CLARIN-IT

The Italian Common Language Resources and Technology Infrastructure



Infrastrutture di ricerca nel settore Umanistico

Monica Monachini - CLARIN Italian National Coordinator

ILC-CNR
National Representative

Perché siamo qui



Infrastrutture: perché saperne di più



- Diego Piacentini, super-cervello, uno dei massimi esperti di **infrastrutture digitali**, 13 anni in Apple e 16 ad Amazon, è stato nominato la settimana scorsa dal governo “Commissario straordinario per il digitale”.
- La sua missione: "Rendere i servizi pubblici accessibili nel modo più **semplice** possibile al cittadino”
- “… condizione necessaria per **sviluppo**, perché **stimola** gli investimenti anziché frenarli".

Definizioni di partenza delle Infrastrutture







Per infrastruttura s'intende

- **una serie di elementi** che media i **rapporti tra i vari componenti** di una struttura
- **una rete** di impianti e servizi **interconnessi** tra loro da specifici **punti nodali**.
- le reti dei **trasporti**, sia di energia che di mezzi e persone, strade, ferrovie, canali, gasdotti, oleodotti, …
- … ma anche reti di **telecomunicazione** (rete telefonica, emittente televisiva, emittente radiofonica ma anche la rete informatica nel suo complesso)
- potenziando le comunicazioni, sono un volano per lo sviluppo **economico**

Le Infrastrutture di Ricerca





- Le Infrastrutture di ricerca (IR) sono una **rete di centri** coordinati che forniscono **accesso** internazionale e multidisciplinare a **dati, strumenti e servizi** per un settore specifico
- rappresentano strumenti essenziali per lo **sviluppo scientifico** e tecnologico del Paese
- hanno un impatto straordinario:
  - realizzano una spinta formidabile sulle metodologie ed organizzazione del lavoro tecnico-scientifico
  - attraggono i migliori talenti
  - catalizzano la concentrazione di attività avanzate
  - creano l'humus per attività fortemente innovative

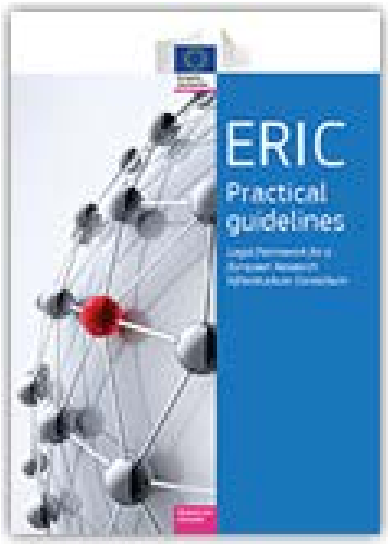
Organi e strumenti di gestione

- L'ESFRI (European Strategy Forum on Research Infrastructures) è un organo creato nel 2002 con cui si **definisce** la **politica** europea per le IR.

- Ha il compito di realizzare e implementare la **roadmap** strategica per la loro costruzione, operatività e integrazione

- L'ERIC (Consorzio Europeo per le Infrastrutture di Ricerca) istituito nel 2009 rappresenta uno **strumento** legale per facilitare la **creazione** e il **funzionamento** delle Infrastrutture di Ricerca europee superando gli ostacoli derivanti dai differenti ordinamenti giuridici vigenti nei vari paesi dell'Unione





Cosa si intende per Infrastruttura di Ricerca?

Una prospettiva europea

- Impianti
- Risorse
- Servizi

+ Attività di livello «top»

× Unicità

=

European Strategy Forum on Research Infrastructures ESFRI

Research Infrastructure

Commissione europea

Localizzate | **Virtuali** | **Distribuite**

Accessibili | **Internazionali**

Grandi strumentazioni | **Centri di attrazione per ricercatori**

Cosa si intende per Infrastruttura di Ricerca?

Una definizione meno esclusiva: "facility" di ricerca scientifico-tecnologica

«…Insieme di strumentazioni, competenze, servizi o funzioni specializzate e dedicate a realizzare in modo continuativo i flussi di servizi necessari alla ricerca scientifica e tecnologica…»

Al dato fisico si aggiunge quello intangibile delle competenze, del saper fare, del facilitare complessivamente un insieme di attività → **Facility**

indica prevalentemente uno strumento o un insieme di strumenti, ma anche una particolare abilità individuale ovvero facilità offerta per un uso pieno di una cosa.

IR: strategia europea condivisa



- La prima **roadmap** italiana delle IR redatta dal MIUR nel Giugno 2010 **individua** un portafoglio di **progetti** infrastrutturali **maturi** ed emergenti per affrontare le sfide che si presentano alla ricerca scientifica

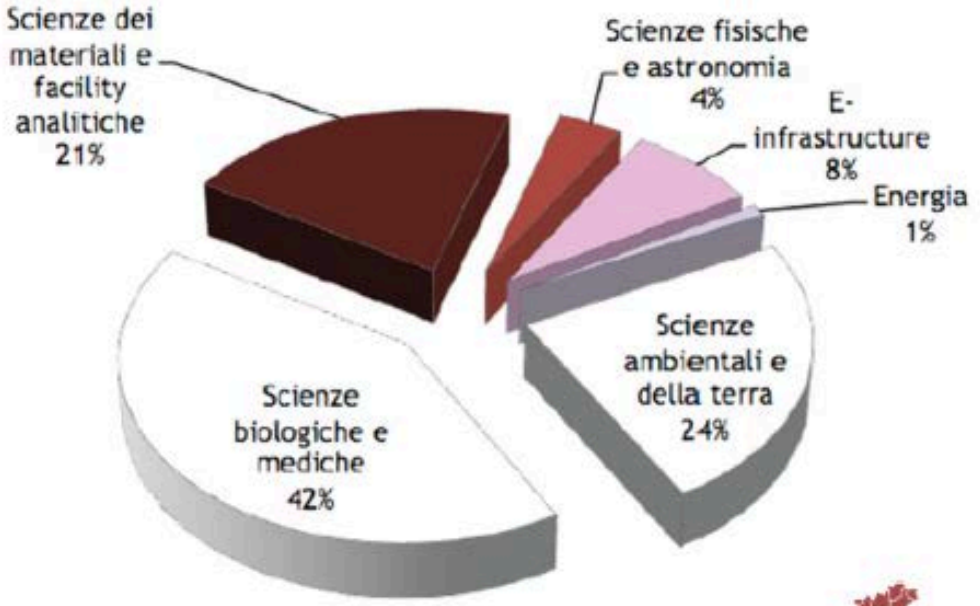
- Le categorie adottate internazionalmente: Materials, Physical Sciences and Engineering, Energy, Biological and Medical Sciences, Environment, **Social Science and Humanities**, e-Infrastructures

Una Mappatura delle IR

Uno sguardo dall'alto



Distribuzione dei Settore Scientifici



La scelta dei domini scientifici è basata, per convenzione, su quelli individuati dall' ***ESFRI***





Credits: Quaderni CARIPLO 2014

IR nel settore delle scienze umane:



Infrastrutture nel settore delle Scienze Umane e Sociali		
Charisma	conservazione e restauro	https://ec.europa.eu/research/infrastructures/pdf/charisma.pdf
Cendari	archivi e risorse per la storia medievale e moderna	http://www.cendari.eu
Dariah	tecnologie digitali per l'arte	www.dariah.eu
Ariadne	archeologia	http://www.ariadne-infrastructure.eu

1° Ottobre 2015

- L'Italia diventa membro della Infrastruttura CLARIN-ERIC
  Common Language Resource Infrastructure for Social Sciences and Humanities



