

Epilexo Search: un'interfaccia di fruizione di lessici antichi per il progetto ItAnt

Rapporto Tecnico v.1.0

Riassunto:

Il documento descrive le caratteristiche tecniche e lo sviluppo dell'interfaccia web, EpiLexo Search, della piattaforma DigItAnt, con annesse informazioni relative al data-model e alle tecnologie utilizzate per la realizzazione della piattaforma.

Autori

Autore principale:

Michele Mallia, Istituto di Linguistica Computazionale "A. Zampolli", CNR, Pisa.

Co-autori: Andrea Bellandi, Michela Bandini, Valeria Quochi, Alessandro Tommasi, Cesare Zavattari

Supervisione: Valeria Quochi

Storia del documento

Versione	Autori	Data	Commenti
0.1	Michele Mallia	06/02/2023	Prima bozza
0.2	Michele Mallia	24/12/2023	1. Introduzione
0.3	Michele Mallia	27/12/2023	2. Background Context 3. Progetti di epigrafia digitale 4. L'applicazione web EpiLexO Search
0.5	Michele Mallia	28/12/2023	Sezioni 4.1, 4.2, 4.3
0.6	Michele Mallia	02/01/2024	Sezione 5
0.7	Michele Mallia	02/12/2024	GeoNames Model

Indice

1. Introduzione	4
2 Background context: il progetto ItAnt	4
2.1 Obiettivi del progetto	5

2.2 Architettura funzionale di EpiLexO Search	5
3 Progetti di epigrafia digitale: stato dell'arte	7
3.1 Epigraphic Database Heidelberg	7
3.2 EDEN—An Epigraphic Web Database of Ancient Inscriptions	9
4. L'applicazione web EpiLexO Search	11
4.1 Interfaccia di fruizione	11
4.2 Sezione dei Testi	12
4.2.1 Requisiti per la sezione dei testi (anteprima)	12
4.2.2 Requisiti per la sezione dei testi (dettaglio)	13
4.3 Sezione del lessico	16
4.3.1 Requisiti per la pagina del lessico (anteprima)	16
4.3.2 Requisiti per la pagina del lessico (dettaglio)	17
4.4 Sezione della Bibliografia	19
4.4.1 Requisiti per la sezione della bibliografia	20
4.5 Sezione delle Concordanze	21
4.6 Sezione della ricerca avanzata	21
4.6.1 Requisiti per la sezione della ricerca avanzata	22
5. Servizi del front-end	24
5.1 Servizi del lessico (LexO-Server)	24
5.1.1 Servizi per il recupero dei dati lessicali	24
Get Lexical Entries	24
Get Lexical Entry	25
Get Filtered Forms	26
Get Forms	26
Get Form	27
Get Filtered Senses	27
Get Senses	28
Get Sense	28
Get Etymologies	29
Get Etymology	29
5.1.2 Servizi per il recupero di dati statistici e di vocabolario	30
Get Lexical Entry Types	30
Get Authors	30
Get Languages	30
Get Pos	31
Get Status	31
5.1.3 LexO Data Model	32
StatisticsCounter	32
LexicalElement	32
Form	34
Sense	35
EtymologySet, EtyLink ed Etymology	35
5.2 Servizi dei testi (CASH)	37
5.2.1 Servizi per le attestazioni e la ricerca	37

Get Tokens	37
Get Content	38
Get Annotations	38
Search Files	39
CountFiles	39
GetUniqueMetadata	40
5.2.2 CASH Data Model	40
Token	40
XmlAndId	41
Annotation	41
AnnotationAttributes	42
Span	42
BodyTextPart	43
ConservationInstitution	43
ObjectDimension	43
DiscoveryYear	44
Facsimile	44
LanguageMetadata	45
License	45
PlaceModel	45
Settlement	45
TraditionalIDs	46
Trismegistos	46
SupportModel	46
TextMetadata	47
5.3 GeoData Services (Pleiades & Geonames)	49
5.3.1 Pleiades	49
GetPlace	50
5.3.2 Pleiades Data Model	50
PleiadesDataModel	50
5.3.3 GeoNames	50
GetPlace	51
5.3.4 GeoNames Data Model	51
GeoNamesDataModel	51
5.3.5 Global GeoData Model	52
GeoNamesDataModel	52
5.3.6 Minor Models	52
BboxModel	53
Markers	53
5.4 Servizi della bibliografia	53
GetAllBooks	53
5.4.1 Bibliography Data Model	54
Book	54
6. Bibliografia	55

1. Introduzione

Il medesimo documento descrive lo sviluppo di “EpiLexO Search”, una piattaforma web per la fruizione di dati relativi ai lessici elettronici collegati a materiali testuali e bibliografici, nonché a risorse lessicografiche esterne, creati attraverso la piattaforma di editing dei lessici “EpiLexO Editor”¹. L’obiettivo del progetto è quello di creare uno strumento che possa permettere agli utenti interessati a visualizzare tutte le informazioni peculiari relative ai dati delle iscrizioni codificate secondo gli standard dell’epigrafia digitale - in questo caso di Epidoc - e dei dati relativi al lessico computazionale e dei Linguistic Linked Open Data, tramite l’utilizzo del modello OntoLex-Lemon².

Oltre alla consultazione dei lessici, EpiLexO Search include un particolare sistema per la renderizzazione dei testi epigrafici secondo le convenzioni di Leida con annessi dati LOD e dei sistemi per la reperibilità di dati da più fonti per consentire agli utenti di effettuare ricerche incrociate e avanzate.

Questo documento esplora le varie fasi della progettazione della piattaforma, dettagliando le specifiche tecniche di ciascun componente e delineando il modello generale per l’organizzazione e la presentazione dei dati.

2 Background context: il progetto ItAnt

L’applicazione web EpiLexo Search è sviluppata nel contesto del progetto ItAnt ovvero “Lingue e culture dell’Italia antica: linguistica storica e modelli digitali” (*finanziato dal MUR - PRIN 2017 - 2017XJLE8J). Tra i suoi obiettivi primari, oltre alla realizzazione di una infrastruttura informatica per la creazione e gestione di materiali linguistici per lingue e arcaiche, c’è anche la necessità di presentare i dati raccolti e gestiti tramite l’interfaccia di editing attraverso un’interfaccia che consenta di visualizzare i dati e combinarli da fonti diverse.

2.1 Obiettivi del progetto

L’obiettivo principale del progetto ItAnt è quello di offrire ai linguisti storici e studiosi di culture antiche uno strumento elettronico per la descrizione e/o lo studio del lessico di lingue arcaiche anche di frammentaria attestazione a partire da e in collegamento con i testi che ne

¹ Mallia, Michele; et al., 2023, EpiLexO, ILC-CNR for CLARIN-IT repository hosted at Institute for Computational Linguistics "A. Zampolli", National Research Council, in Pisa, <http://hdl.handle.net/20.500.11752/ILC-1005>.

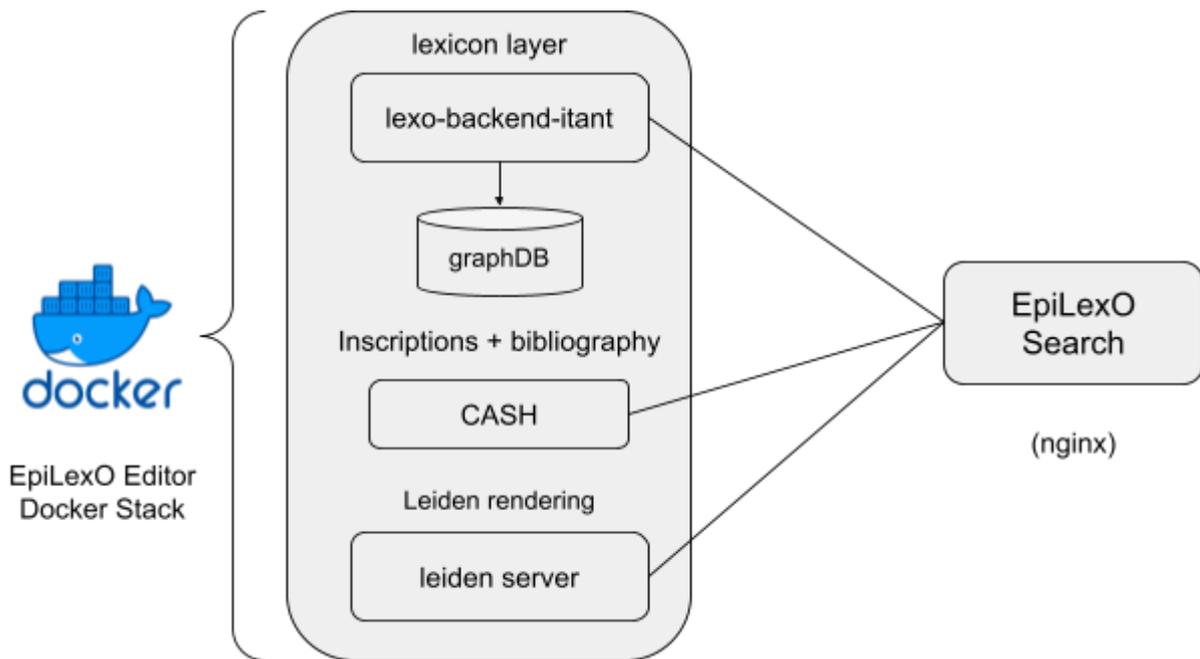
² McCrae, John P., et al. "The Ontolex-Lemon model: development and applications." *Proceedings of eLex 2017 conference*. 2017.

costituiscono le attestazioni /i testimoni e la bibliografia della letteratura secondaria di riferimento. L'idea di base è quella di creare uno strumento che consenta agli studiosi di poter fruire dei lessici elettronici di lingue antiche e arcaiche basati e collegati ai relativi testimoni al fine di facilitare la condivisione e la collaborazione (tra studenti, professori e centri di ricerca) anche a scopi didattici e di divulgazione. Da un punto di vista prettamente tecnico al momento non sembrano esistere soluzioni software web disponibili che permettano di poter effettuare una renderizzazione di dati relativi ai lessici computazionali (in LOD) e dei dati relativi alle iscrizioni e ai testi annotati (codificati in TEI-EpiDoc XML).

Un obiettivo tecnico di questa piattaforma consiste dunque nella possibilità di realizzare in modalità user-friendly lessici compatibili e pubblicabili come Linguistic Linked Open Data, ovvero basati su ontologie, modelli e vocabolari condivisi nella comunità del web semantico di modo che questi stessi dati possano essere riutilizzati all'interno della comunità scientifica, magari con scopi diversi.

