

Firenze

...avanzarono. Per la qual cosa con quelli, successivamente or l'uno or l'altro a Firenze tornandosi, gran parte delle loro possessioni ricompararono e molte dell'altre...

1 OF 7 ANNOTATIONS JUMP TO TEXT

geonames 3176959 Florence

Recogito

Tiziana Mancinelli, Simone Rebora, Rachele Sprugnoli

ASSOCIAZIONE PER
L'INFORMATICA UMANISTICA
E LA CULTURA DIGITALE



What is Recogito?

- Free online **annotation tool** and **publishing** platform
- Semi-automatic annotation of place references in maps and texts. It was developed as part of the Pelagios 3 research project
- Semantic annotation Linked Open Data
- Allow to create collaborative project and simultaneous annotations of the same document
- First developed in 2014, the community-oriented philosophy behind **Pelagios Commons** has made users an active agent in shaping its functionality and interface.

What is Recogito?

- Recogito allows people to create semantic annotation without worrying too much about the formal language and structure of Linked Open Data.

Why? What is a semantic annotation?

Producing semantic annotations usually requires a certain amount of knowledge of digital technologies such as RDF, ontologies and/or text encoding. These techniques can sometimes act as a barrier for users that are not already familiar with Semantic Web theory.

Recogito motto:

“Linked Data annotation without the pointy brackets”

Linked (Geo) Data Annotation

Definitions

- LINKED DATA

Connecting structured data on the Web / Linked Data is simply about using the Web to create typed links between data from different sources. (Bizer et al., 2011)

- LINKED GEO-DATA:

The text annotation area to demarcate place names in digital text, an image annotation area to mark up and transcribe place names on map or manuscript scans, and a geo-resolution area, where the identified (and transcribed) place names are mapped to a gazetteer (and, thus, to geographical coordinates).

Gazeeters

What is a gazeteer? What are the gazeeters in Recogito?

A gazetteer is a geographical dictionary or directory used in conjunction with a map or atlas. It typically contains information concerning the geographical makeup, social statistics and physical features of a country, region, or continent (...) - Wikipedia

- ✓ **Pleiades** Pleiades Gazetteer of the Ancient World
- ✓ **CHGIS** China Historical GIS
- ✓ **DPP Places** Places from the Digitizing Patterns of Power project
- ✓ **DARE** Digital Atlas of the Roman Empire
- ✓ **GeoNames** A subset of GeoNames populated places, countries and first-level administrative divisions
- ✓ **MoEML** Map of Early Modern London
- ✓ **HGIS de las Indias** Historical-Geographic Information System for Spanish America (1701-1808)