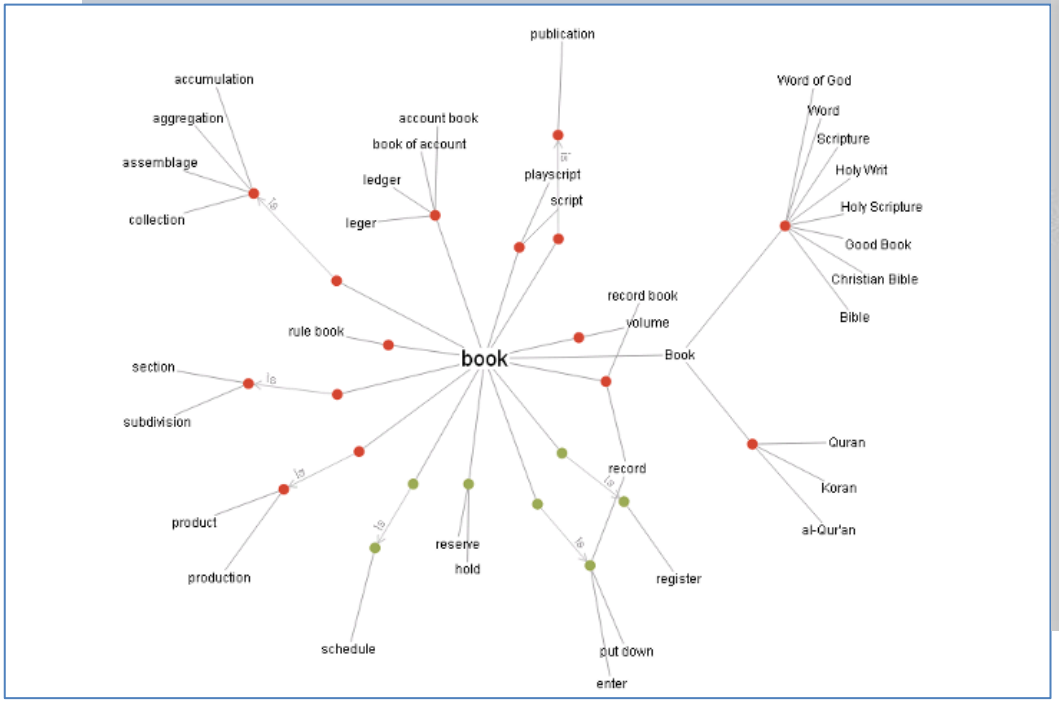
- Una opportunità per chi si occupa di Discipline Umanistiche

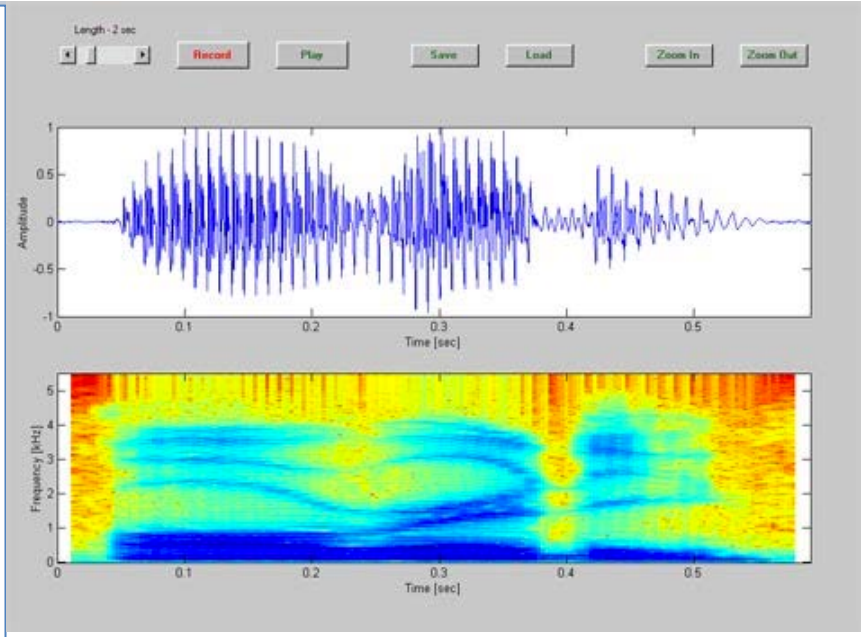
Cosa sono le Risorse Linguistiche

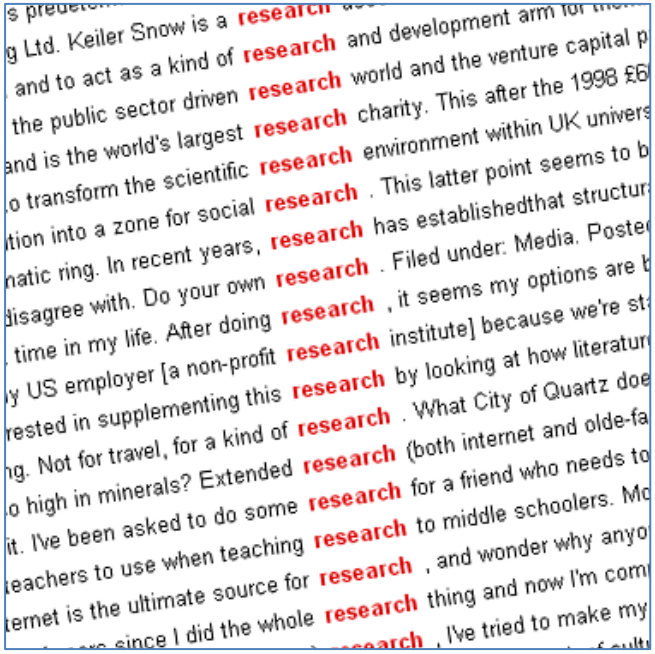


- Il termine Risorse linguistiche si riferisce ad un insieme di dati relativi al linguaggio scritto o parlato in formato digitale, utilizzati per la costruzione, il miglioramento o la valutazione di algoritmi per l'analisi del linguaggio naturale

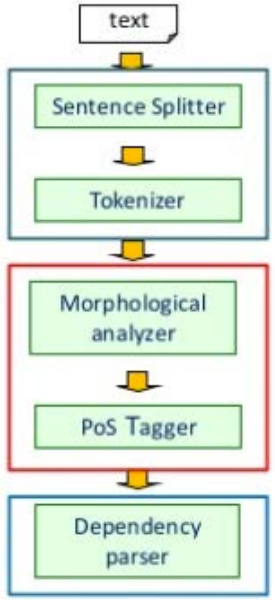
- Esempi di risorse linguistiche sono i corpora (scritti o parlati, annotati o non) lessici, terminologie, collezioni di parlato ma anche strumenti per l'analisi linguistica automatica…"







- **Segmentazione in frasi e tokenizzazione** (ovvero segmentazione del testo in parole ortografiche o tokens
- **Annotazione morfo-sintattica**
  - a ogni token del testo viene associata informazione relativa alla categoria grammaticale che la parola ha nel contesto specifico e il relativo lemma
- **Annotazione sintattica a dipendenze**
  - analisi della struttura sintattica della frase in termini di relazioni di dipendenza (es. soggetto, oggetto, etc.)



A cosa servono?



"A corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research."

"A Lexical resource is a collection of lexical items with additional linguistic information and/or classification of these items. Digital lexical resources are organised as lexical databases, where information is stored in a structured way"

A cosa servono?



- Per lo studio e l'insegnamento delle lingue costituiscono la risorsa di riferimento

- I lessici possono essere usati per studiare fenomeni specifici che non sono facilmente isolabili nei dizionari cartacei (pensiamo agli studi derivazionali)

- I corpora consentono lo studio del linguaggio nel reale contesto di uso e sono più affidabili dell'introspezione; consentono di osservare la distribuzione delle parole e la frequenza, i costrutti sintattici e i sensi.

  - Nei sistemi computazionali sono la materia prima per i sistemi automatici che richiedono dati linguistici per imparare a eseguire l'analisi del linguaggio

  - I corpora annotati consentono di stabilire quale sistema funziona meglio.
  - un corpus contenente testi in inglese e le loro traduzioni in spagnolo può essere utilizzato per la formazione di un sistema di traduzione automatica,
  - e un altro simile corpus verrà utilizzato per valutare i risultati
  - i lessici computazionali sono inoltre richieste da molti algoritmi che hanno necessità di accedere a informazioni lessicale

… e nelle Humanities?



