizzate in basso a sinistra le possibili proprietà personalizzabili, es. "Nome del giudice" in un tag "giudice")

6. L'utente seleziona il testo (tenendo premuto il tasto sinistro del mouse) da voler annotare con il tag selezionato, nell'area destra dell'interfaccia.
7. L'utente rilascia il tasto sinistro del mouse, terminando la selezione, ed il sistema mostra il tag intorno al contenuto testuale annotato (in una sorta di overlay).
8. L'utente clicca sul tasto "Completato", una volta terminato il lavoro di annotazione.

Ricerca documenti simili (basata sul contenuto testuale)

1. L'utente apre la dashboard di Heimdall che mostra il set di fascicoli a disposizione
2. L'utente clicca sull'icona di un fascicolo, accedendovi
3. L'utente visualizza tutti i documenti del fascicolo con delle icone disposte in una griglia con tre colonne
4. L'utente clicca il pulsante a forma di cervello con l'acronimo "NLP" accanto all'icona del documento d'interesse (root document)
6. L'utente visualizza, a destra dell'interfaccia, una colonna con l'insieme dei documenti simili in termini di contenuto testuale. Sotto ciascuna delle icone che rappresentano i documenti, il sistema riporta un valore numerico reale (tra -1 e 1) che equivale alla similarità (espressa con la cosines similarity) tra il documento selezionato (root document) e il documento simile trovato nel database.

e. Gestione dei dati

Heimdall gestisce i dati attraverso un database orientato ai documenti, MongoDB. Le collezioni contenute nel database sono le seguenti. Per ciascuna collezione verranno mostrati i diversi campi - e relativo tipo - che la compongono.

Fascicolo	
fascicoloID	String
nome	String
dataCreazione	Date
ultimaMod	Date
partiCoinvolte	String[]
documentIDs	ArrayList<String>

Documento	
documentID	String
name	String
date	Date

dataUltimaModifica	Date
text	String [INDEXED]
djangoCaseFileId	String
mongoCaseFileId	String
documentObject	Bson
ContenutoBert	Double[]
idDocSenTag	Int
refsNorm	ArrayList<String>

Riferimento normativo

RiferimentoID;	String
NomeLegge;	String
url	String [INDEXED, UNIQUE]
type	String

É possibile effettuare operazioni CRUD accedendo direttamente alla base di dati (usando le istruzioni da linea di comando), oppure è possibile aggiungere, rimuovere, modificare dati attraverso le funzionalità messe a disposizione da Heimdall tramite GUI.

f. Personalizzazione delle impostazioni

L'aggiornamento di schemi di annotazione esistenti o importazione di nuovi schemi xsd deve essere svolto da esperti IT in quanto l'utente non può accedere al file di configurazione di uno schema pre-esistente.

Altre personalizzazioni non sono attualmente possibili.