Example: <https://sites.fas.harvard.edu/~chgjs/>

Connecting Knowledge

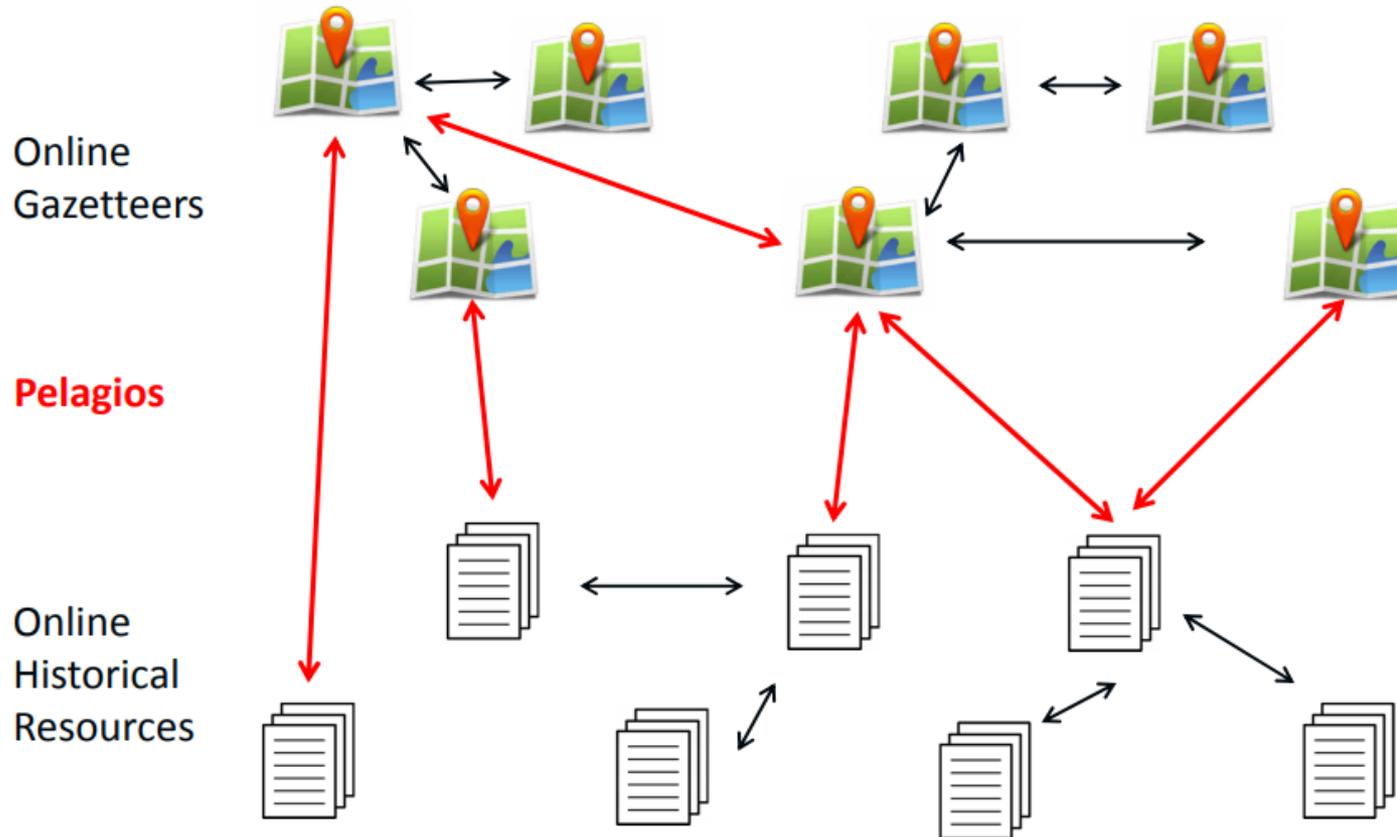
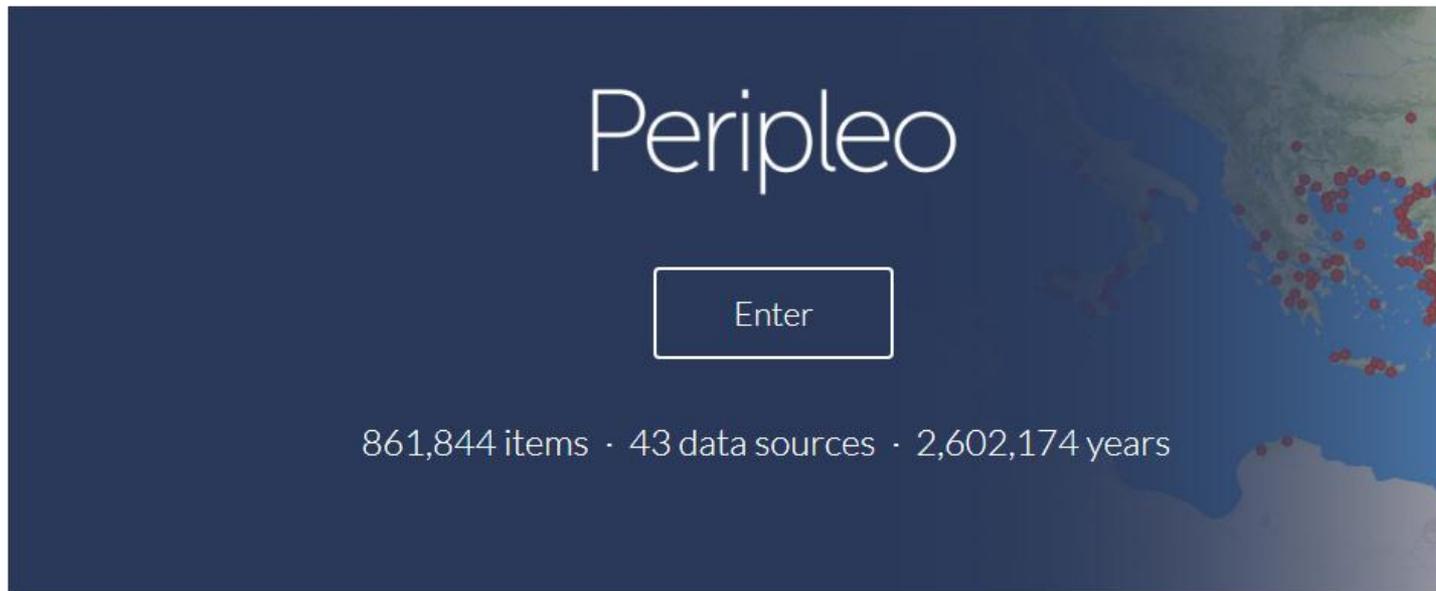