Forniscono gli strumenti per assistere gli studiosi del testo

Nelle edizioni digitali le risorse e gli strumenti linguistici consentendo di studiare il testo a rendere esplicita la struttura del testo e la sua semantica

Potenziano lo studio della informazione linguistica e della semantica del testo rendendola esplicita per mezzo dell'annotazione e per mezzo del legame a fonti esterne

Situazione e bisogni





- Le risorse
  - risiedono presso il produttore
  - relegate agli scopi e ambiti scientifici del progetto
  - gravate da vincoli
  - difficili da reperire

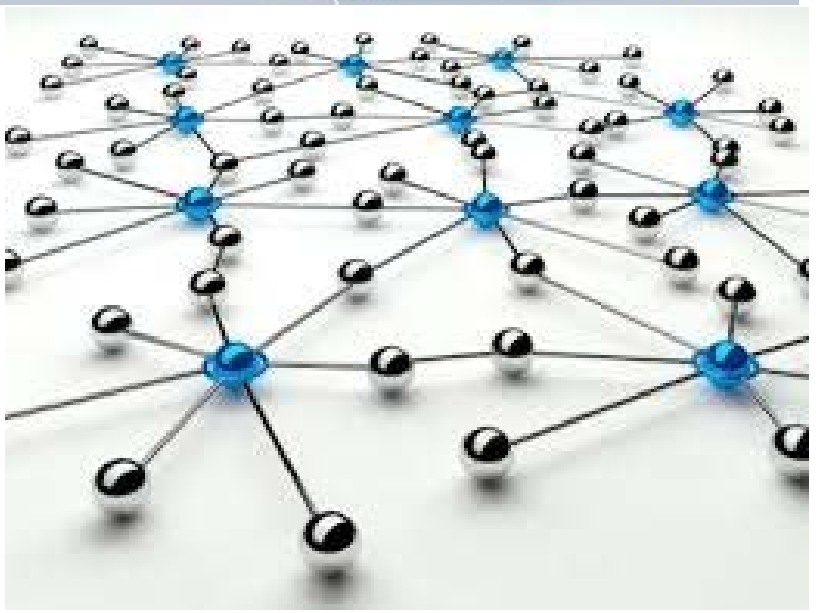


- Costruire le risorse richiede tempo e denaro:
  - cruciale garantire un facile accesso alle risorse per i ricercatori di diverse istituzioni e paesi
  - fondamentale garantire la loro persistenza e fruibilità nel tempo



- Un cambio di paradigma:
  - condivisione, accesso, replicabilità dei risultati

CLARIN: alcune risposte



- CLARIN si configura come una federazione di centri di vario tipo, i pilastri dell'infrastruttura dedicata a soddisfare le esigenze della comunità di studiosi

- Ciascuna nazione paga una tassa annua di iscrizione

- Identifica un Esecutore Nazionale (tipicamente in Italia CNR)

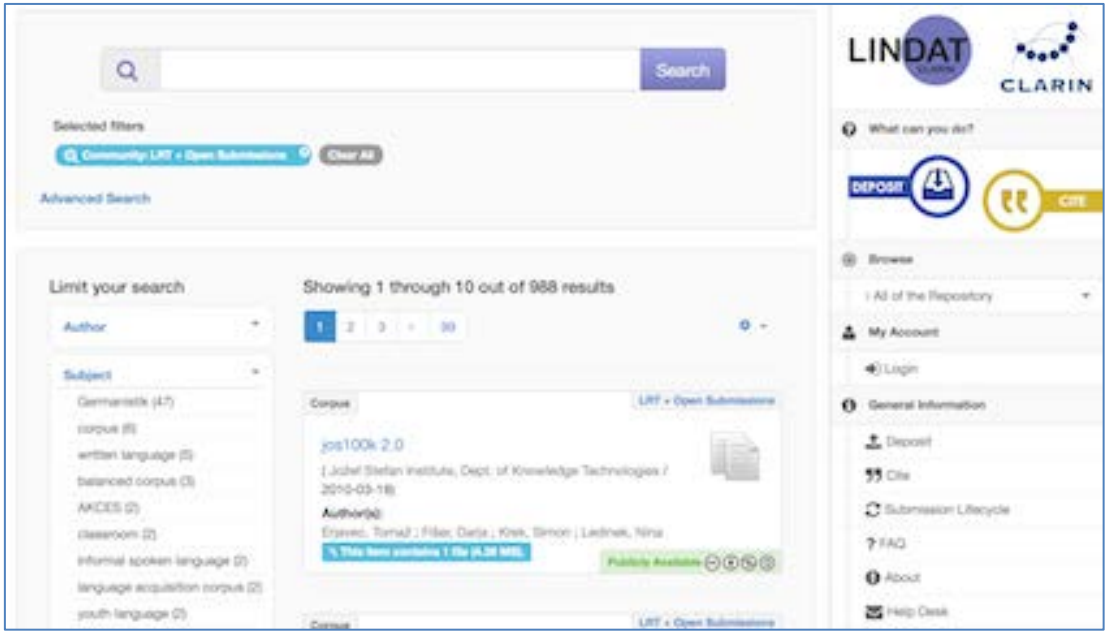
- Uno (o più) CENTRI CLARIN mettono a disposizione le proprie risorse
- offrono una serie di servizi linguistici in modo stabile e persistente fruibili liberamente o tramite autenticazione
- promuovono la collaborazione
- coordinano le azioni necessarie


Mappa
Satellite
Svezia
Norvegia
Finlandia
Mar Baltico
Mar del Nord
Danimarca
Regno Unito
Irlanda
Polonia
Bielorussia
Mosca
Москва
Germania
Belgio
Ucraina
Moldavia
Ungheria
Francia
Romania
Croazia
Serbia
Italia
Roma
Bulgaria
Mar Nero
Georgia
Portogallo
Madrid
Spagna
Mar Tirreno
Grecia
Istanbul
Turchia
Azerbaiian
Siria
Libano
Tunisia
Mar Mediterraneo
Google

CLARIN: alcune risposte

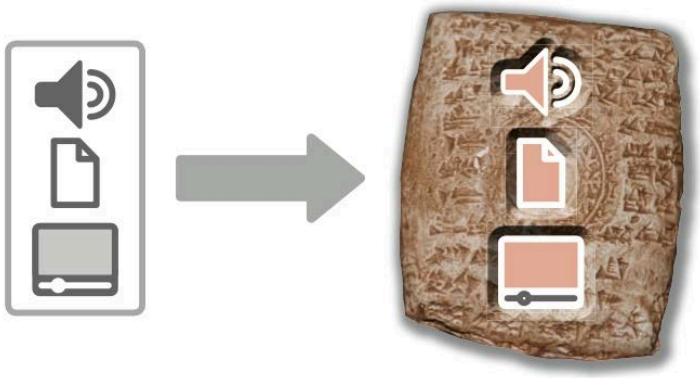
I servizi di base che ogni centro deve assicurare

- documentare le risorse linguistiche e depositare la loro descrizione tramite descrittori armonizzati, metadati
- riversare le descrizioni dei metadati nel catalogo centrale (tramite protocollo di harvesting)
- reperire dati e strumenti da parte degli studiosi di ogni parte del mondo
- accedere alle informazioni relative a dati e strumenti di interesse per la ricerca dello studioso



CLARIN: alcune risposte

- depositare in maniera sostenibile corpora, lessici, tracciati audio/video, annotazioni, grammatiche, modelli linguistici
- archiviare a lungo termine
- descrivere dati in maniera univoca e ricercabile attraverso l'infrastruttura
- attribuire un identificativo persistente
- proteggere le risorse attraverso un sistema di autenticazione controllato
- integrare le risorse nella infrastruttura e combinarle con strumenti linguistici



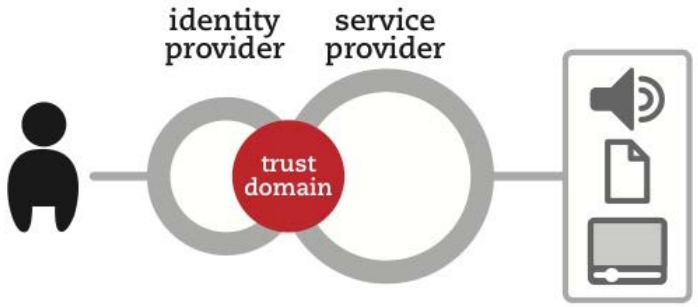
CLARIN: alcune risposte

- Scoprire le risorse linguistiche attraverso una ricerca libera oppure facilitata tramite il sistema a faccette



CLARIN: alcune risposte

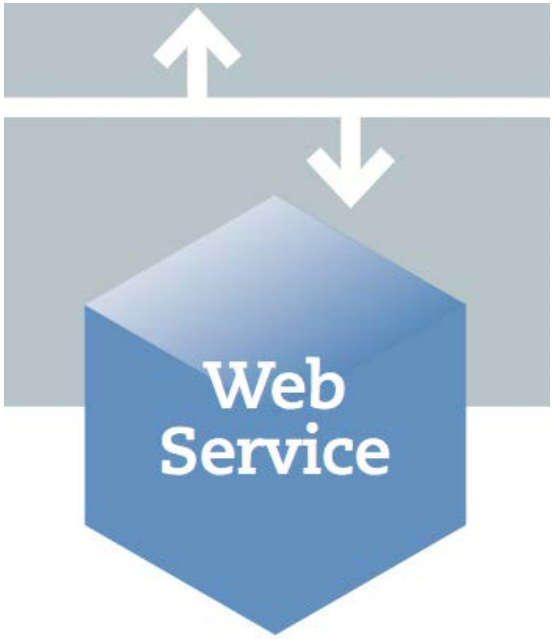
- Avere accesso a risorse protette o disponibili solo tramite autenticazione semplice e unificata mediante il proprio identificativo istituzionale