2.2 Architettura funzionale di EpiLexO Search

A differenza del precedente applicativo menzionato, EpiLexO Editor, questa applicazione web non appartiene ad uno stack Docker ben specifico. L'architettura di questa applicazione risulta essere quindi composta da un solo componente, ovvero l'interfaccia stessa, che però è collegata ai servizi presenti nello stack dell'applicazione per l'editing dei lessici e delle iscrizioni precedentemente costruito. L'applicazione EpiLexO Search non prevede inoltre dei meccanismi di autorizzazione e di autenticazione come per l'applicazione di editing, ma prevede un accesso totalmente libero per qualsiasi tipologia di utente, trattandosi in effetti di uno strumento per mostrare al "pubblico" i risultati prodotti mediante l'applicazione EpiLexO Editor.



3 Progetti di epigrafia digitale: stato dell'arte

A partire dai primi anni del 2000 si iniziano a intravedere i primi progetti legati alla conservazione e alla consultazione di iscrizioni antiche; tra i maggiori portali di epigrafia digitale tenuti sotto osservazione per lo sviluppo di questo progetto e che costituiscono punti di riferimento importanti di questa disciplina ritroviamo il "Database Epigrafico di Heidelberg", i "database epigrafici di Roma e Bari e il portale del progetto Papyri. In questi progetti è possibile consultare / esplorare corpus di risorse attraverso una vasta gamma di filtri e metodi di ricerca che consentono anche ad un utente non esperto di poter ricercare la risorsa desiderata, anche attraverso l'ausilio di mappe ed altri elementi interattivi.

Oltre ad avvalersi in maniera sempre più consistente delle tecnologie digitali per poter affrontare lo studio delle iscrizioni in maniera innovativa, potendo condividere meglio i risultati all'interno della comunità accademica e al di fuori di essa, il mondo dell'epigrafia digitale ha visto anche nascere un'integrazione tra queste tecnologie e quelle relative alla gestione del sapere e della conoscenza, orientandosi in un paradigma del web affermatosi nell'ultimo decennio come uno standard: quello dei linked open data (LOD) e del web-semantic.

Di seguito vengono illustrati due dei progetti più salienti di epigrafia digitale che stabiliscono lo stato dell'arte della disciplina e/o costituiscono un riferimento nel settore dell'epigrafia digitale, e che hanno ispirato lo sviluppo della piattaforma EpiLexO Search, descrivendone le caratteristiche peculiari.

3.1 Epigraphic Database Heidelberg

"Epigraphic Database Heidelberg" (EDH) è un progetto iniziato nel 1986 con l'obiettivo di digitalizzare e rendere accessibili online tutte le iscrizioni latine dell'antichità. La base dati è stata sviluppata con l'intento di servire sia la comunità accademica che il pubblico più vasto interessato all'epigrafia romana.³ Il progetto EDH, diventato parte integrante dell'Accademia delle Scienze e delle Lettere di Heidelberg nel 1993, si distingue per il suo approccio scientifico alla raccolta e verifica delle iscrizioni. Questo approccio include l'utilizzo di disegni, fotografie e, ove possibile, autopsie delle iscrizioni, garantendo un alto livello di accuratezza e affidabilità delle informazioni raccolte. Il database è diventato accessibile online nel 1997, fornendo agli utenti la possibilità di consultare trascrizioni, record bibliografici e metadati relativi alle immagini delle iscrizioni.

La struttura dei dati nell'EDH è composta da quattro database principali: il database dei testi, delle immagini, dei riferimenti bibliografici e dei dati geografici. Il database dei testi, in particolare, è estremamente dettagliato, contenendo oltre 1,4 milioni di campi metadati. Il database delle immagini, con circa 38.700 immagini, è una risorsa preziosa per gli studiosi che cercano di collegare testi e contesti visivi delle iscrizioni. Inoltre, sono in fase di sviluppo ulteriori database, come quello delle persone menzionate nelle iscrizioni, per arricchire ulteriormente la risorsa.

Dal punto di vista tecnico, l'EDH ha visto una notevole evoluzione dalla sua creazione. Originariamente basata su un database Oracle offline, la piattaforma è stata successivamente trasformata in un'applicazione web complessa con un server di indice di ricerca dedicato. Per gestire in modo efficiente i dati e le immagini, l'EDH utilizza vari server e database.

L'integrazione dei LOD in questo progetto segna un punto di svolta nella modalità di condivisione e interconnessione delle informazioni epigrafiche. La EDH, con la sua vasta collezione di iscrizioni latine, si posiziona all'avanguardia nell'adozione di questa tecnologia che trasforma il modo in cui i dati storici e archeologici possono essere esplorati e analizzati. L'implementazione dei LOD nell'EDH va oltre il semplice accumulo di dati; rappresenta un impegno verso una maggiore interconnessione delle conoscenze. Utilizzando gli identificatori unici (URI), l'EDH stabilisce collegamenti tra le sue entità e una vasta gamma di risorse digitali correlate. Questo approccio non solo migliora la navigabilità all'interno del database stesso, ma crea anche un tessuto di collegamenti esterni che arricchiscono il contesto di ogni iscrizione. Gli studiosi possono così navigare facilmente da un'iscrizione latina a informazioni correlate in altri database, collegando le iscrizioni a contesti storici, geografici e culturali più ampi. L'uso dei LOD trasforma la EDH in un ecosistema digitale

³ Feraudi-Gruénais, Francisca. "Epigraphic Database Heidelberg (EDH)-Or: the challenge of getting all ancient Latin inscriptions outside Rome available" by one click". (2018).

dinamico, in cui le iscrizioni non sono più viste come entità isolate, ma come parti di una rete di conoscenza più ampia. Questa integrazione permette agli studiosi di realizzare analisi transdisciplinari, unendo i dati epigrafici con altre fonti storiche per ottenere una comprensione più ricca e sfaccettata del passato. I LOD aprono nuove prospettive per la ricerca in campi come la storia, l'archeologia e le Digital Humanities, offrendo nuove strade per esplorare le connessioni tra i dati.

3.2 EDEN—An Epigraphic Web Database of Ancient Inscriptions

Il progetto EDEN (Epigraphic Database Erlangen-Nürnberg) è una banca dati online in continua evoluzione che raccoglie iscrizioni greche antiche provenienti dalle città dell'odierna Turchia occidentale. La creazione di EDEN, iniziata nel 2012 all'Università di Erlangen-Nürnberg, rappresenta un'unione sinergica tra storici antichi, archeologi e informatici, che lavorano insieme per mantenere e sviluppare sia il contenuto che il quadro tecnico della piattaforma.⁴

EDEN si avvale del sistema WissKI, un ambiente di ricerca virtuale sviluppato e mantenuto da un gruppo di ricerca in Digital Humanities presso la stessa università, il Museo Nazionale Germanico di Norimberga e il Museo di Ricerca Zoologica Alexander Koenig di Bonn. WissKI, basato sul sistema di gestione dei contenuti Drupal, offre un'interfaccia web unica per l'acquisizione, la modifica e la presentazione dei dati, sfruttando tecniche del web semantico per arricchire i dati con significati aggiuntivi. L'uso di RDF (Resource Description Framework) e OWL (Web Ontology Language) all'interno di WissKI permette a EDEN di strutturare i dati in modo che possano essere facilmente interrogati e collegati non solo all'interno del database stesso, ma anche con risorse esterne. Questo significa che le iscrizioni possono essere collegate a dati storici, geografici e archeologici da altre fonti, arricchendo la comprensione delle iscrizioni e fornendo agli studiosi un contesto più ampio per la loro interpretazione.

Il progetto EDEN, all'epoca della stesura del documento, conteneva circa 600 iscrizioni greche e latine del periodo ellenistico e romano imperiale, arricchite con metadata. Il database riveste una grande importanza sia per gli storici antichi, focalizzati sul testo delle iscrizioni, sia per gli archeologi, interessati agli aspetti fisici dei supporti delle iscrizioni come pietre e monete. Il layout di EDEN è ispirato a Wikipedia, con testo libero al centro e dati tabellari e immagini sul lato destro, permettendo agli utenti di navigare facilmente tra i record. Ogni segmento di testo e metadata tabellare può essere riutilizzato in altri record, creando una rete di dati interconnessi.

⁴ Scholz, Martin, et al. "EDEN: an epigraphic Web Database of ancient inscriptions." *Proceedings of CIDOC 2014*. ICOM/CIDOC, 2014. 1-7.

Un aspetto rivoluzionario di EDEN è l'integrazione di modelli 3D degli siti di scavo, che permette agli utenti di esplorare virtualmente i siti archeologici. Questo è stato possibile grazie all'adozione di WebGL, una tecnologia che consente la resa efficiente di contenuti tridimensionali sulle pagine web. Tuttavia, ci sono state sfide nel ridurre le dimensioni dei file dei modelli 3D per renderli gestibili sul web, mantenendo allo stesso tempo un livello di dettaglio sufficiente per visualizzare le iscrizioni.

4. L'applicazione web EpiLexO Search

In questa sezione si illustrano le caratteristiche relative all'applicazione EpiLexO Search; vengono chiariti gli obiettivi dell'applicazione stessa e i suoi requisiti funzionali al fine di offrire una panoramica quanto più completa al lettore di questa relazione tecnica.

La realizzazione di EpiLexO Search prende spunto da altri progetti di fruizione di dati epigrafici, tra i quali vale la pena menzionare lo stesso EDH citato precedentemente e CGRN (Collection of Greek Ritual Norms)⁵.

Per questo progetto è stato utilizzato Angular 14 per la realizzazione della single page application; rispetto all'applicazione EpiLexO Editor, però, è stata utilizzata un'altra libreria grafica, ovvero PrimeNG⁶ (nella sua versione v13), che contiene numerosi componenti e strumenti per la gestione grafica dell'interfaccia.

Rispetto all'applicazione EpiLexO Editor, trattandosi di uno strumento di fruizione, non risulta particolarmente complessa a livello di componentistica e di struttura, ma vi sono state numerose difficoltà nella realizzazione di componenti custom e per l'implementazione di sistemi di ricerca incrociata e di visualizzazione di dati geografici. Nel corso della sezione verranno trattati tutti gli argomenti nel dettaglio.

4.1 Interfaccia di fruizione

L'interfaccia di fruizione si presenta suddivisa in quattro sotto-sezioni. Sono presenti infatti le sezioni relative a:

- Testi e iscrizioni: sono presenti tutte le informazioni relative alle iscrizioni con annesse informazioni geografiche, resa del testo epigrafico secondo le convenzioni di Leida e la relativa bibliografia del testo; una prima parte di questa sezione si occupa di mostrare tutti i testi presenti nella collezione e di effettuare delle ricerche semplici all'interno della repository, secondo parametri arbitrati e categorie.