Image taken from: <https://f-origin.hypotheses.org/wp-content/blogs.dir/2899/files/2017/04/2017-04-Recogito-Paris.pdf>

From Annotation to Exploration

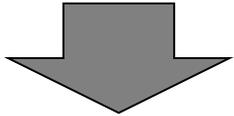
- From Recogito to Peripleo: <https://peripleo.pelagios.org/>
- Example “Herodotus: The Histories”:
<https://recogito.pelagios.org/document/tjrrsqn4dwmgep>



How Recogito Works

Step by step with Recogito

- Uploading a file - TXT (it is the only choice)



The screenshot displays the Recogito user interface. On the left sidebar, the user profile 'Tizzja' is shown with the text 'Joined on 21 Aug 2018'. Below the profile, a blue button with a white plus sign and the word 'New' is circled in red. Other sidebar options include 'My Documents', 'Shared with me', 'Storage' (0.34 of 200 MB used), 'Activity' (13 edits), and 'My top collaborators' (with a collaborator 'R' visible).

The main content area features a search bar labeled 'Search Recogito...'. Below it, a section titled 'Shared with me (1)' contains a table with the following data:

Author	Title	Language	Date	Uploaded at	Last edit	...
Giovanni Boccaccio	Decameron Seconda Giornata	IT		31 Dec 2018	1 week ago	