CLARIN: alcune risposte

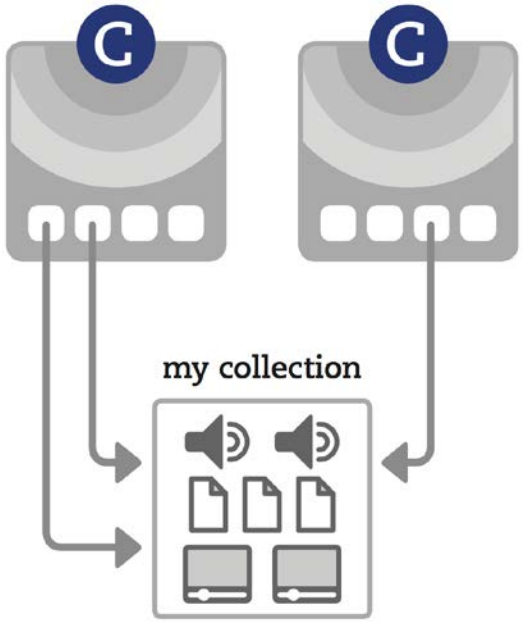
- Esplorare ed analizzare dati linguistici con un'ampia varietà di strumenti linguistici forniti incapsulati sottoforma di servizi per i quali non c'è necessità di competenze tecniche
- Uno di questi è WebLicht fornito dal centro tedesco per l'annotazione di corpora. Offre strumenti come:
  - Tokenizzatori
  - Annotatori per parti del discorso
  - Annotatori sintattici
- che possono essere combinati a seconda delle esigenze dell'utente e le risposte visualizzate in maniera user/friendly



CLARIN: alcune risposte

- Creare collezioni virtuali di oggetti, dati e strumenti che provengono da centri diversi di interesse per uno studioso

  - *una collezione di link che vengono menzionati in una serie articoli, oppure*
  - *una serie di file audio che vengono utilizzati in un libro che parla di lingue a rischio di estinzione…)*



Come funziona un CLARIN nazionale



I CLARIN nazionali sono organizzati in consorzi supportati finanziariamente dalla propria nazione grazie a progetti strategici nazionali:

- raggruppano varie istituzioni (centri di ricerca, università, biblioteche, istituzioni di supporto alla ricerca…)
- forniscono dati, strumenti e servizi

CLARIN-ERIC: i pilastri

https://www.clarin.eu/content/governance

General Assembly: responsabili designati dal ministero di ogni nazione. Massimo organismo decisionale.



Clarin office + **Board of directors**: gestione esecutiva



National coordinators' forum: capi dei consorzi nazionali + **Standing committee for technical centers**: integrazione e interoperabilità a livello trans-nazionale



CLARIN Nazionale in Italia: CLARIN-IT



CLARIN-IT		
Università di Siena	archivi orali	Silvia Calamai
Scuola Normale Superiore	archivi orali	Pier Marco Bertinetto
EURAC Bolzano	dati e strumenti per le lingue regionali	Andrea Abel
FBK Trento	strumenti per applicazioni di NLP	Bernardo Magnini e Sara Tonelli
Università Cattolica Milano	strumenti per le lingue classiche	Marco Passarotti
Università Ca' Foscari	annotazione tematica in DH	Paolo Mastandrea
Università di Parma	edizioni digitali per il greco	Anika Nicolosi
Università di Pisa	dati e strumenti per NLP	Alessandro Lenci
Università di Roma	ontologie per DH	Fabio Ciotti e Daniele Silvi

Concretamente…



- Uno studioso vuole fare una ricerca sulla sintassi
- È interessato a studiare le lingue classiche
- Cerca lessici e dati annotati sintatticamente (tree-bank), strumenti di annotazione, visualizzatori …

Concretamente…

…per esempio per il latino…

A Valency Lexicon for Latin extracted from the Latin Dependency Treebank

1

⊟ This resource is a corpus-based valency lexicon for Latin verbs, automatically extracted from the Latin Dependency Treebank. It contains the full list of patterns of arguments occurring with each verbal form in the Latin Dependency Treebank, with syntactic information on the dependency role of the arguments, morphological information on the case/mood of the arguments, and lexical information on the lemmas of the arguments. It also contains metadata specifying the authors of the works where the verbal forms were found.

Name	• A Valency Lexicon for Latin extracted from the Latin Dependency Treebank
Collection	• Oxford Text Archive
Language	• Latin
Subject	• text_and_corpus_linguistics

More details...

Resources

2546.xml

Concretamente…

…oppure…

Index Thomisticus Treebank

The Index Thomisticus Treebank is the syntactically annotated subset of the Index Thomisticus corpus. The Index Thomisticus corpus features the opera omnia by Thomas Aquinas in Latin, for a total of around 11 million words. The Index Thomisticus Treebank is a dependency-based treebank, whose annotation style resembles that of the Prague Dependency Treebank of Czech

Name	• Index Thomisticus Treebank
Collection	• Tübingen Curated Resource
Language	• Latin
Modality	• written
Country	• Italy
Genre	• historical corpus • treebank • syntactically annotated corpus
Organisation	• Università Cattolica del Sacro Cuore, Milan (Italy)
National project	• CLARIN-D

More details...

Resources

0000-0001-3338-7

Concretamente…



http://weblicht.sfs.uni-tuebingen.de/weblichtwiki/index.php/Main_Page

http://weblicht.sfs.uni-tuebingen.de/weblichtwiki/index.php/Tundra

Concretamente…



WEBLICHT

Search Go Search

History View source Discussion Page

User

- Getting Started
- Use Cases and Examples
- User Manual
- FAQ

Developer

- Developer Manual

Associated Applications

- WebLicht
- Annotation Viewer
- TViewer
- TüNDRA - the Tübingen aNnotated Data Retrieval Application
- FCS Aggregator
- CiNaViz - Visualization of European City Names

Main Page

Welcome to WebLicht!

WebLicht is an execution environment for automatic annotation of text corpora. Linguistic tools such as tokenizers, part of speech taggers, and parsers are encapsulated as web services, which can be combined by the user into custom processing chains. The resulting annotations can then be visualized in an appropriate way, such as in a table or tree format.

Many of the tools incorporated into WebLicht have existed as command-line or desktop tools for many years and are well know within the linguistic community. Others have been developed specifically for use with WebLicht. By making these tools available on the web and by use of a common data format for storing the annotations, WebLicht provides a way to combine them into processing chains.

It is often the case that a tool relies on one or more annotation layers to exist in the input data before it can begin processing. WebLicht assists the user in building a processing chain by determining which tools are able to operate on the data at any given point and offering only those tools as options to add to the chain.

Development of WebLicht started in October 2008 as part of the D-SPIN project, the predecessor project of CLARIN-D. Further development and enhancement of WebLicht is an important goal of CLARIN-D, aiming to make WebLicht a fully-functional virtual research environment.

Tools

Tools and processing tasks are described in more detail in the User Manual under Tools in Detail.