⁵ Carbon, J.-M., & Pirenne-Delforge, V. (December 2019). Two Notes on the Collection of Greek Ritual Norms: Looking Back, Looking Forward. Axon : Greek historical inscriptions, 3 (2). doi:10.30687/Axon/2532-6848/2019/02/007

⁶ <https://primeng.org/>

- Lessici: in questa sezione è presente un'alberatura contenente tutti i lemmi e le relative forme dove, cliccando su una di esse, si apre una vista con i dettagli dell'elemento lessicale selezionato; è presente anche una sezione dove poter effettuare una ricerca delle voci lessicali mediante alcune categorie
- Bibliografia: sono presenti tutti i materiali bibliografici legati alle iscrizioni e ai lessici inseriti dagli /dalle studiosi/e tramite la piattaforma di editing; anche qui è presente una sezione dove poter effettuare una ricerca semplice mescolando anche dei dati relativi alle iscrizioni (simple cross query search)
- Concordanze: in questa sezione è presente una tabella di corrispondenze tra le sigle dei testi usate in ItAnt e le altre sigle utilizzate da altri sistemi di riferimento (eg: Trismegistos) e altre sigle lasciate dagli/dalle studiosi/e
- Advanced Search: questa è una delle sezioni più importanti, dove vengono effettuate le query complesse mescolando i vari dati provenienti da fonti diverse (LexO e CASH) per ottenere le iscrizioni corrispondenti ai parametri richiesti

4.2 Sezione dei Testi

La sezione dei testi si presenta con una schermata principale dove sono mostrati tutti i testi che sono presenti in CASH e a sinistra una barra laterale con due sezioni relative a: categorie e simple search.

Le categorie con cui è possibile effettuare le ricerche sono quattro: All (tutti i materiali epigrafici, indistintamente), per Date, Luoghi e Tipologia delle iscrizioni.

4.2.1 Requisiti per la sezione dei testi (anteprima)

Qui vengono elencati i requisiti per la sezione relativa ai testi:

- Lazy loading dei dati: una volta che l'utente entra in questa sezione, viene effettuato un fetching dei dati delle iscrizioni, effettuando una richiesta di una porzione di dati da CASH; si è notato che effettuando il caricamento totale delle iscrizioni, si andava incontro ad un rallentamento generale della piattaforma per via della chiamata "onerosa" al sistema di gestione delle iscrizioni
- Raggruppamento: i dati, per essere categorizzati e per poter effettuare le chiamate tramite le "faccette", necessitano di essere raggruppati; questo avviene mediante un API che consente all'interfaccia di ottenere un set di informazioni relative ad un metadato specifico; questo sistema permette anche di organizzare meglio il form con cui viene effettuata la Simple Search.

- Ricerca semplice: in questa sezione, la ricerca prevede soltanto un solo endpoint di riferimento, ovvero CASH, nel senso che i criteri di ricerca del form prevedono di ricercare soltanto all'interno del repository delle iscrizioni (eg: chiedimi tutte le iscrizioni che hanno titolo X, materiale di tipo Y, etc....)

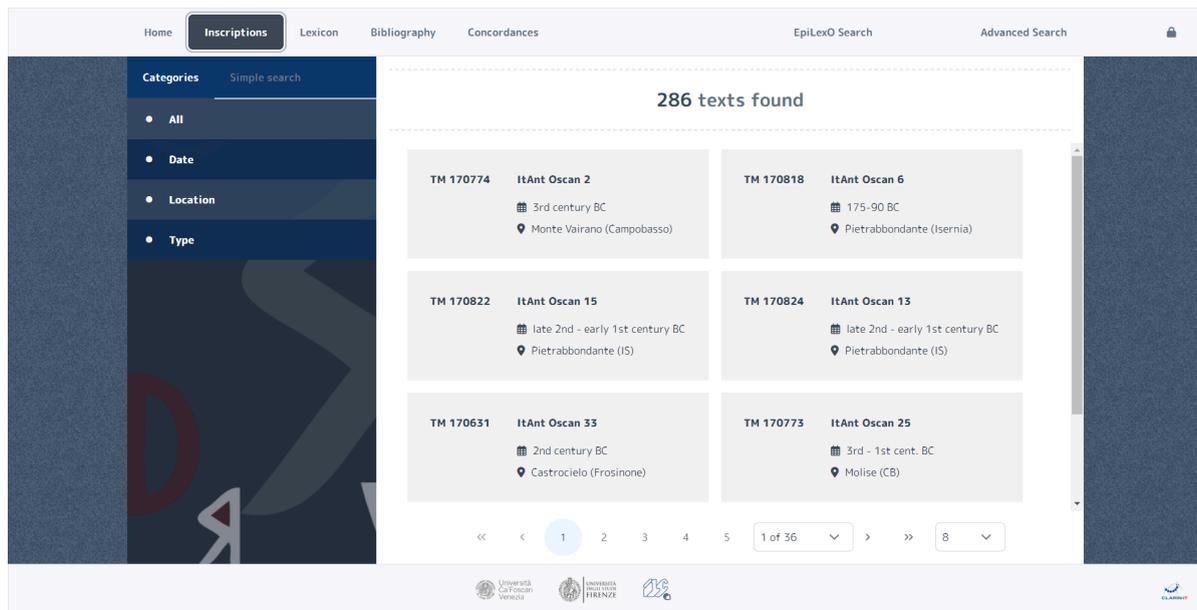


Fig. 1: sezione relativa ai testi

4.2.2 Requisiti per la sezione dei testi (dettaglio)

Per quanto riguarda il dettaglio delle iscrizioni, è necessario fare un'analisi dettagliata dei passi effettuati dalla piattaforma per il recupero dati.

Una volta che l'utente seleziona un'iscrizione, si scatena una serie di eventi che si può riassumere in questo schema:

- Tramite il sistema di controllo di routing viene elaborata la stringa che corrisponde alla sigla del file da analizzare; con questa, viene richiamato un servizio di CASH che ci restituisce tutti i suoi metadati; viene triggerato il sistema per il recupero dei dati geografici della singola locazione presente tra i metadati del file;
- Dopo aver ottenuto i metadati, viene richiamato un altro servizio che, tramite l'identificativo del file (campo element-id), ci restituisce il contenuto XML del file, che servirà per ottenere i dati relativi al contenuto della trascrizione interpretativa dell'iscrizione, bibliografia, immagini, contenuto dell'apparato e traduzione;
- Parallelamente, vengono effettuate le operazioni di conversione del contenuto epigrafico del file XML secondo le convenzioni di Leida; in questa parte viene preso il

contenuto XML del file, analizzato e, per ciascun nodo TEI presente, gli vengono assegnate le proprietà relative dei dati lessicali legati a quel nodo⁷; una volta recuperate le informazioni sugli identificativi e lo span di caratteri relativo per ciascun token, viene effettuata una ricerca delle annotazioni presenti per ciascun token; se ci sono delle informazioni lessicali e delle forme associate ad un token, queste vengono inserite all'interno dell'elemento HTML relativo alla parola annotata nell'editor; viene anche creato un componente dinamico che consentirà all'utente di visualizzare le informazioni lessicali cliccando sulla parola annotata mediante un click.

Una volta effettuate queste operazioni in parallelo, tutti i dati recuperati verranno presi e renderizzati opportunamente per ciascuna sezione. Tutte le operazioni asincrone vengono performate dalla piattaforma mediante l'utilizzo della libreria javascript RxJs che consente di effettuare queste operazioni in maniera controllata ed efficiente.

Tramite questa vista è possibile inoltre visualizzare l'XML sorgente del file caricato su CASH, visualizzare o nascondere le ricostruzioni del testo e interagire interattivamente con la resa della versione interpretativa attraverso una corrispondenza tra gli identificativi dei nodi XML presenti nel commentario e il testo renderizzato.

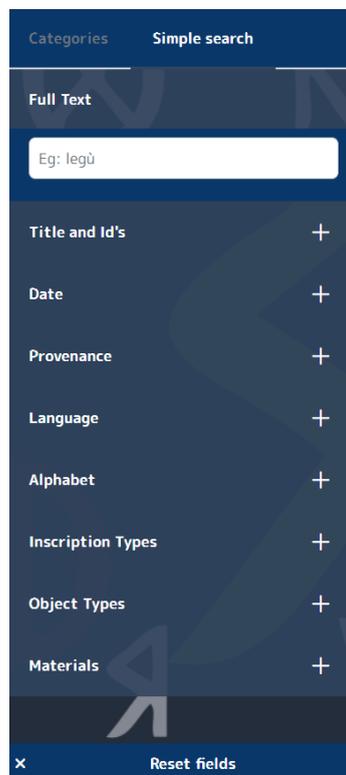


Fig. 2: sezione del simple search per i testi

⁷ Vengono effettuate delle comparazione tra i nodi presenti nell'XML e i token presenti nel testo, effettuando un matching tra gli identificativi dell'XML e gli identificativi presenti nei token estratti da CASH.

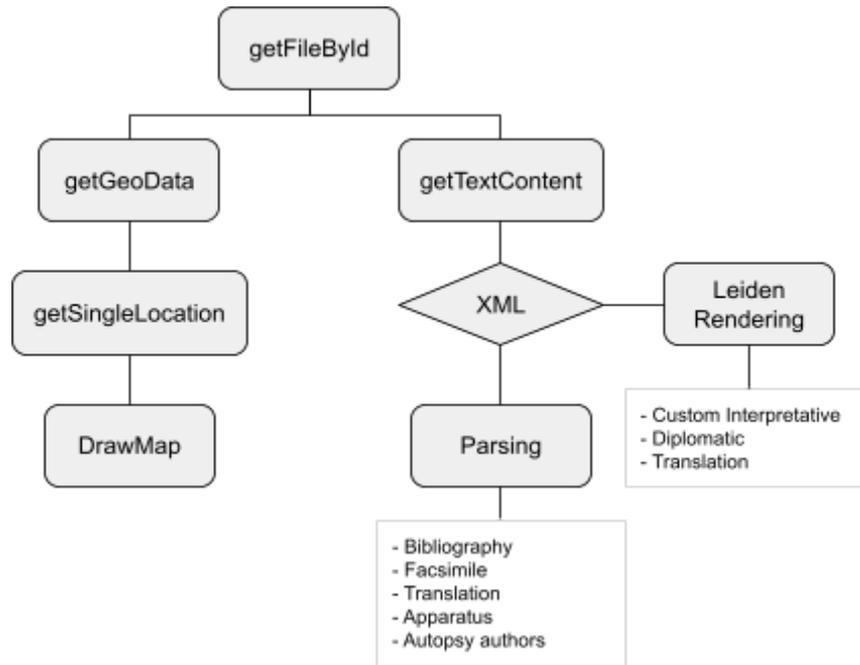


Fig. 3: workflow per il recupero di informazioni di un testo

Home Inscriptions Lexicon Bibliography Concordances EpiLexO Search Advanced Search

ItAnt Oscan 2
Curse tablet from Monte Vairano

Text

Intepretative

Face_a	Face_b
1 pakis : heleviis . tre(bieis)	1 statis : betitis : [---]

Translation

Pacio Helevio figlio di di ... ius | Status Betitius [son

Support

Object typ

Material: **lead**

Dimensions: **Width:** 1,5 cm
Height: 6,7 cm

Layout notes: Inscribed on both sides, with a line on each.