Annotating Documents

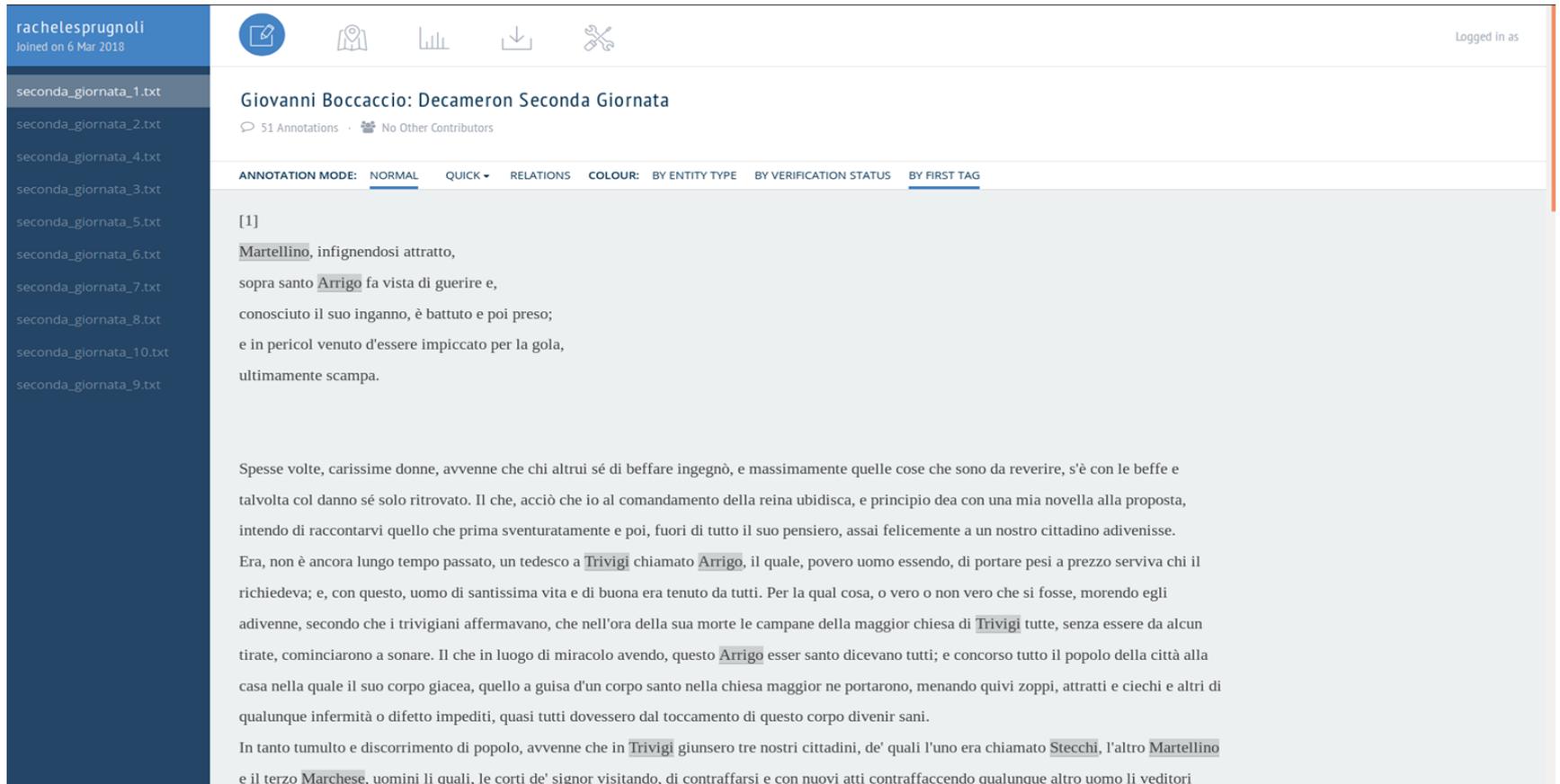
- Automatic annotation - Named Entity Recognition (only EN)

The screenshot displays a user interface for document management. On the left, a sidebar shows the user's profile (Tiziana, joined on 21 Aug 2018) and navigation options like '+ New', 'My Documents', 'Shared with me', 'Storage', and 'Activity'. The main area features a search bar and a document list. A document titled 'Giornata_3.txt' is selected, and a context menu is open over it, listing actions such as 'Open', 'Open in new tab', 'Move to', 'Duplicate', 'Delete', and 'Named Entity Recognition'. The 'Named Entity Recognition' option is circled in red.