Concretamente…



WEBLICHT

Search

History

User

- Getting Started
- Use Cases and Examples
- User Manual
- FAQ

Developer

- Developer Manual

Associated Applications

- WebLicht
- Annotation Viewer
- TViewer
- TüNDRA - the Tübingen aNnotated Data Retrieval Application

Tundra

TüNDRA - the *Tübingen aNnotated Data Retrieval Application* - is a treebank search application based in part on the popular but no longer

Using TüNDRA

TüNDRA is available via the WebLicht website at https://weblicht.sfs.uni-tuebingen.de/Tundra. Login is with any valid WebLicht acces

TüNDRA can be used to search the following treebanks:

- TüBa-D/Z treebank of German
- TüBa-D/S treebank of spoken German
- TüBa-E/S treebank of spoken English
- TüBa-J/S treebank of spoken Japanese
- HPSG-based Syntactic Treebank of Bulgarian
- Index Thomisticus Treebank

Concretamente…



LINDAT CLARIN | Repository | TreeQuery | Treex | More Apps | Events | About | CLARIN

PML Tree Query

Tool for searching and browsing treebanks online

Browse Treebanks | Login

Filter treebanks …

Show only publicly accessible treebanks

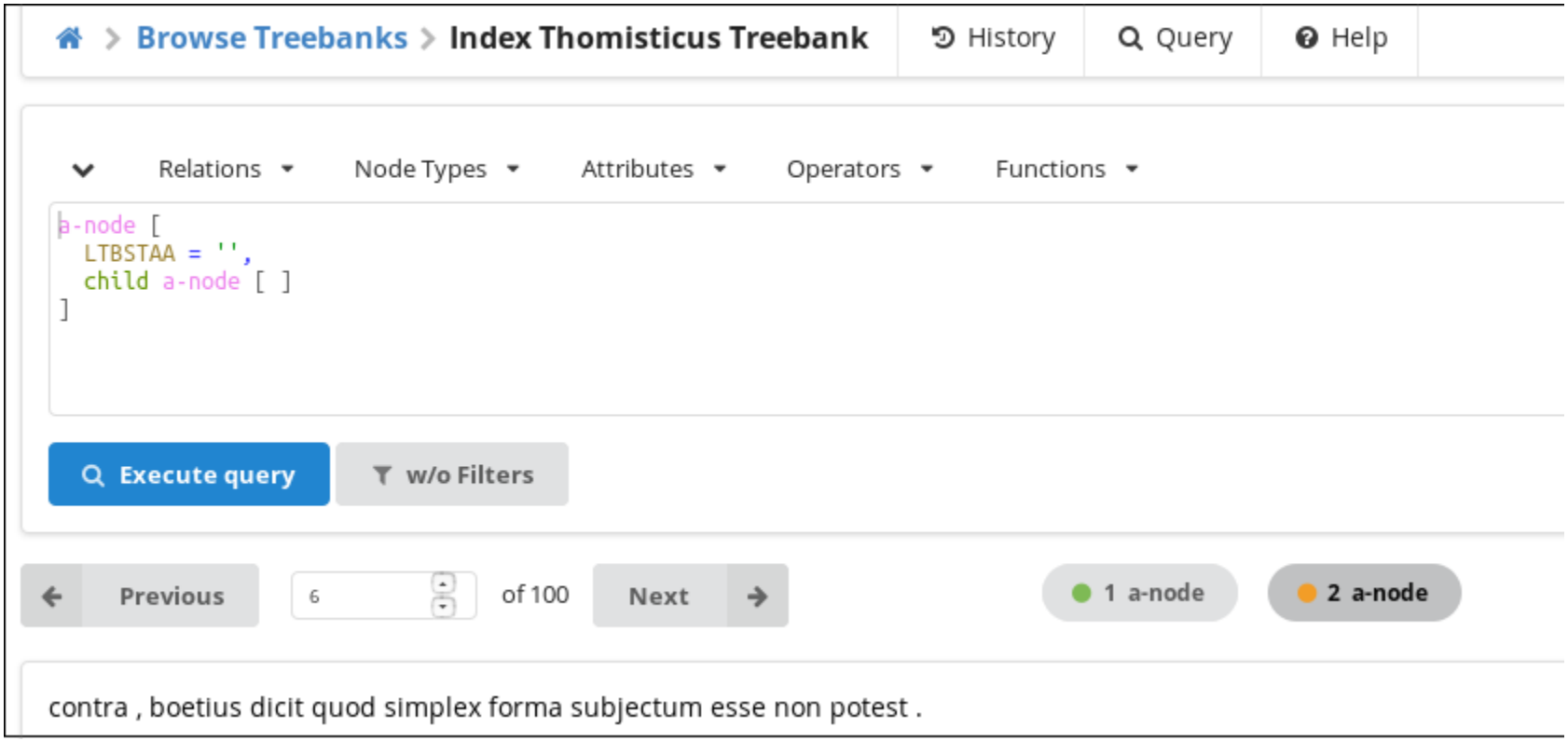
LANGUAGES:

Ancient Greek 3 | Arabic 2 | Basque 1 | Bengali 0 | Bulgarian 2 | Catalan 1 | Chinese 1
Croatian 1 | Czech 9 | Danish 1 | Dutch 3 | English 7 | Estonian 2 | Finnish 2 | French 1
Galician 1 | German 1 | Gothic 1 | Greek 1 | Hebrew 1 | Hindi 1 | Hungarian 1 | Indonesian 1
Irish 1 | Italian 1 | Japanese 0 | Kazakh 1 | Latin 7 | Latvian 1 | Norwegian 1 | Old Church Slavonic 1
Persian 2 | Polish 2 | Portuguese 3 | Romanian 2 | Russian 2 | Slovak 0 | Slovenian 4 | Spanish 2
Swedish 3 | Tamil 2 | Telugu 0 | Turkish 1

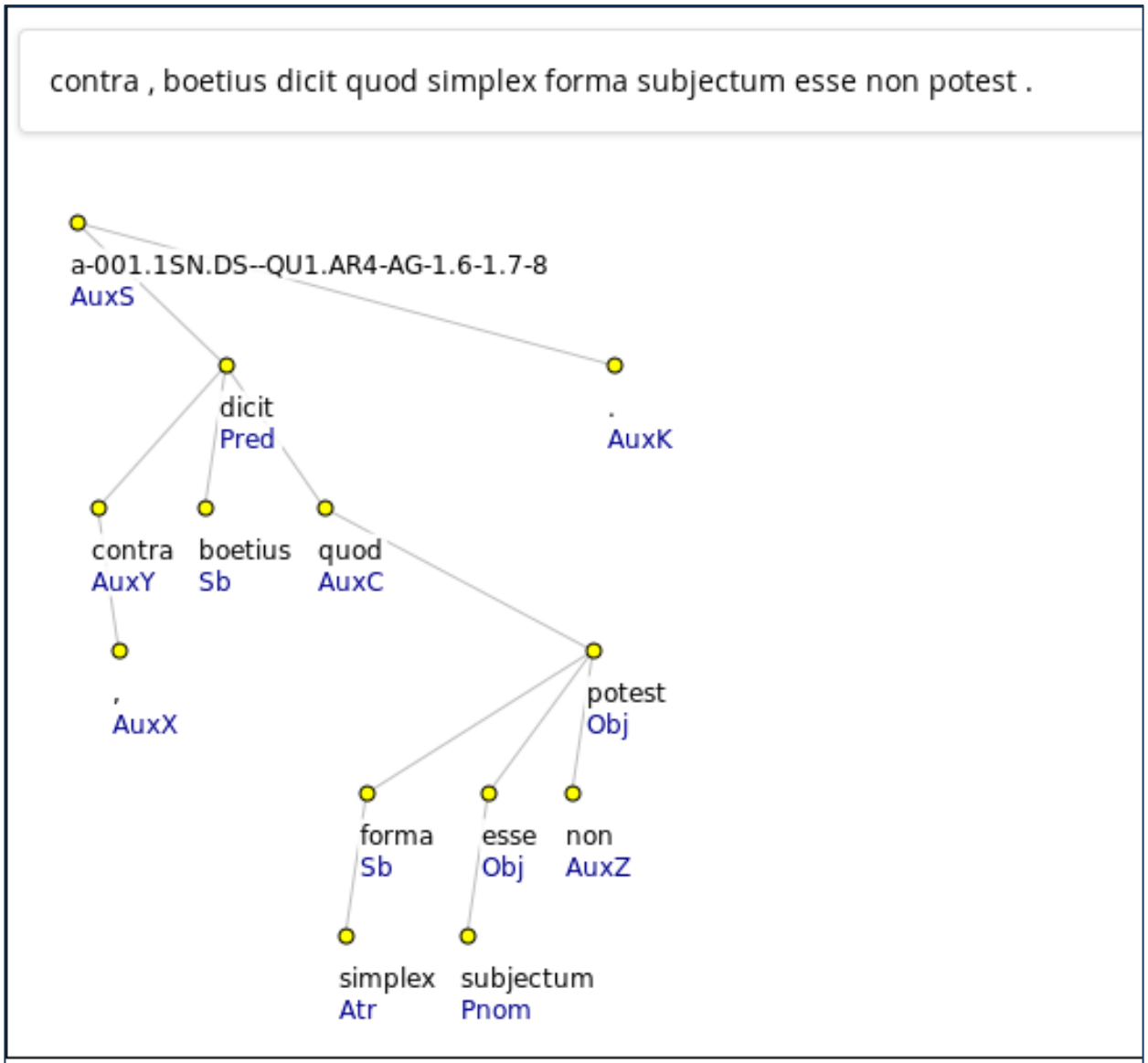
TAGS: CoNLL 1 | HamleDT 13 | PDT 2 | Penn Treebank 4 | Treex 1 | Universal Dependencies 52