Palaeographic notes: Words are separated by double dots; the second word-break on face A is a single

trebis@osc
http://lexica/mylexicon#LexO_2023-04-1710_19_29_081

Type: lexicalForm

Morphology:
partOfSpeech properNoun
case genitiveCase
gender masculine
number singular

Fig. 4: vista del dettaglio di un testo

4.3 Sezione del lessico

La sezione del lessico si presenta con una schermata principale dove sono mostrati tutti gli elementi lessicali presenti in LexO; a sinistra c'è sempre una barra laterale con le due sezioni relative alle categorie con cui effettuare le ricerche per tipologia (all, pos e lingua) e simple search. In quest'ultima, non è presente un form dove effettuare le ricerche, ma un albero dove sono elencate tutte le entrate lessicali (eccetto le radici del proto-indoeuropeo e proto-italico). Espandendo le entrate lessicale, verranno mostrate, ove presenti, le forme appartenenti all'entrata. Cliccando su ogni elemento, l'utente verrà portato nella pagina di dettaglio.

4.3.1 Requisiti per la pagina del lessico (anteprima)

Qui vengono elencati i requisiti per l'interfaccia di anteprima del lessico:

- Lazy loading dei dati: una volta che l'utente entra in questa sezione, viene effettuato un fetching dei dati del lessico, effettuando una richiesta di una porzione di dati da LexO; rispetto al recupero dei dati da CASH, qui le operazioni sono molto più dinamiche e veloci, ma per questioni di resa grafica e scelta stilistica, si è deciso di continuare con l'approccio della paginazione.
- Raggruppamento: il lessico inizialmente viene raggruppato in ordine alfabetico; per fare ciò è necessario scaricare tutto il lessico presente su LexO ed effettuare in maniera programmatica un raggruppamento per lettere dell'alfabeto; per le categorie Part of Speech e Language viene effettuata la medesima operazione, ma puntando a metadati diversi.
- Alberatura: i dati prelevati dal repository semantico andranno a popolare l'albero (componente simile a quello già presente nella piattaforma di editing), potendo così rendere chiara la struttura gerarchica che costituisce il lessico computazionale.

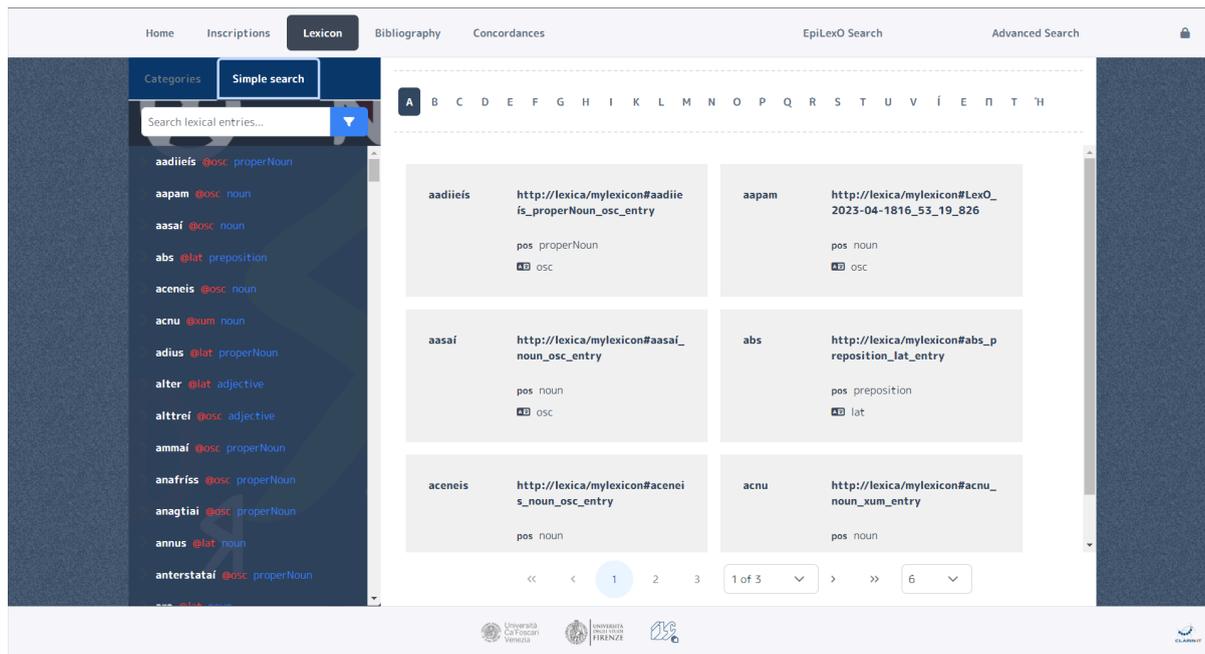


Fig. 5: sezione del lessico (anteprima), a sinistra c'è l'albero delle entrate lessicali

4.3.2 Requisiti per la pagina del lessico (dettaglio)

Per quanto riguarda il dettaglio di un elemento lessicale, è necessario fare un'analisi dettagliata dei passi effettuati dalla piattaforma per il recupero dati.

Una volta che l'utente seleziona un'entrata lessicale, si scatenano una serie di eventi che si può riassumere in questo schema:

- Tramite il sistema di controllo di routing viene elaborata la stringa che corrisponde all'IRI dell'entrata lessicale; tramite questa, vengono richiamati alcuni servizi per recuperare informazioni sui sensi, etimologie, cognates, forme e bibliografia relative all'entrata lessicale; successivamente viene chiamata l'API per ottenere i dettagli dell'entrata lessicale selezionata.
- Parallelamente viene chiamata l'API per il recupero delle attestazioni che hanno forme provenienti da questa entrata lessicale.

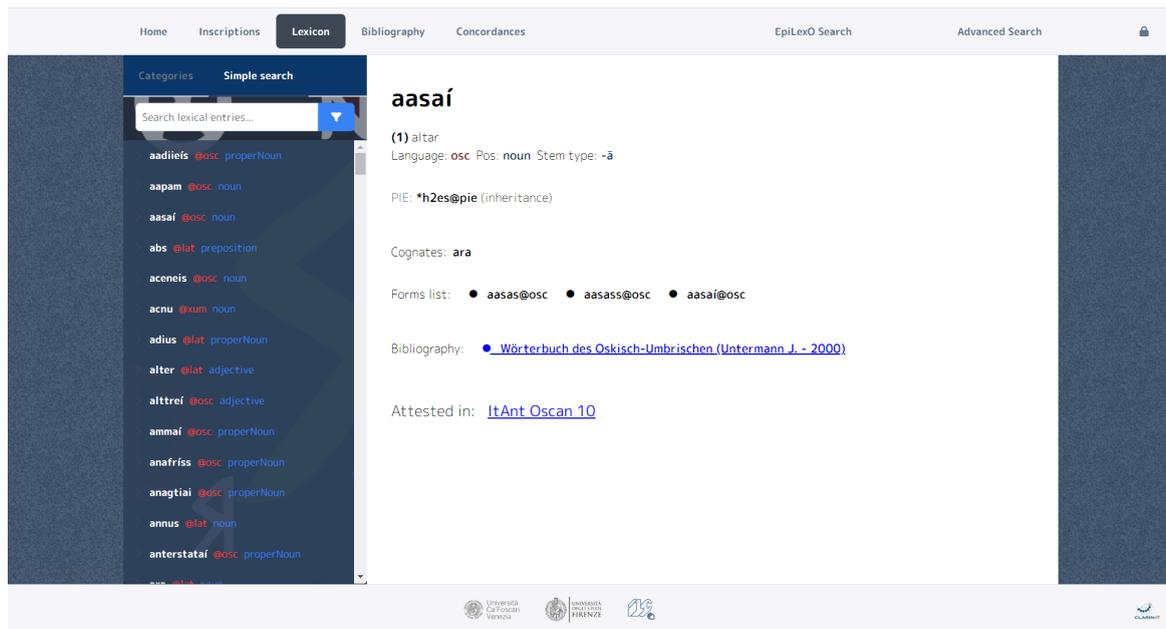


Fig. 6: vista in dettaglio di un'entrata lessicale

Se invece l'utente seleziona una forma vengono performatte queste operazioni:

- Viene richiamata la catena di eventi che serve per ottenere le informazioni sull'entrata lessicale a cui appartiene la forma (i sensi, le etimologie e le informazioni minori dell'entrata lessicale non verranno renderizzate, ma soltanto alcune informazioni basilari come la label).
- Parallelamente, vengono chiamate le API per recuperare le informazioni più dettagliate sulla forma selezionata.
- Allo stesso tempo, viene richiamata l'API per estrarre una lista di attestazioni in cui la forma è associata.