Author	Title	Language	Date	Uploaded
	Giornata_3.txt			19 Jan 2019

Step by step with Recogito

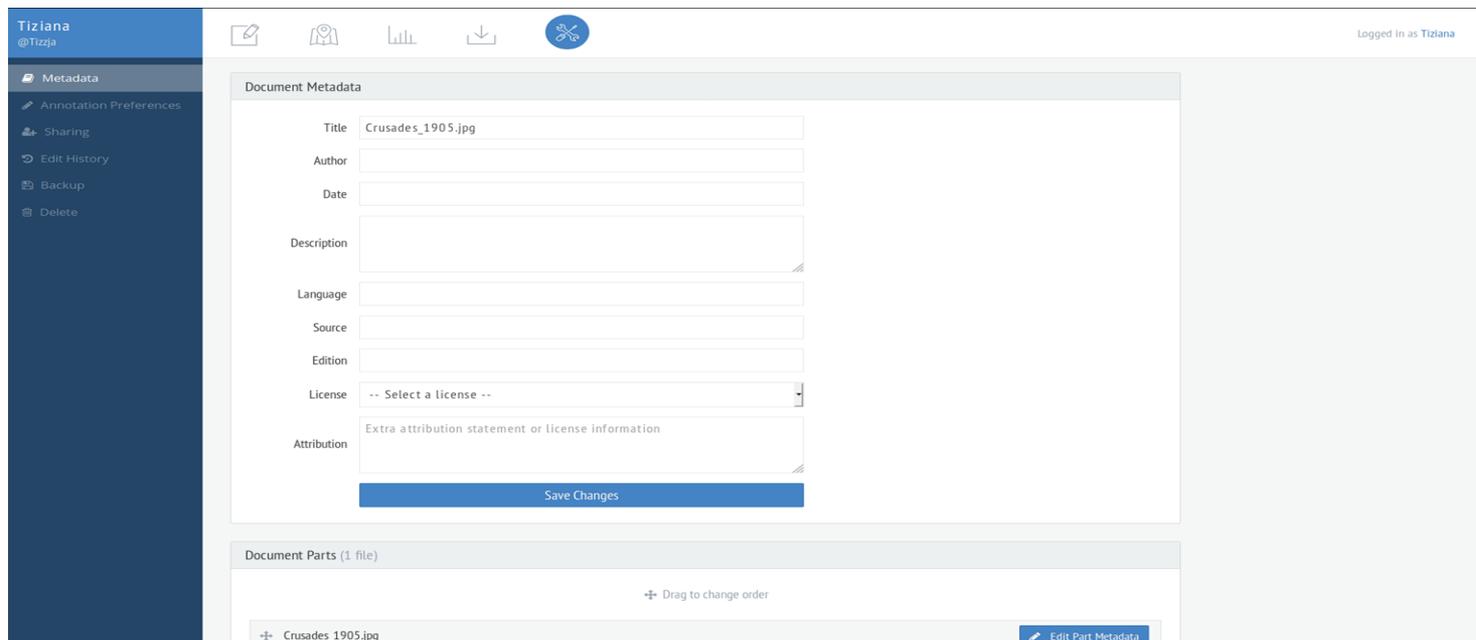
- Double-click on the text and you are in the working space for annotations for text



The screenshot displays the Recogito web interface. On the left, a dark blue sidebar shows the user profile 'rachelesprugnoli' (joined on 6 Mar 2018) and a list of document files named 'seconda_giornata_1.txt' through 'seconda_giornata_10.txt'. The main content area is titled 'Giovanni Boccaccio: Decameron Seconda Giornata' and shows '51 Annotations' and 'No Other Contributors'. Below the title, there is a navigation bar with 'ANNOTATION MODE:' followed by 'NORMAL' (selected), 'QUICK', 'RELATIONS', 'COLOUR', 'BY ENTITY TYPE', 'BY VERIFICATION STATUS', and 'BY FIRST TAG'. The text of the document is displayed with several words highlighted in yellow, indicating annotations: 'Martellino', 'Arrigo', 'Trivigi', 'Stecchi', and 'Marchese'. The text is a passage from the Decameron, describing a scene where a man is executed and his body is used for medicinal purposes.