Reset filter

Concretamente…



Concretamente…



Concretamente…



LINDAT CLARIN | Repository | TreeQuery | Treex | More Apps | Events | About | CLARIN

PML Tree Query

Tool for searching and browsing treebanks online

Browse Treebanks | Login

Filter treebanks …

Show only publicly accessible treebanks

LANGUAGES:
Ancient Greek 3 | Arabic 2 | Basque 1 | Bengali 0 | Bulgarian 2 | Catalan 1 | Chinese 1
Croatian 1 | Czech 9 | Danish 1 | Dutch 3 | English 7 | Estonian 2 | Finnish 2 | French 1
Galician 1 | German 1 | Gothic 1 | Greek 1 | Hebrew 1 | Hindi 1 | Hungarian 1 | Indonesian 1
Irish 1 | Italian 1 | Japanese 0 | Kazakh 1 | Latin 7 | Latvian 1 | Norwegian 1 | Old Church Slavonic 1
Persian 2 | Polish 2 | Portuguese 3 | Romanian 2 | Russian 2 | Slovak 0 | Slovenian 4 | Spanish 2
Swedish 3 | Tamil 2 | Telugu 0 | Turkish 1

TAGS: CoNLL 1 | HamleDT 13 | PDT 2 | Penn Treebank 4 | Treex 1 | Universal Dependencies 52

Reset filter

Concretamente…



Browse Treebanks > Universal Dependencies - Ancient Greek | History | Query | Help | Get data | Monica Monachini

Relations | Node Types | Attributes | Operators | Functions | Timeout 30 | Result limit 100

```
a-node [
  afun = '',
  child a-node [ ]
]
```

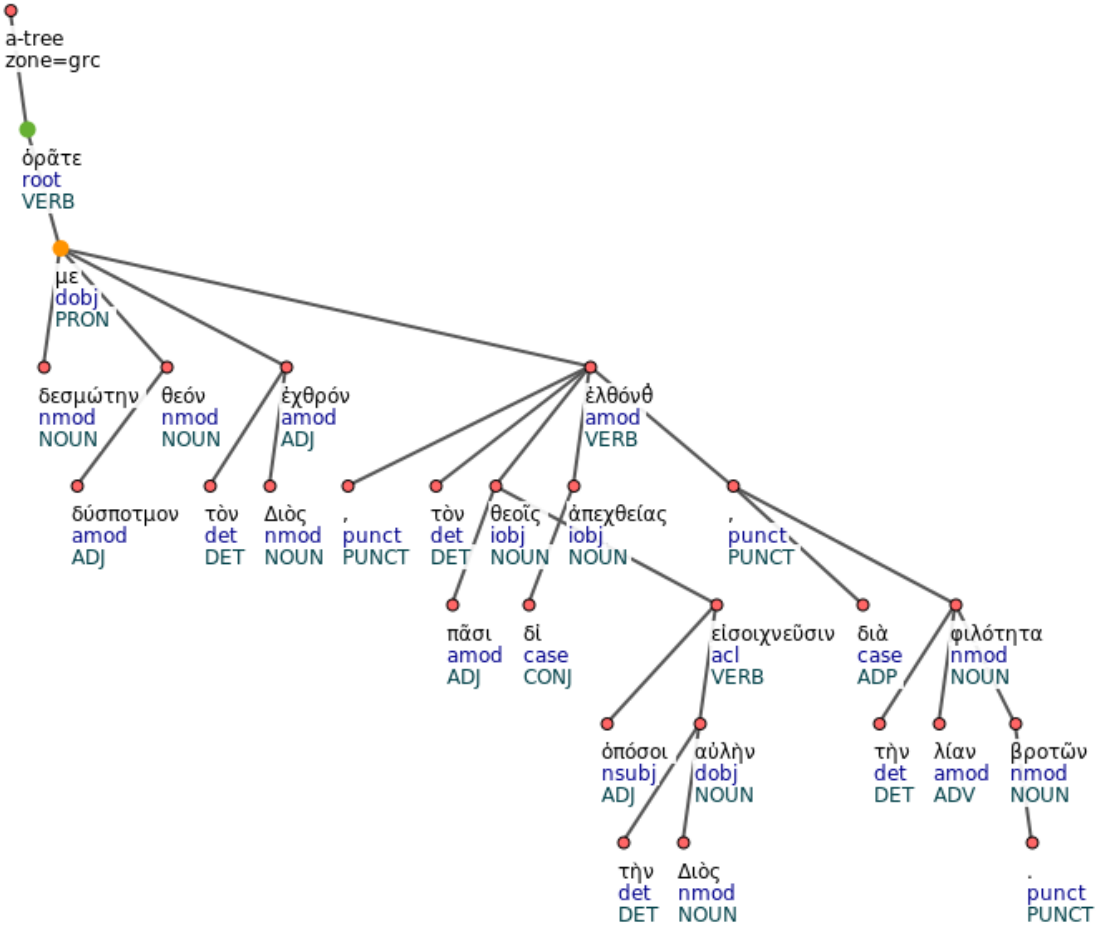
Execute query | w/o Filters | Permanent link | Clear

Previous | 1 | of 100 | 1 a-node | 2 a-node | 1 / 50 | dev-001-003.treex.gz | Next

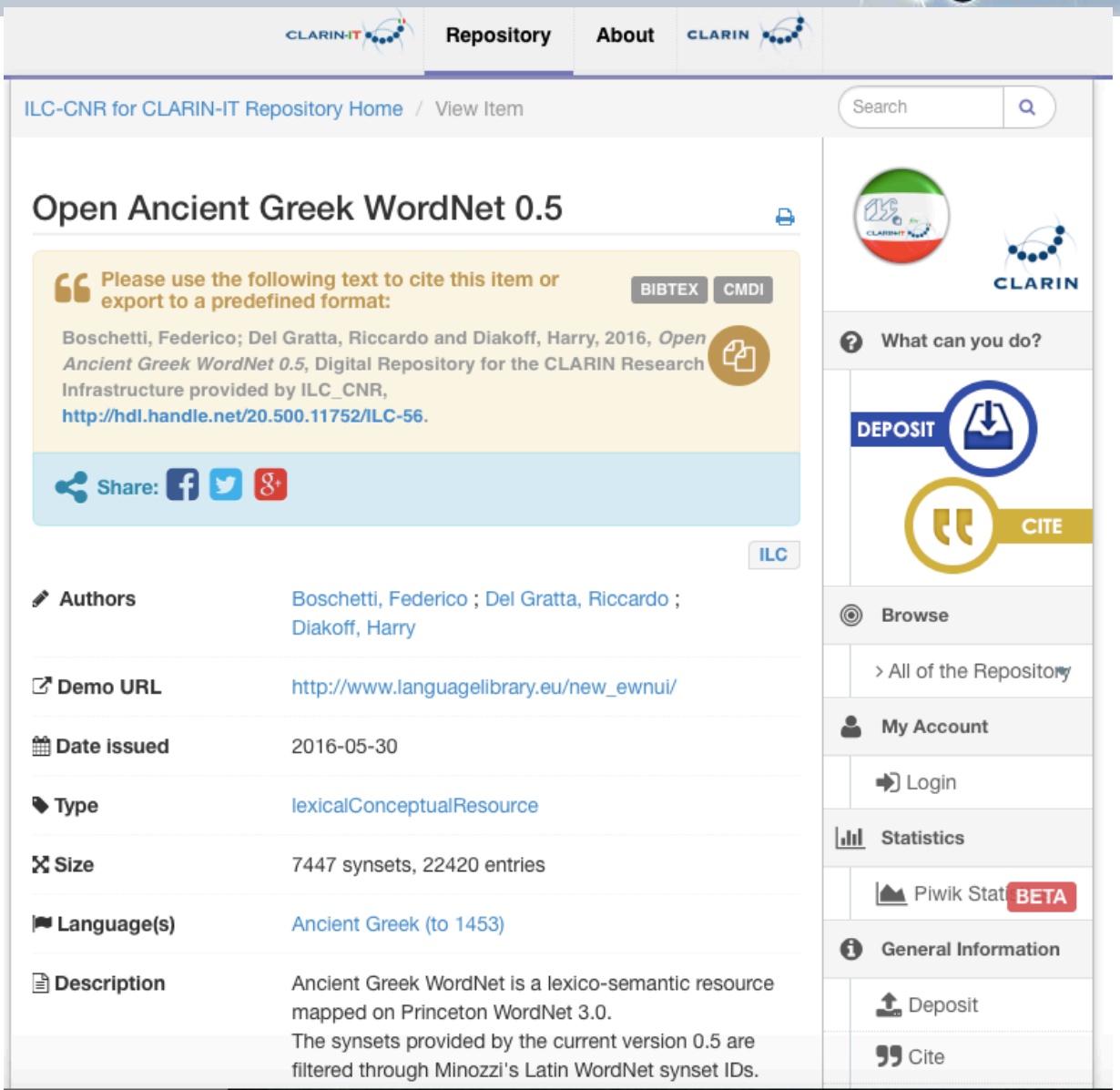
[grc] ὁρᾶτε δεσμώτην με δύσποτμον θεόν τὸν Διὸς ἐχθρόν , τὸν πᾶσι θεοῖς δί ἀπεχθείας ἐλθόνθ ὁπόσοι τὴν Διὸς αὐλὴν εἰσοιχνεῦσιν , διὰ τὴν λίαν φιλότητα βροτῶν .