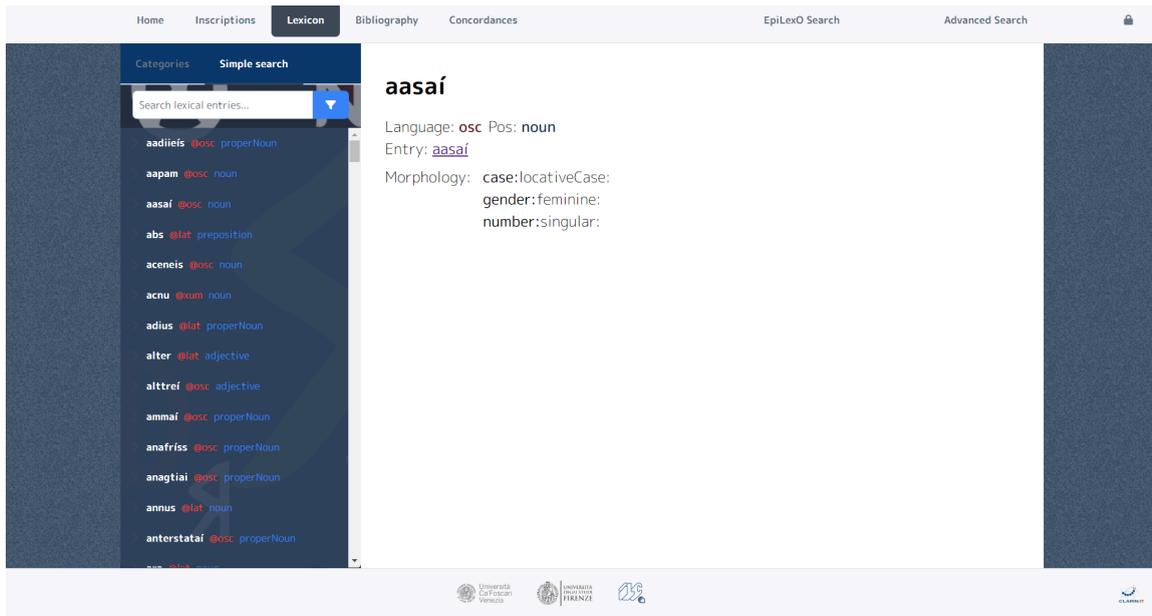


Fig. 7: vista in dettaglio di una forma

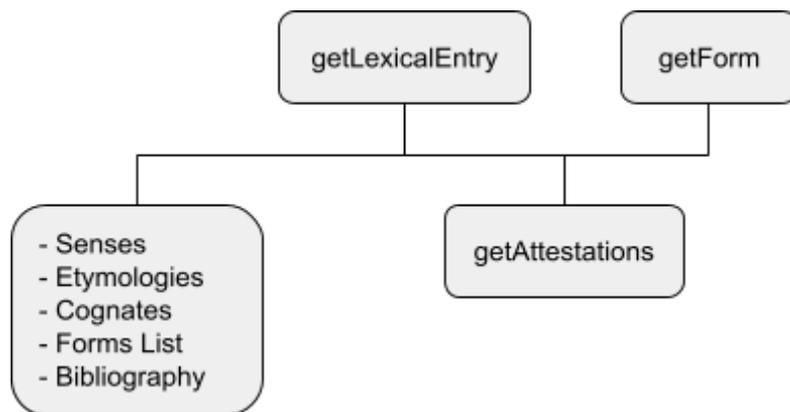


Fig. 8: workflow per il recupero di informazioni di un testo

4.4 Sezione della Bibliografia

La sezione della bibliografia si presenta con una schermata principale dove sono mostrati tutti gli elementi bibliografici presenti nel repository di Zotero (e importati su CASH). Gli elementi sono suddivisi e raggruppati in ordine alfabetico in base al cognome dell'autore. Anche in questa sezione è possibile effettuare una ricerca semplice che ci consente di andare a ricercare elementi bibliografici prendendo come riferimento anche le informazioni sulle iscrizioni (eg: dammi tutti gli elementi bibliografici presenti nell'iscrizione X).

Contrariamente alle altre sezioni, non c'è una vista per il dettaglio della bibliografia, in quanto, una volta cliccato un elemento bibliografico, viene aperto un link diretto al repository bibliografico (Zotero).

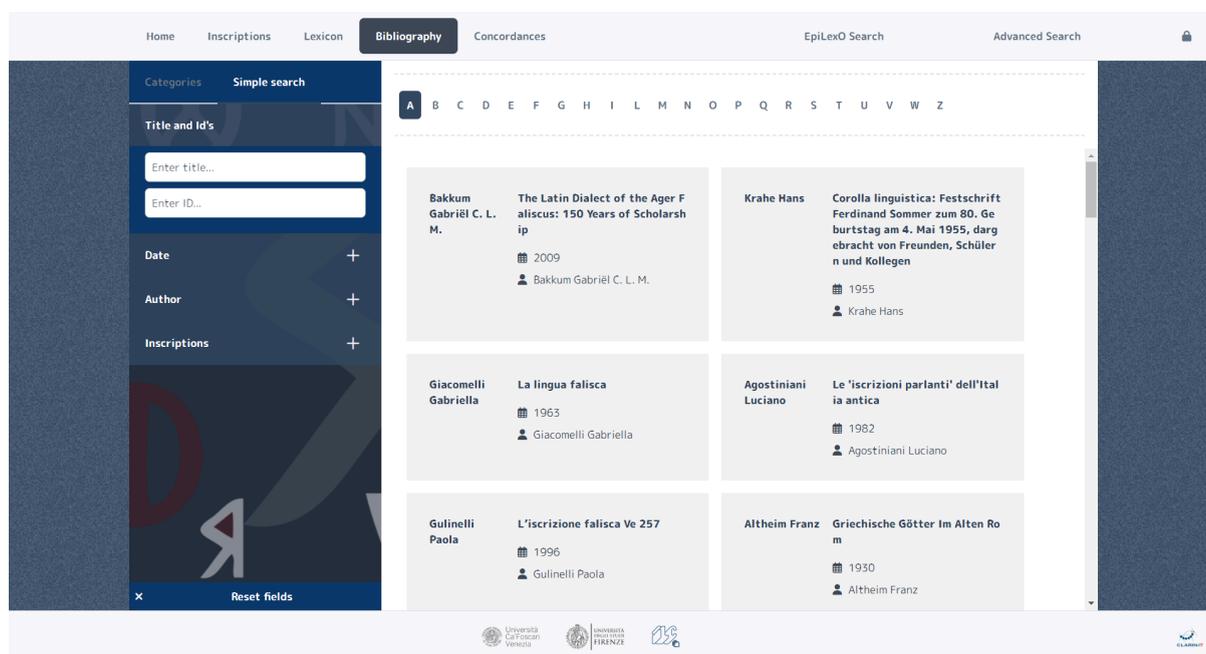


Fig. 9: vista della sezione della bibliografia

4.4.1 Requisiti per la sezione della bibliografia

La sezione degli elementi bibliografici non differisce molto rispetto alle altre sezioni; tra i vari requisiti che richiede questa sezione vi sono il fetching dei dati da CASH per ottenere la lista di tutti gli elementi bibliografici, il raggruppamento degli elementi in base all'ordinamento alfabetico e il sistema di ricerca attraverso la composizione di query in CQL⁸ in base ai parametri impostati nel form di ricerca.

4.5 Sezione delle Concordanze

In questa vista sono presenti, attraverso una tabella, tutti i riferimenti delle iscrizioni presenti nel repository delle iscrizioni e le diverse nomenclature in base alle convenzioni. Nella tabella, sono presenti quattro colonne: identificativo dell'epigrafe nel progetto ItAnt, titolo dell'epigrafe, identificativo Trismegistos⁹ e altri identificativi (con relativi riferimenti a Zotero).

⁸ Stesso sistema utilizzato per quanto riguarda il sistema di ricerca semplice per le iscrizioni. Non si usa questo sistema per la ricerca degli elementi lessicali.

⁹ <https://www.trismegistos.org/>

Attraverso il componente table di PrimeNG è stato possibile realizzare una vista tabellare con la possibilità di effettuare una serie di operazioni di filtering e ordinamento dei risultati in modo da consentire agli studiosi di poter trovare subito le iscrizioni in base a diversi criteri di ricerca.

ItAnt ID ↑↓	Title ↑↓	Trismegistos ID ↑↓	Other IDs
ItAnt Oscan 2	Curse tablet from Monte Vairano	TM 170774	ST Sa 36 Imlit Bouianum 98 Murano 6
ItAnt Oscan 6	Inscription ItAnt Oscan 6	TM 170818	ST Sa 5 Imlit Teruentum 11 Ve 153
ItAnt Oscan 15	Basin from Pietrabbondante	TM 170822	ST Sa 11 Imlit Teruentum 15
ItAnt Oscan 13	Inscription ItAnt Oscan 13	TM 170824	ST Sa 12 Imlit Teruentum 17
ItAnt Oscan 33	Inscription ItAnt Oscan 33	TM 170631	ST Sa 61 Imlit Aquinum 2
ItAnt Oscan 25	Inscription ItAnt Oscan 25	TM 170773	ST Sa 25 Imlit Bouianum 97 Ve 156

Fig. 10: sezione delle concordanze

4.6 Sezione della ricerca avanzata

La sezione delle ricerche avanzate (Advanced Search) è la vista tra le più importanti della piattaforma, in quanto è quella che consente all'utente di effettuare delle ricerche incrociate attingendo da più fonti di dati. La vista si presenta con una lista di iscrizioni in anteprima nella sezione centrale della vista e, lateralmente, c'è una barra con tre sezioni corrispondenti alle tre fonti di dati: inscription dataset (CASH), lexicon dataset (LexO) e bibliography dataset (CASH + Zotero).

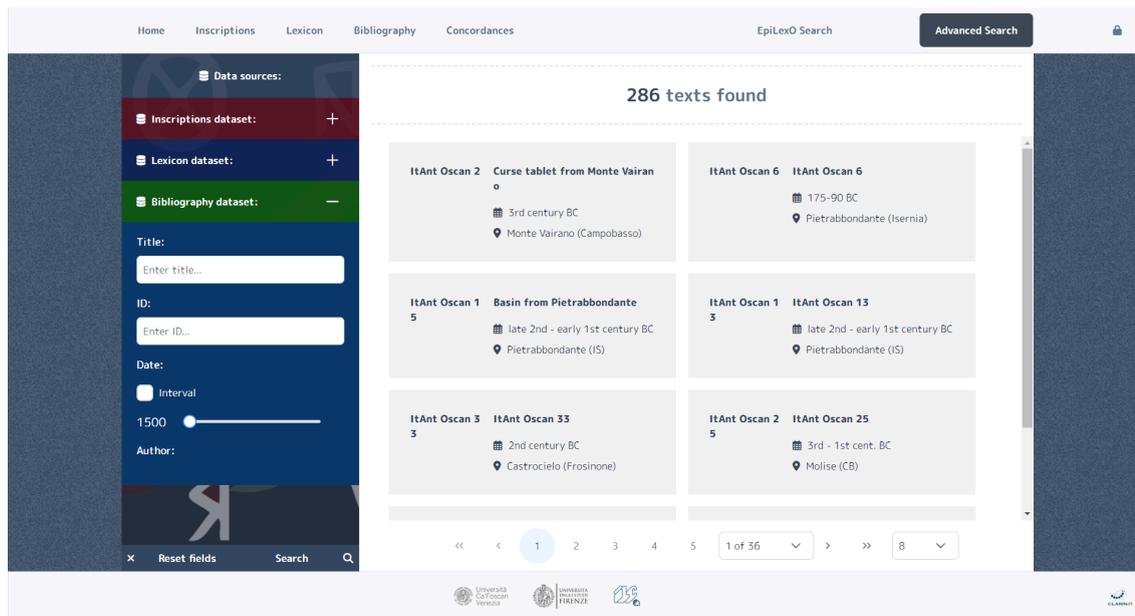


Fig. 11: vista della *advanced search*

Il form nella barra laterale è un'unione di tutti e tre i requisiti di ricerca per le sezioni dei lessici, iscrizioni e bibliografia, con qualche parametro in più in ciascuna di esse per poter effettuare delle query più precise e approfondite.

4.6.1 Requisiti per la sezione della ricerca avanzata

In questa sezione ci sono diversi requisiti funzionali che devono essere soddisfatti:

- Fetching dei dati da CASH: inizialmente, la sezione è popolata da elementi epigrafici caricati tramite la piattaforma di editing;
- Ricerca complessa: questa funzionalità avviene tramite un sistema decisionale che prevede la richiesta di dati da più fonti e di concatenare i risultati in una stringa CQL da passare a CASH per ricercare le iscrizioni; la sintassi è quella relativa al Corpus Query Language con qualche modifica per migliorare le prestazioni del server nel poter restituire i dati;
- Lazy loading: i dati restituiti in maniera parziale e saranno paginati in modo da ottimizzare le performance del sistema.