Adding metadata for your document

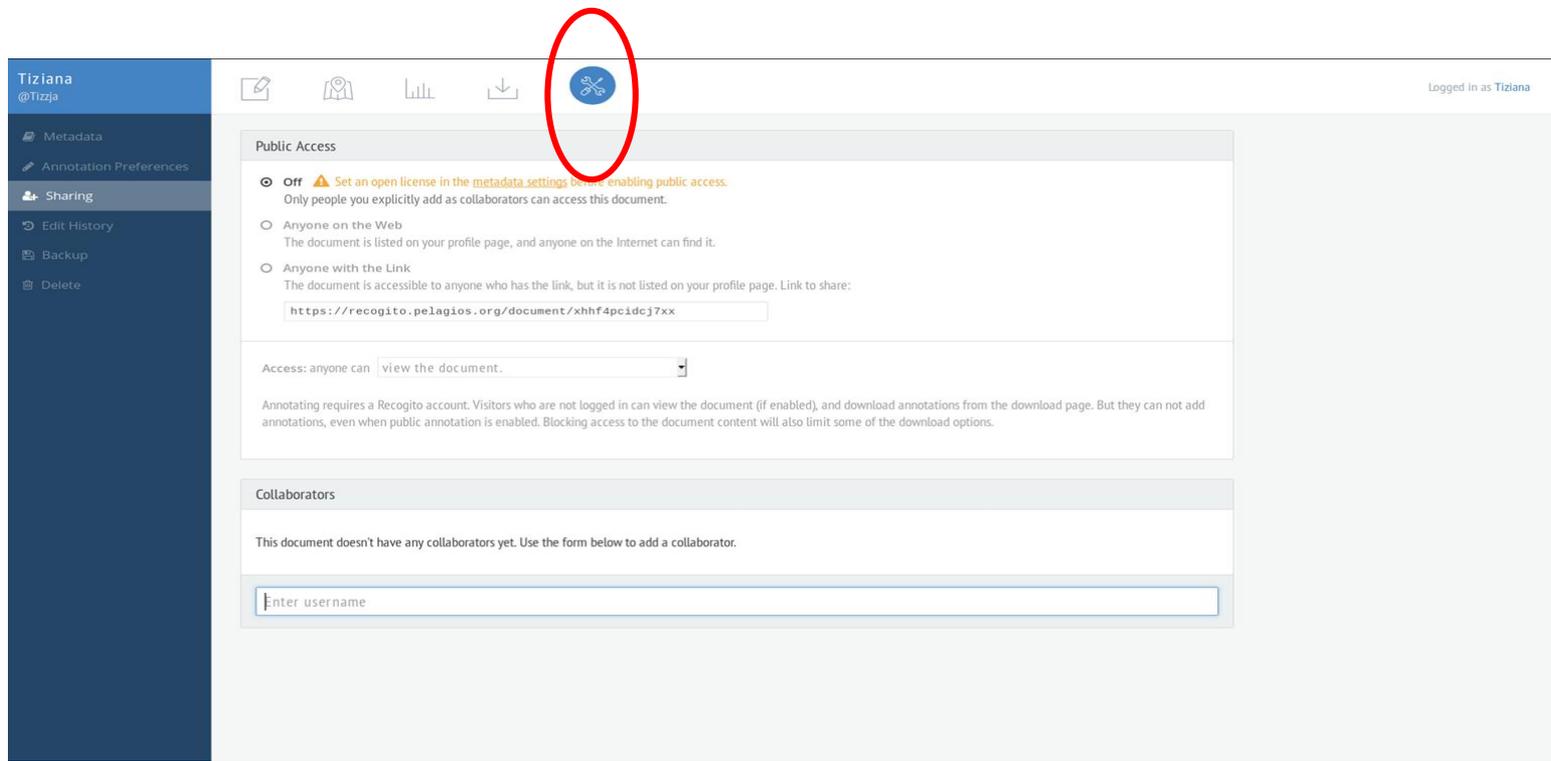
- The new interface does currently have the option to do it:
- Go to “Document Setting”



The screenshot shows a web interface for editing document metadata. On the left is a dark blue sidebar with the user's name 'Tiziana @Tizja' and a menu with options: Metadata, Annotation Preferences, Sharing, Edit History, Backup, and Delete. The main content area is titled 'Document Metadata' and contains several input fields: Title (filled with 'Crusades_1905.jpg'), Author, Date, Description, Language, Source, Edition, License (a dropdown menu showing '-- Select a license --'), and Attribution (with a placeholder text 'Extra attribution statement or license information'). A blue 'Save Changes' button is at the bottom of the form. Below the form is a section for 'Document Parts (1 file)' with a '+ Drag to change order' instruction and a list item for 'Crusades_1905.jpg' with an 'Edit Part Metadata' button.