Concretamente…

[grc] ὁρᾶτε δεσμώτην με δύσποτμον θεόν τὸν Διὸς ἐχθρόν , τὸν πᾶσι θεοῖς δί ἀπεχθείας ἐλθόνθ ὁπόσοι τὴν Διὸς αὐλὴν εἰσοιχνεῦσιν , διὰ τὴν λίαν φιλότητα βροτῶν .



Concretamente…





User Login



Search Panel

Specify the input language, type the word or the synset to search and at least one output language

Select the input language

- English
- Greek
- Italian
- Latin
- Arabic
- Croatian

Type the word or the synsetid

Type a word

Word

Synsetid

Select the target languages

- Greek
- Latin
- Italian
- Arabic
- Croatian
- English

Select Domain Resources

- Italian Geodomain
- English Geodomain
- Mariterm

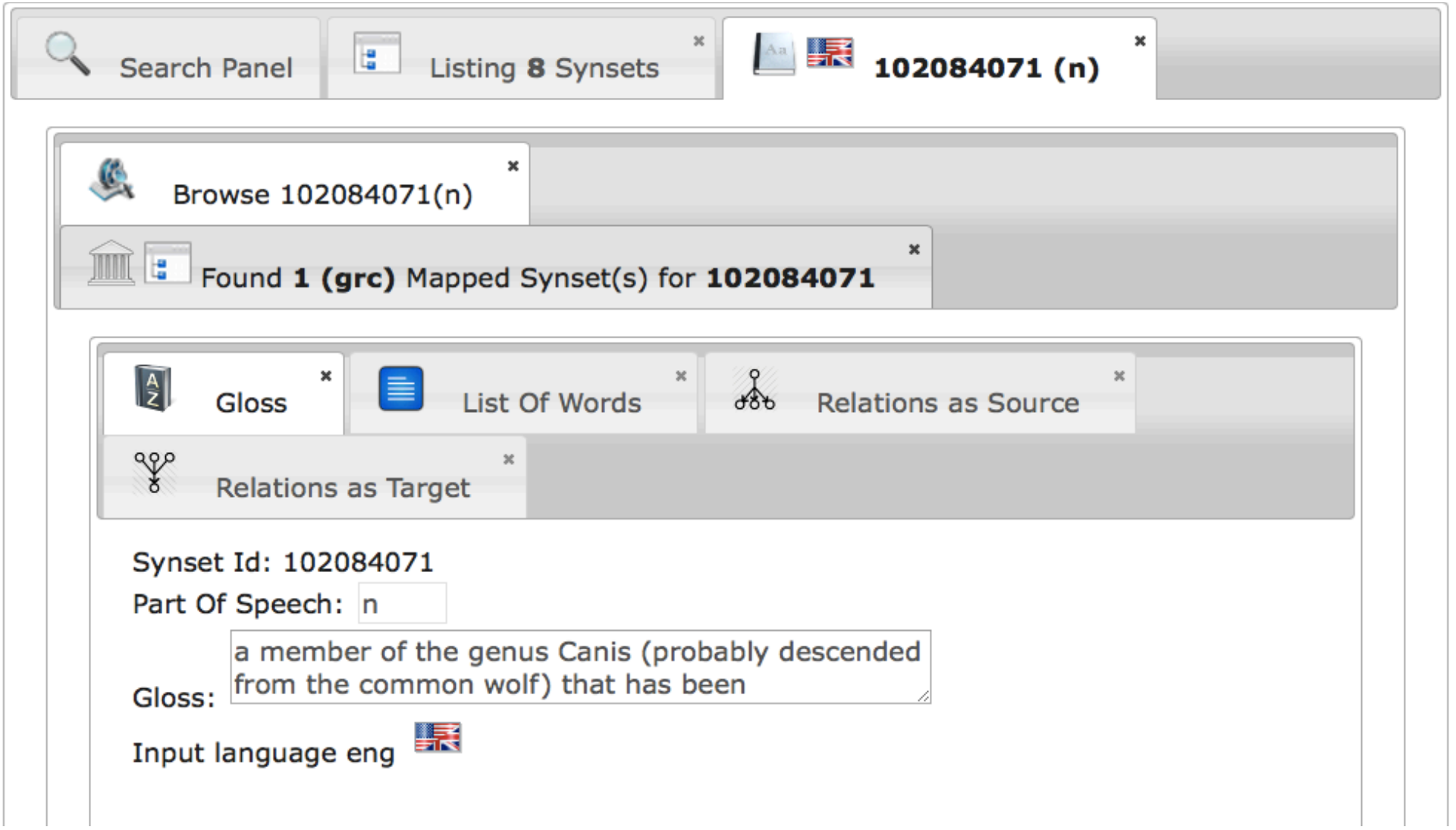
View Results | Reset Form | New Search

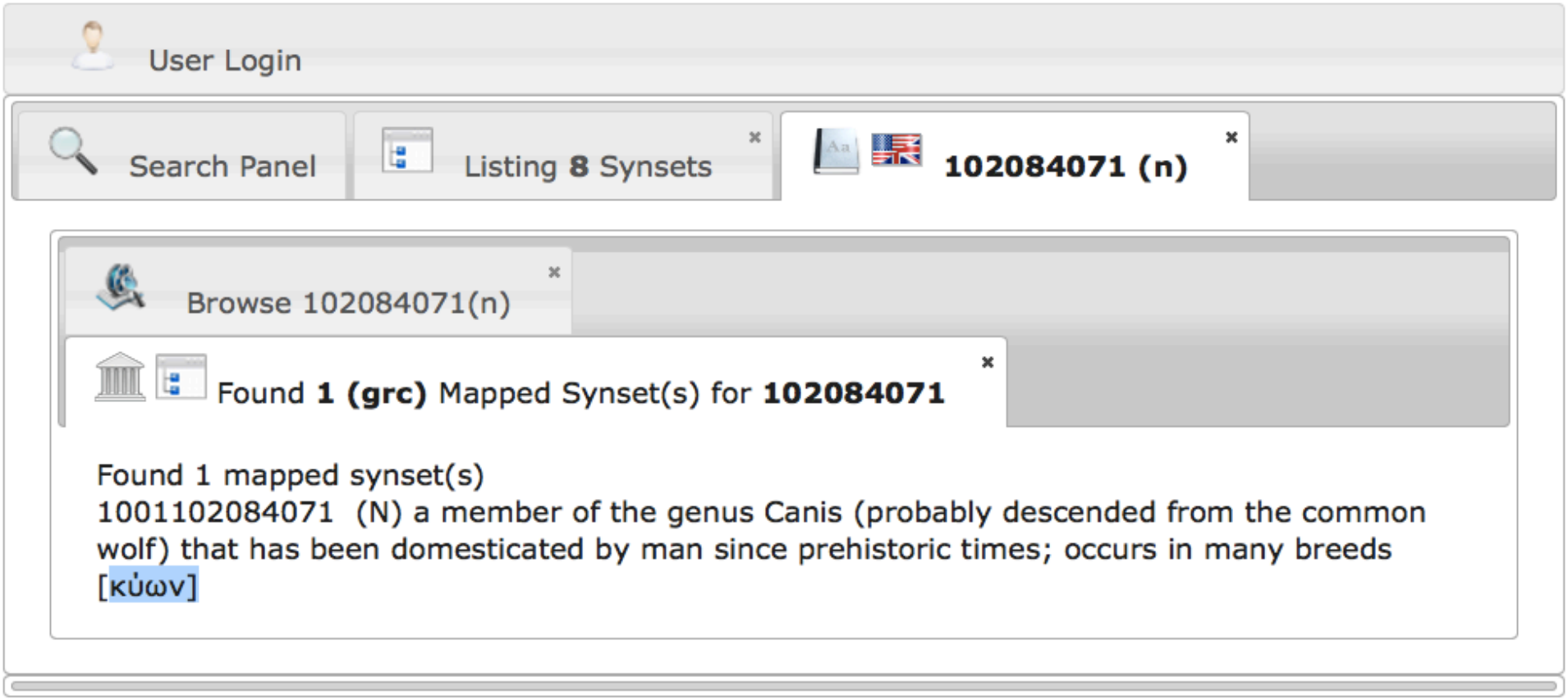
User Login

Search Panel | Listing **8** Synsets

Found 8 synsets containing **dog** and input language: **eng**

102084071 (n) a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds [canis familiaris, dog, domestic dog]
102710044 (n) metal supports for logs in a fireplace [andiron, dog, dog-iron, firedog]
103901548 (n) a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward [click, detent, dog, pawl]
107676602 (n) a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll [dog, frank, frankfurter, hot dog, hotdog, ...]
109886220 (n) someone who is morally reprehensible [blackguard, bounder, cad, dog, heel, ...]
110023039 (n) informal term for a man [dog]
110114209 (n) a dull unattractive unpleasant girl or woman [dog, frump]
202001858 (v) go after with the intent to catch [chase, chase after, dog, give chase, go after, ...]


Search Panel
Listing 8 Synsets
102084071 (n)
Browse 102084071(n)
Found 1 (grc) Mapped Synset(s) for 102084071
Gloss
List Of Words
Relations as Source
Relations as Target
Synset Id: 102084071
Part Of Speech: n
Gloss: a member of the genus Canis (probably descended from the common wolf) that has been
Input language eng


User Login
Search Panel
Listing 8 Synsets
102084071 (n)
Browse 102084071(n)
Found 1 (grc) Mapped Synset(s) for 102084071
Found 1 mapped synset(s)
1001102084071 (N) a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds
[κύων]

Questa dove va?



Tale tipologia si sviluppa nelle IR dell'area delle scienze biomediche che si caratterizzano spesso per la loro natura distribuita, multi-centrica, basata su sistemi distribuiti