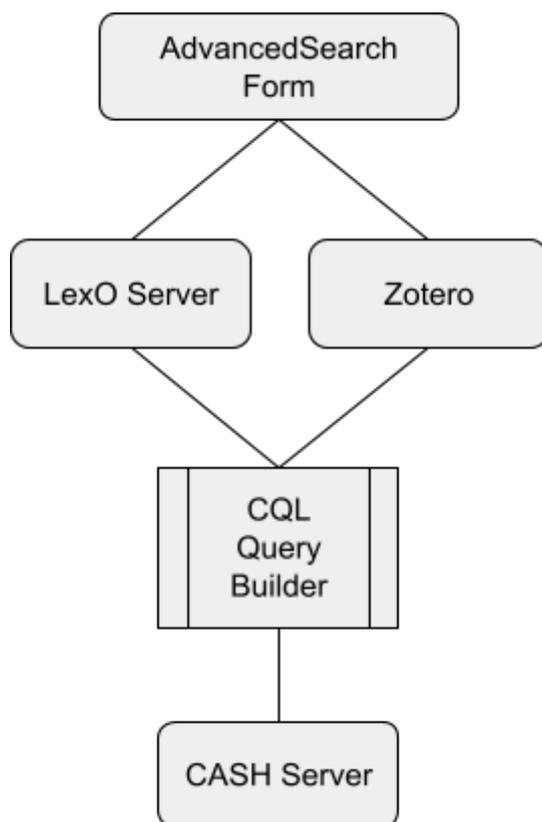


Fig. 12: workflow dell'Advanced search

5. Servizi del front-end

In questa sezione vengono elencati i servizi utilizzati dalla piattaforma per la fruizione di testi, lessici e apparato bibliografico. Il paragrafo è diviso in ulteriori sotto-sezioni relative agli ambiti di competenza dei servizi, i quali sono elencati in forma tabellare con annessi dati e modelli di chiamata.

5.1 Servizi del lessico (LexO-Server)

I servizi relativi al lessico sono erogati dal servizio LexO-Server, il quale comunica con GraphDB per la gestione del lessico computazionale sotto-forma di collezione di triple nel formato RDF. A differenza della piattaforma di editing, qua vengono utilizzati soltanto i servizi per quanto riguarda il recupero delle informazioni, in quanto nessuna azione di modifica viene effettuata all'interno di quest'ambiente.

5.1.1 Servizi per il recupero dei dati lessicali

In questa sezione vengono riportati i servizi per l'ottenimento di dati lessicali, come ad esempio informazioni su un'entrata lessicale, forma, senso, etimologia ed altri elementi che fanno riferimento al modello OntoLex-Lemon.

Get Lexical Entries

Questa chiamata serve per recuperare l'alberatura delle entrate lessicali e per effettuare ricerche attraverso i filtri dell'albero per recuperare le entrate lessicali in base ai parametri impostati dall'utente

- API: POST /lexicon/data/lexicalEntries
- PAYLOAD:
 - key : string
- BODY:
 - text : string
 - searchMode : string ("startsWith" | "contains" | "end" | "equals")
 - type : string
 - pos : string
 - formType : string ("flexed" | "entry")
 - author : string
 - lang : string
 - status : string
 - offset : number
 - limit : number
- RESPONSE: [LexicalElement\[\]](#)

Get Lexical Entry

Questa chiamata serve per recuperare i dati relativi a una lexical entry, si attiva quando un utente clicca sull'albero delle entrate lessicali, i dati in risposta vengono inviati al componente dell'editor lessicale.

- API: GET /lexicon/data/lexicalEntry
- PAYLOAD:
 - id : string
 - key : string
 - aspect: string ("core" | "decomposition" | "vartrans" | "synsem")

- RESPONSE: [LexicalElement](#)

Get Filtered Forms

Chiamata per ottenere una lista di forme in base ai parametri di filtering impostati dall'utente

- API: POST /lexicon/data/filteredForms
- PAYLOAD:
 - key : string
- BODY:
 - text : string
 - searchMode : string ("startsWith" | "contains" | "end" | "equals")
 - author : string
 - representationType : string
 - writtenRep
 - phoneticRep
 - pronunciation
 - romanization
 - segmentation
 - transliteration
 - offset : number
 - limit : number
- RESPONSE: [LexicalElement\[\]](#)

Get Forms

Chiamata utilizzata per ottenere la lista completa delle forme di una Lexical Entry.

- API: GET /lexicon/data/forms
- PAYLOAD:
 - id : string
 - key : string
- RESPONSE: [Form\[\]](#)

Get Form

Chiamata utilizzata per ottenere i dati di una forma selezionata nell'albero delle entrate lessicali.

- API: GET /lexicon/data/form
- PAYLOAD:
 - id : string

- aspect : string ("core" | "decomposition" | "vartrans" | "synsem")
- key : string
- RESPONSE: [Form](#)

Get Filtered Senses

Chiamata per ottenere una lista di forme in base ai parametri di filtering impostati dall'utente

- API: POST /lexicon/data/filteredSense
- PAYLOAD:
 - key : string
- BODY:
 - text : string
 - searchMode : string ("startsWith" | "contains" | "end" | "equals")
 - author : string
 - type : string
 - field : string
 - definition
 - description
 - gloss
 - senseExample
 - senseTranslation
 - pos : string
 - formType : string
 - lang : string,
 - status : string
 - offset : number
 - limit : number
- RESPONSE: [Form\[\]](#)

Get Senses

Chiamata utilizzata per ottenere la lista completa dei sensi di una Lexical Entry.

- API: GET /lexicon/data/senses
- PAYLOAD:
 - id : string
 - key : string

- RESPONSE: [Sense\[\]](#)

Get Sense

Chiamata utilizzata per ottenere i dati di un senso selezionato nell'albero delle entrate lessicali.

- API: GET /lexicon/data/lexicalSense
- PAYLOAD:
 - id : string
 - aspect : string ("core" | "decomposition" | "vartrans" | "synsem")
 - key : string
- RESPONSE: [Sense](#)

Get Etymologies

Chiamata utilizzata per ottenere la lista completa delle etimologie di una Lexical Entry

- API: GET /lexicon/data/etymologies
- PAYLOAD:
 - id : string
 - key : string
- RESPONSE: [Etymology\[\]](#)

Get Etymology

Chiamata utilizzata per ottenere i dati di un'etimologia selezionata nell'albero delle entrate lessicali.

- API: GET /lexicon/data/etymology
- PAYLOAD:
 - id : string
 - key : string
- RESPONSE: [Etymology](#)

5.1.2 Servizi per il recupero di dati statistici e di vocabolario

In questa sezione vengono elencati i servizi per il recupero di dati relativi alle statistiche sul lessico, come ad esempio le lingue, tipi di entrate lessicali presenti e le part of speech delle unità lessicali con le relative occorrenze.

Get Lexical Entry Types

Chiamata per ottenere i tipi di entrate lessicali in un lessico di riferimento.

- API: GET /lexicon/statistics/type

- PAYLOAD:
key : string
- RESPONSE: [StatisticsCounter\[\]](#)

Get Authors

Chiamata per ottenere l'elenco degli autori delle entrate lessicali, con annesso il numero delle occorrenze delle entrate lessicali su cui hanno lavorato.

- API: GET /lexicon/statistics/authors
- PAYLOAD:
key : string
- RESPONSE: [StatisticsCounter\[\]](#)

Get Languages

Chiamata per ottenere l'elenco delle lingue presenti nel lessico, con annesso il numero delle occorrenze delle entrate lessicali a cui fanno riferimento.

- API: GET /lexicon/statistics/languages
- PAYLOAD:
key : string
- RESPONSE: [StatisticsCounter\[\]](#)

Get Pos

Chiamata per ottenere l'elenco delle part of speech utilizzate per modificare le entrate lessicali, con annesso il numero relativo delle occorrenze di tali PoS distribuite nel lessico.

- API: GET /lexicon/statistics/languages
- PAYLOAD:
key : string
- RESPONSE: [StatisticsCounter\[\]](#)

Get Status

Chiamata per ottenere l'elenco degli stati delle entrate lessicali, con allegato il numero delle occorrenze di entrate lessicali legate per ogni stato.

- API: GET /lexicon/statistics/status
- PAYLOAD:
key : string

- RESPONSE: [StatisticsCounter\[\]](#)

5.1.3 LexO Data Model

In questa sezione vengono elencate le interfacce/modelli dei dati relativi agli elementi del lessico computazionale, ovvero tutti i dati provenienti da LexO-Server, conformi alle specifiche del modello Ontolex-lemon e sue estensioni.

StatisticsCounter

Questo tipo di elemento corrisponde a tutti gli elementi statistici relativi ai tipi di elementi lessicali, alle loro *part of speech*, autori, lingue e stati.

CAMPO	TIPO	DESCRIZIONE
label	string	Tipologia del campo
count	number	Numero di elementi

LexicalElement

Un *lexical element* è un oggetto JSON composto da vari campi che contengono le informazioni relative a un'entrata lessicale. Essa costituisce l'unità base del lessico computazionale.

CAMPO	TIPO	DESCRIZIONE
creator	string null	Nome del creatore della lexical entry
lastUpdate	string	Timestamp dell'ultima modifica
creationData	string	Timestamp della creazione
confidence	number	Affidabilità delle informazioni sulla lexical entry. Può avere due valori (-1 e 0)
lexicalEntry	string	IRI della lexical entry
label	string	Etichetta contenente il valore da visualizzare nell'albero

type	Array<string>	Contiene i valori relativi al tipo della lexical entry, più contenere i seguenti valori: - LexicalEntry - Word - Affix - MultiWordExpression - Cognate - Etymon
pos	string null	Part of Speech della lexical entry
language	string null	Lingua a cui appartiene la lexical entry
stemType	string null	Radice della lexical entry
morphology	Array< Morphology >	Tratti morfologici della lexical Entry
links	Array< Link >	Riferimento esterni, bibliografia, note, seeAlso, sameAs
author	string	Autore dell'entrata
revisor	string null	Revisore dell'entrata
status	string	Stato dell'entrata lessicale, gli stati possibili sono: - working - completed - reviewed
note	string	Note relative all'entrata lessicale
completionDate	string	Data di completamento dell'entrata lessicale
revisionDate	string	Data di revisione dell'entrata lessicale
hasChildren	boolean	Indica se l'elemento ha figli (sensi, forme, etimologie)

Form

Una *form* è un oggetto JSON composto da vari campi che contengono le informazioni relative a una forma di un'entrata lessicale.