How to share a document with collaborators

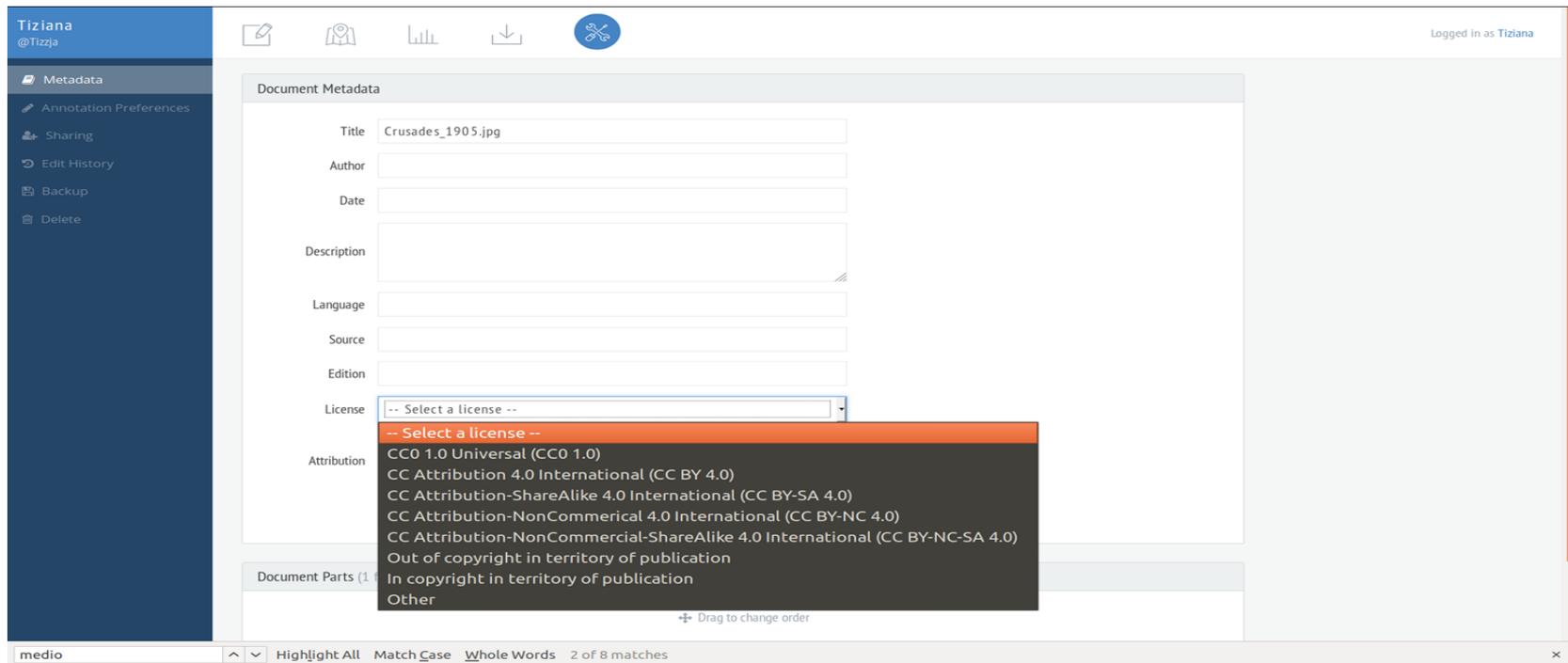
- In “setting document”, go to > sharing > you can also share a document with other collaborators.



The screenshot shows a user interface for document management. On the left is a dark blue sidebar with the user's name 'Tiziana @Tizza' and a menu with options: Metadata, Annotation Preferences, Sharing (highlighted), Edit History, Backup, and Delete. The main content area has a top navigation bar with icons for edit, user, chart, download, and share. The share icon, a blue circle with a white scissors, is circled in red. Below the navigation bar, the 'Public Access' section is visible, showing a radio button for 'Off' (selected) with a warning icon and text: 'Set an open license in the metadata settings before enabling public access. Only people you explicitly add as collaborators can access this document.' Other options include 'Anyone on the Web' and 'Anyone with the Link'. A link to share is provided: 'https://recogito.peLAGIOS.org/document/xhhf4pcidcj7xx'. Below this, there is a dropdown menu for 'Access: anyone can view the document.' and a note about annotation requirements. The 'Collaborators' section below shows a message: 'This document doesn't have any collaborators yet. Use the form below to add a collaborator.' and a text input field labeled 'Enter username'.