territorialmente che richiedono non solo un mantenimento e operatività di lungo periodo, ma anche
e soprattutto un continuo aggiornamento,

Le infrastrutture di ricerca, infatti, non possono prescindere da un'impostazione
di accesso aperto (ed in una certa misura, gratuito) dell'infrastruttura per lo svolgimento della ricerca di base. Ciò non vuol dire che una infrastruttura
di ricerca non possa operare sul mercato (vendere beni e servizi, ad esempio, applicare user charges per singoli utenti)

Riepilogando … le infrastrutture



mettono a disposizione di ricercatori e studiosi nel settore del patrimonio culturale digitale:

- un accesso online semplice e controllato a servizi, risorse e strumenti di collaborazione,
- potenza delle tecnologie ICT per l'elaborazione di calcolo,
- deposito e conservazione a lungo termine
- strumentazione
- servizi avanzati
- strumento per coordinare il lavoro nelle arti e scienze umane attraverso la creazione di un "ambiente" in cui condividere tecnologie digitali, dati e metodologie innovative.
- modo per integrare a livello europeo iniziative già presenti a livello nazionale, come centri di ricerca, progetti, reti e archivi di dati.

Riepilogando … le infrastrutture

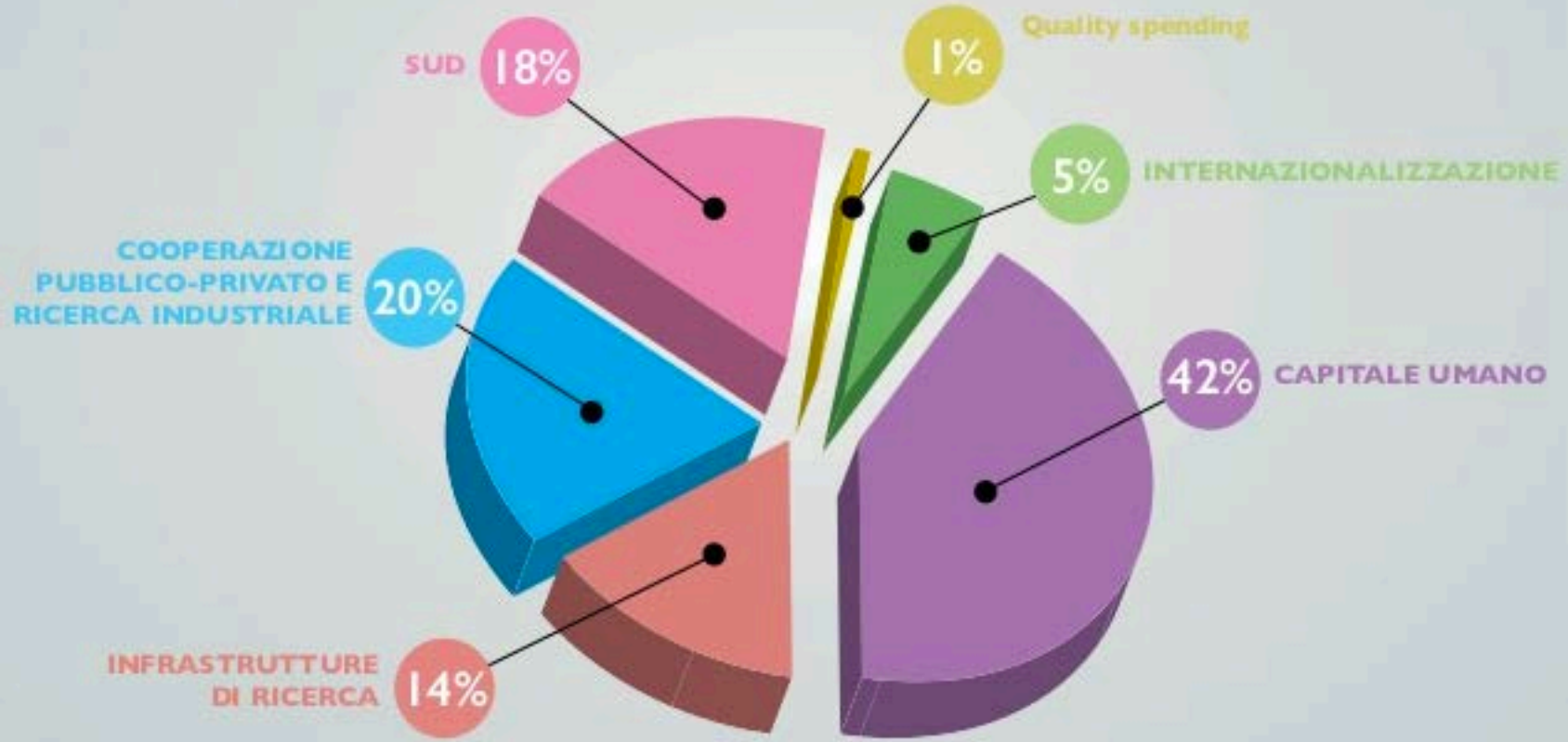


Piattaforma di servizi stabile e persistente:

- rende facilmente accessibili risorse e tecnologie per le lingue di tutti gli stati membri ma anche altre lingue studiate in Europa
- da accesso al contenuto multilingue ma anche multiculturale
- aiuta a conservare e sostenere il patrimonio multiculturale
- facilita la creazione di un nuovo paradigma di sviluppo collaborativo
- permette la riusabilità dei dati e degli strumenti software, di aggregare contributi e garantire nuovi servizi
- fornisce strumenti e soluzioni nuove a problemi esistenti
- assicura la formazione di nuove figure professionali
- fornisce la consulenze necessarie per personalizzare le risorse al fine soddisfare le necessità della comunità umanistica.

Le risorse: distribuzione sui Programmi

2.428,6 milioni €



Il volume dei testi scritti e parlati e del materiale audiovisivo è enorme ed è in crescita esponenziale. dimensioni di questo materiale rendono

indispensabile l’utilizzo di sistemi informatici da parte delle aree umanistiche che fanno linguistico (scritto o parlato o

Attualmente circa 179 istituzioni di 33