CAMPO	TIPO	DESCRIZIONE
-------	------	-------------

creator	string null	Nome del creatore della lexical entry
lastUpdate	string	Timestamp dell'ultima modifica
creationDate	string	Timestamp della creazione
confidence	number	Affidabilità delle informazioni sulla forma. Può avere due valori (-1 e 0)
form	string	IRI della forma
inheritedMorphology	Array< Morphology >	Tratti morfologici (PoS) ereditati dalla lexical entry di riferimento
lexicalEntry	string	IRI della lexical entry
lexicalEntryLabel	string	Etichetta con il titolo della lexical entry madre
label	Array< Property >	Elementi caratterizzanti di una forma
type	Array<string>	Contiene i valori relativi al tipo della lexical entry, può contenere i seguenti valori: <ul style="list-style-type: none"> - LexicalEntry - Word - Affix - MultiWordExpression - Cognate - Etymon
language	string null	Lingua a cui appartiene la forma (ereditata dalla lexical entry)
morphology	Array< Morphology >	Tratti morfologici della forma
links	Array< Link >	Riferimento esterni, bibliografia, note, seeAlso, sameAs
note	string	Note relative all'entrata lessicale

Sense

Un *sense* è un oggetto JSON composto da vari campi che contengono le informazioni relative a un senso di un'entrata lessicale.

CAMPO	TIPO	DESCRIZIONE
creator	string null	Nome del creatore della lexical entry
lastUpdate	string	Timestamp dell'ultima modifica
creationDate	string	Timestamp della creazione
confidence	number	Affidabilità delle informazioni sul senso
explanation	string null	Spiegazione del senso
sense	string	IRI del senso
senseExample	string null	Esempio di uso del senso
senseTranslation	string null	Traduzione del senso
gloss	string null	Glossa di un significato / senso
links	Array<Link>	Riferimento esterni, bibliografia, note, seeAlso, sameAs
status	string	Stato dell'entrata lessicale, gli stati possibili sono: - working - completed - reviewed
note	string	Note relative all'entrata lessicale
usage	string	Uso del senso

EtymologySet, EtyLink ed Etymology

Un *EtymologySet* è un oggetto che viene restituito quando si clicca su un'etimologia nell'albero delle entrate lessicali; esso contiene due campi relativi all'etimologia in sé e agli *etylinks* (riferimenti dell'etimologia) che essa contiene

CAMPO	TIPO	DESCRIZIONE
etyLinks	Array<EtyLink>	Array con tutti gli etyLinks legati all'etimologia

etymology	Etymology	Oggetto contenente i dati dell'etimologia
-----------	-----------	---

Ecco un modello di EtyLink e, successivamente, di Etymology.

CAMPO	TIPO	DESCRIZIONE
confidence	number	Affidabilità delle informazioni sull'etimologia
creationDate	string	Data di creazione
creator	string	Creatore del link
etyLinkType	string	Tipo di etyLink. Può avere due valori (inheritance, borrowing)
etySource	string	IRI all'etymon di partenza
etySourceLabel	string	Label dell'etymon di partenza
etyTarget	string	IRI all'etymon di destinazione
etyTargetLabel	string	Label dell'etymon di destinazione
etymologicalLink	string	IRI dell'etyLink
externalIRI	boolean	Se è un IRI esterno o no
label	string	Etichetta dell'etyLink
lastUpdate	string	Ultima modifica
note	string	Note dell'etyLink

CAMPO	TIPO	DESCRIZIONE
confidence	number	Affidabilità delle informazioni sull'etimologia
creationDate	string	Data di creazione
creator	string	Creatore del link
etymology	string	IRI dell'etimologia
hypotesisOf	string	Autore dell'etimologia

label	string	Etichetta dell'etyLink
lastUpdate	string	Ultima modifica
links	Array< Link >	Riferimento esterni, bibliografia, note, seeAlso, sameAs
note	string	Note dell'etyLink

5.2 Servizi dei testi (CASH)

I servizi relativi ai testi sono erogati dal servizio di CASH, il quale comunica con un database MySQL per la gestione dei documenti XML caricati sulla piattaforma e delle associazioni tra testo e lessico. Tutte le API funzionano senza la necessità di avere un'autorizzazione o autenticazione particolari.

5.2.1 Servizi per le attestazioni e la ricerca

In questo paragrafo vengono elencati i servizi che si occupano del recupero e della ricerca delle attestazioni.

Get Tokens

Questo servizio serve per ottenere tutti i token di un testo a partire da un suo identificativo.

- API: `GET /api/public/token`
- PAYLOAD:

```
requestUUID : string
nodeid : number
```
- RESPONSE: [Token\[\]](#)

Get Content

Questo servizio serve per ottenere il contenuto XML di un file a partire da un suo identificativo.

- API: GET /api/public/getcontent
- PAYLOAD:
 - requestUUID : string
 - nodeid : number
- RESPONSE: [XmlAndId](#)

Get Annotations

Questo servizio serve per ottenere la lista di attestazioni presenti in un file a partire da un suo identificativo.

- API: GET /api/public/annotation
- PAYLOAD:
 - requestUUID : string
 - nodeid : number
- RESPONSE: [Annotation\[\]](#)

Search Files

Questo servizio serve per effettuare una ricerca all'interno della collezione di testi presenti in CASH.

- API: POST /api/public/searchFiles
- PAYLOAD:
 - query: CQL string
 - offset: number
 - limit: number

- BODY: {}
- RESPONSE: [TextMetadata\[\]](#)

CountFiles

Questo servizio serve per restituire il numero totale di iscrizioni che corrispondono ad una ricerca.

- API: POST /api/public/countFiles
- PAYLOAD:
 - query: CQL string
 - offset: number
 - limit: number
- BODY: {}
- RESPONSE: number

GetUniqueMetadata

Questo servizio serve per restituire tutti i valori presenti a partire da un metadato dell'iscrizione

- API: POST /api/public/uniqueMetadataValues
- PAYLOAD:
 - field: string
 - offset: number
 - limit: number
- BODY: {}
- RESPONSE: string[]

5.2.2 CASH Data Model

In questa sezione vengono elencate le interfacce/modelli dei dati relativi alle iscrizioni presenti nel repository di CASH.

Token

Questo modello corrisponde a un body da mandare come richiesta per creare un token all'interno di un file a partire da una stringa di testo.

CAMPO	TIPO	DESCRIZIONE
text	string	Testo selezionato dell'attestazione
xml-id	number	Identificativo del token
position	number	Posizione del token all'interno del testo
begin	number	Porzione iniziale dell'intervallo corrispondente alla sottostringa del testo dell'annotazione
end	number	Porzione finale dell'intervallo corrispondente alla sottostringa del testo dell'annotazione
node	number	Identificativo corrispondente all'ID del testo
source	string	Origine del testo
imported	boolean	Se il testo è stato importato o creato da zero
id	number	Identificativo del testo

XmlAndId

In questo modello sono presenti un campo relativo al contenuto XML di un'iscrizione epigrafica e l'ID relativo per il recupero di altre informazioni da CASH da allegare al contenuto del testo.

CAMPO	TIPO	DESCRIZIONE
xml	string	Testo dell'attestazione
id	number	Identificativo dell'attestazione

Annotation

Questo modello corrisponde a un body da mandare come richiesta per creare una nuova attestazione tra testo e lessico. L'oggetto contiene valori come l'IRI della forma attestata, l'intervallo di caratteri del token in cui è attestata l'entità lessicale ed eventuali metadati.

CAMPO	TIPO	DESCRIZIONE
layer	string	Livello attestazione
value	string	Valore contenuto dall'attestazione; generalmente, contiene l'IRI della forma
imported	boolean	Booleano per indicare se la risorsa è importata o meno
attributes	AnnotationAttributes	Attributi dell'attestazione
spans	Span[]	Intervallo corrispondente alla sotto-stringa del token a cui fa riferimento
id	number	Identificativo dell'attestazione

AnnotationAttributes

Questo modello corrisponde a un body da mandare come richiesta per creare una nuova attestazione tra testo e lessico. L'oggetto contiene valori come l'IRI della forma attestata, l'intervallo di caratteri del token in cui è attestata l'entità lessicale ed eventuali metadati.

CAMPO	TIPO	DESCRIZIONE
author	string	Autore dell'attestazione
creator	string	Creatore dell'attestazione
note	string	Nota dell'attestazione
confidence	number	Confidenza dell'attestazione
form_id	string	Nome dell'istanza lessicale
validity	string	Validità dell'attestazione
label	string	Etichetta da visualizzare
timestamp	string	Timestamp della creazione

		dell'attestazione
bibliography	Array< AnnotationBibliography >	Elementi bibliografici legati all'attestazione
node_id	number	Identificativo del nodo in cui è avvenuta l'attestazione; corrisponde al testo

Span

Lista di oggetti contenente gli indici corrispondenti all'intervallo della stringa di testo selezionata per effettuare l'attestazione

CAMPO	TIPO	DESCRIZIONE
start	number	Parte iniziale della stringa
end	number	Parte finale della stringa

BodyTextPart

Oggetto relativo alle tecniche e ai modi con cui è stata incisa o scritta un'iscrizione. Può essere un array o anche un oggetto singolo.

CAMPO	TIPO	DESCRIZIONE
ductus	string	Modo con cui viene eseguita una scrittura
section	string	Sezione dell'iscrizione
textDirection	string	Direzione del testo

ConservationInstitution

Coppia di chiave valore contenente le informazioni sul luogo di conservazione dell'iscrizione.

CAMPO	TIPO	DESCRIZIONE
-------	------	-------------

name	string	Nome dell'istituzione che conserva l'epigrafe
URL	string	Indirizzo web dell'istituzione

ObjectDimension

Oggetto contenente le informazioni sulle dimensioni dell'iscrizione.

CAMPO	TIPO	DESCRIZIONE
depth	string	Profondità supporto
height	string	Altezza supporto
type	string	Tipologia del supporto
unit	string	Unità di misura
width	string	Larghezza supporto
precision	string	Precisione della misurazione

DiscoveryYear

Oggetto relativo alla datazione di un'iscrizione.

CAMPO	TIPO	DESCRIZIONE
notAfter	string	Limite superiore della datazione
notBefore	string	Limite inferiore della datazione
when	string	Data precisa

Facsimile

Oggetto relativo alle immagini di un'iscrizione e la relativa licenza.

CAMPO	TIPO	DESCRIZIONE
desc	string	Descrizione immagine

license	string	Informazioni sulla licenza
licenseUrl	string	URL della licenza
url	string	URL dell'immagine

LanguageMetadata

Oggetto relativo al linguaggio dell'iscrizione secondo lo standard ISO.

CAMPO	TIPO	DESCRIZIONE
ident	string	Identificativo, codice lingua
source	string	URL sul codice ISO della lingua

License

Oggetto relativo alle licenze e al copyright dell'iscrizione.

CAMPO	TIPO	DESCRIZIONE
distributionLicense	string	Nome della licenza
distributionLicense_URL	string	URL della licenza

PlaceModel

Oggetto relativo al luogo di appartenenza dell'epigrafe.

CAMPO	TIPO	DESCRIZIONE
ancientName	string	Nome antico
ancientNameUrl	string	URL al repository Pleiades
modernName	string	Nome nuovo
modernNameUrl	string	URL al repository GeoNames

Settlement

Oggetto relativo al luogo di conservazione

CAMPO	TIPO	DESCRIZIONE
placeofConservation	string	Nome del luogo di conservazione
placeOfConservationExternalReference	string	URL del luogo di conservazione

TraditionalIDs

Oggetto relativo agli identificativi tradizionali dell'iscrizione. Può essere un oggetto singolo o un array di oggetti.

CAMPO	TIPO	DESCRIZIONE
traditionalID	string	Identificativo tradizionale
traditionalID_URL	string	URL dell'identificativo tradizionale

Trismegistos

Oggetto relativo all'identificativo Trismegistos dell'iscrizione.

CAMPO	TIPO	DESCRIZIONE
trismegistosID	string	Identificativo Trismegistos
trismegistosID_URL	string	URL dell'identificativo Trismegistos

SupportModel

Oggetto relativo alle informazioni del supporto epigrafico.

CAMPO	TIPO	DESCRIZIONE
material	string	Materiale usato
material_conceptUrl	string	Riferimento al concetto del materiale usato

objectType	string	Tipo oggetto
ObjectType_conceptUrl	string	Riferimento al concetto del tipo di oggetto

TextMetadata

Oggetto contenente coppie chiave valore, ognuna corrispondente a un metadato ben preciso.

CAMPO	TIPO	DESCRIZIONE
alphabet	string	Alfabeto di appartenenza
alphabet_url	string	URL per informazioni su alfabeto di appartenenza
ancientFindSpotName	string	Nome antico del luogo di ritrovamento
authority	string	Nome dell'autorità
autopsy	string	Tecnica dell'autopsia
autopsyAuthor	string	Nome di chi ha condotto i lavori autoptici sull'iscrizione
autopsyDate	string	Data autopsia
bodyTextPart	BodyTextPart[] & BodyTextPart	Informazioni sul corpo dell'iscrizione
condition	string	Condizioni supporto
conditionDesc	string	Descrizione su condizioni supporto
conservationInstitution	ConservationInstitution	Dati su istituzione che conserva l'epigrafe
dateOfOrigin	string	Data origine
dateOfOriginNotAfter	string	Limite inferiore data origine
dateOfOriginNotBefore	string	Limite superiore data origine
datingCertainty	string	Certezza sulla datazione
decoration	string	Informazioni sulle decorazioni
detailedFindSpot	string	Luogo preciso di ritrovamento

dimensions	ObjectDimension	Dimensioni dell'iscrizione
discoveryYear	DiscoveryYear	Datazione dell'iscrizione
element-id	number	Identificativo dell'elemento in CASH
editor	string	Editore dell'iscrizione
encoding	string	Informazioni sulla codifica
encodingSourceUrl	string	URL della codifica
execution	string	Tipo di esecuzione
execution_conceptUrl	string	URL dell'esecuzione
facsimile	Facsimile[]	Immagini dell'iscrizione
fileID	string	Identificativo del nome del file in CASH
inscriptionTitle	string	Titolo dell'iscrizione
inscriptionType	string	Tipologia dell'iscrizione
inventoryNumber	string	Numero di inventario
itAnt_ID	string	Identificativo dell'iscrizione nel progetto ItAnt
language	LanguageMetadata	Informazioni sulla lingua
layoutNotes	string	Note sul layout dell'iscrizione
license	License	Informazioni sulla licenza
opistography	string	Stringa per indicare se l'iscrizione è scritta su entrambi i lati, fronte e retro
originalPlace	PlaceModel	Informazioni sul luogo di appartenenza
palaeographicNotes	string	Informazioni paleografiche
publicationPlace	string	Luogo di pubblicazione
publicationPID	string	Identificativo del luogo di pubblicazione
reuse	string	Informazioni sul riuso
settlement	Settlement	Informazioni sul luogo di conservazione
summary	string	Riassunto delle informazioni

		sull'iscrizione
support	SupportModel	Informazioni sul supporto dell'iscrizione
title	string	Titolo iscrizione
traditionalIDs	TraditionalIDs[] & TraditionalIDs	Informazioni sugli identificativi tradizionali
trismegistos	Trismegistos	Informazioni su identificativo Trismegistos
wordDivisionType	string	Informazioni su divisione delle parole
writingSystem	string	Sistema di scrittura

5.3 GeoData Services (Pleiades & Geonames)

I servizi relativi ai dati geografici sono offerti da servizi esterni, ovvero Pleiades e Geonames, che mettono a disposizione le loro API per poter effettuare interrogazioni ai loro database ed ottenere quindi informazioni su, ad esempio, latitudine e longitudine di un determinato luogo (che è rappresentato da un identificativo numerico) e altre informazioni minori. Questo servizio serve per la renderizzazione delle mappe (la libreria open source in uso è Leaflet¹⁰)

5.3.1 Pleiades

In questa sezione viene illustrato il servizio per il recupero delle informazioni dal repository Pleiades.

GetPlace

Questo servizio serve per ottenere tutti i token di un testo a partire da un suo identificativo.

- API: GET <https://pleiades.stoa.org/places/{ancientPlaceId}/json>
- PAYLOAD: none
- RESPONSE: [PleiadesDataModel](#)

5.3.2 Pleiades Data Model

In questa sezione viene illustrato il modello dei dati in risposta ad una richiesta avanzata ai sistemi di Pleiades

¹⁰ <https://leafletjs.com/>

PleiadesDataModel

CAMPO	TIPO	DESCRIZIONE
bbox	number[]	Lista contenente latitudine e longitudine del box per circoscrivere l'area del luogo
id	string	Identificativo del luogo
uri	string	URI del luogo
description	string	Descrizione del luogo
reprPoint	number[]	Lista contenente le coordinate del luogo
title	string	Titolo del luogo

5.3.3 GeoNames

In questa sezione viene illustrato il servizio per il recupero delle informazioni dal repository GeoNames.

GetPlace

Questo servizio serve per ottenere tutti i token di un testo a partire da un suo identificativo.

- API: GET `http://api.geonames.org/getJSON`
- PAYLOAD:
 - `formatted: boolean`
 - `geoNameId: string`
 - `username: string`
 - `style: string`
- RESPONSE: [GeoNamesDataModel](#)

5.3.4 GeoNames Data Model

In questa sezione viene illustrato il modello dei dati in risposta ad una richiesta avanzata ai sistemi di Pleiades

GeoNamesDataModel

CAMPO	TIPO	DESCRIZIONE
bbox	BBoxModel	Lista contenente latitudine e longitudine del box per circoscrivere l'area del luogo
name	string	Nome del luogo
lat	string	Latitudine del luogo
lng	string	Longitudine del luogo

5.3.5 Global GeoData Model

Questo modello serve per unire le caratteristiche dei due modelli precedenti in modo da combinare entrambe le informazioni (ove possibile, a volte può capitare di non avere informazioni per una o l'altra fonte di dati)

GeoNamesDataModel

CAMPO	TIPO	DESCRIZIONE
ancientName	string	Nome antico
modernName	string	Nome nuovo
ancientBbox	BBoxModel	Bounding Box del luogo antico
attestations	TextMetadata[]	Lista di attestazioni da appendere ad ogni luogo disegnato sulla mappa
modernBbox	BBoxModel	Bounding Box del luogo moderno
modernId	string	Identificativo del luogo moderno
ancientId	string	Identificativo del luogo antico
description	string	Descrizione del luogo
reprPoint	Markers	Coordinate del marcatore
ancientUri	string	Uri del luogo antico
modernUri	string	Uri del luogo moderno

5.3.6 Minor Models

Modelli accessori per la realizzazione delle mappe.

BboxModel

CAMPO	TIPO	DESCRIZIONE
north	number	Nord
south	number	Sud
west	number	Ovest
east	number	Est

Markers

CAMPO	TIPO	DESCRIZIONE
latitude	number	Latitudine
longitude	number	Longitudine

5.4 Servizi della bibliografia

I servizi relativi alla bibliografia sono gestiti da CASH. Per questo tipo di servizio è stato importata tutta la bibliografia presente su Zotero all'interno di questa repository, in quanto i servizi base di quest'ultimo non sono performanti per effettuare ordinamenti e query puntuali. A tal proposito, è stato deciso di ingerire tutte le informazioni pertinenti e di creare un'apposita API che consente di effettuare query più puntuali, precise e veloci.

searchBiblio

Questo servizio serve per richiedere tutti gli elementi bibliografici presenti in CASH. E' possibile inoltre effettuare delle ricerche in base al body passato alla richiesta: è possibile, infatti, inserire all'interno di "filters", una chiave di ricerca, con il relativo valore, e un operatore di ricerca per raffinare la ricerca.

- API: `POST /api/public/searchbiblio`
- PAYLOAD:
 - `offset: number`

- limit: number
- BODY: {


```

"requestUUID": "string",
"filters": [
  {
    "key": "string",
    "value": "string",
    "op": "string"
  }
],
"user-id": "string"
}

```
- RESPONSE: [Book\[\]](#)

5.4.1 Bibliography Data Model

In questa sezione sono presenti i modelli di dati relativi ai servizi della bibliografia.

Book

CAMPO	TIPO	DESCRIZIONE
key	string	Identificativo dell'elemento bibliografico
author	string	Autore dell'elemento bibliografico
editor	string	Editore del riferimento bib.
date	string	Data di pubblicazione
itemType	string	Tipologia riferimento bib.
key	string	Identificativo del libro
notes	string	Note dell'editore
pages	string	Pagine citate

place	string	Luogo di pubblicazione
publisher	string	Editore
references	string	Referenze ad altri elementi bibliografici
title	string	Titolo dell'elemento bibliografico
url	string	URL del riferimento bib.

6. Bibliografia

1. Mallia, Michele; et al., 2023, EpiLexO, ILC-CNR for CLARIN-IT repository hosted at Institute for Computational Linguistics "A. Zampolli", National Research Council, in Pisa, <http://hdl.handle.net/20.500.11752/ILC-1005> .
2. McCrae, John P., et al. "The Ontolex-Lemon model: development and applications." *Proceedings of eLex 2017 conference*. 2017.
3. Scholz, Martin, et al. "EDEN: an epigraphic Web Database of ancient inscriptions." *Proceedings of CIDOC 2014*. ICOM/CIDOC, 2014. 1-7.
4. Carbon, J.-M., & Pirenne-Delforge, V. (December 2019). Two Notes on the Collection of Greek Ritual Norms: Looking Back, Looking Forward. Axon : Greek historical inscriptions, 3 (2). doi:10.30687/Axon/2532-6848/2019/02/007