How to share a document with collaborators

- By default your document is private (only you can see your work)
To make your document public → go to the metadata



The screenshot displays a web application interface for document management. On the left, a dark blue sidebar contains the user's name 'Tiziana @Tizzja' and a menu with options: 'Metadata', 'Annotation Preferences', 'Sharing', 'Edit History', 'Backup', and 'Delete'. The main content area is titled 'Document Metadata' and contains several input fields: 'Title' (filled with 'Crusades_1905.jpg'), 'Author', 'Date', 'Description', 'Language', 'Source', and 'Edition'. The 'License' field is currently open, showing a dropdown menu with the following options: '-- Select a license --', 'CC0 1.0 Universal (CC0 1.0)', 'CC Attribution 4.0 International (CC BY 4.0)', 'CC Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)', 'CC Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)', 'CC Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)', 'Out of copyright in territory of publication', and 'In copyright in territory of publication'. Below the license dropdown, there is a section for 'Document Parts (1)' with a list containing 'In copyright in territory of publication' and 'Other'. At the bottom of the interface, a search bar shows the word 'medio' and search options: 'Highlight All', 'Match Case', 'Whole Words', and '2 of 8 matches'. The top right corner indicates 'Logged in as Tiziana'.

Annotating Documents

- Annotated manually in the text annotation area
- Recogito also includes a “transcription assistant”, which automatically suggests potential matches based on the proximity to other, already transcribed, toponyms
- Re-apply the same annotation function:
“There are Ntimes other annotated occurrences of Name in the text. Do you want to re-apply this annotation?”

Annotating Documents

- Annotated manually in the text annotation area
Sometime it doesn't really guess what you want!

The screenshot shows a text annotation interface. The user is logged in as Tiziana. The document is titled "Giornata_3.txt". The annotation mode is set to "NORMAL". The text being annotated is:

ogni cosa piena di quei fiori che concedeva il tempo e di frondi, postesi a sedere, venne il discreto siniscalco e loro con preziosissimi confetti e ottimi vini ricevette e riconfortò.

Appresso la qual cosa, fattosi aprire un giardino che di costa era al palagio, in quello, che tutto era da torno murato, se n'entrarono; e parendo loro nella prima entrata di meravigliosa bellezza tutto insieme, più attentamente le parti di quello cominciarono a riguardare. Esso avea dintorno da sé e per lo mezzo in assai parti vie ampissime, tutte diritte come strale e coperte di pergolati di viti, le quali facevano gran vista di dovere quello anno assai uve fare, e tutte allora fiorite sì grande odore per lo giardin rendevano, che, mescolato insieme con quello di molte altre cose che per lo giardino olivano, pareva loro essere tra tutta la spezieria che mai nacque in Oriente. Le latora delle quali vie tutte di rosa' bianchi e vermigli e di gelsomini erano quasi chiuse: per le quali cose, non che la mattina, ma qualche volta da quello, vi si poteva per tutto andare. Quante e quali e come ordina niuna n'è laudevole la quale il nostro aere patisca, di che quivi non che altra cosa che vi fosse ma molto più, era un prato di minutissimi fiori, chiuso dintorno di verdissimi e vivi aranci e di cedri, li quali occhi ma ancora all'odorato facevan piacere. Nel mezzo del quale si so se da natural vena o da artificiosa, per una figura, la quale sopra il cielo, che poi non senza dilettevol suono nella fonte chiarissima sopra soprabondava al pieno della fonte, per occulta via del pratello usciva tutto lo 'ntorniava; e quindi per canaletti simili quasi per ogni parte del giardin discorrea, raccogliendosi ultimamente in una parte dalla quale del bel giardino avea l'uscita, e quindi verso il pian discendendo chiarissima, avanti che a quel divenisse, con grandissima forza e con non piccola utilità del signore due mulina volgea.

Il veder questo giardino, il suo bello ordine, le piante e la fontana co' ruscelletti precedenti da quella tanto piacque a ciascuna donna e a' tre giovani, che tutti cominciarono a affermare che, se Paradiso si potesse in terra fare, non sapevano conoscere che altra forma che quella di quel giardino gli si potesse dare, né pensare, oltre a questo, qual bellezza gli si potesse agiugnere. Andando adunque contentissimi dintorno per quello, facendosi di vari

The interface shows a modal window for manual annotation of the word "Oriente". The modal has tabs for "Place", "Person", and "Event". The "Place" tab is active, showing a map of San Rafael Oriente with geonames 3583379. The modal also includes an "Automatic Match" warning, a comment field, a tag field, and buttons for "Cancel", "OK & Next", and "OK".

Annotating Documents

- But when Recogito guesses that word is a place, it matches with the digital gazetteers that are integrated with Recogito.
- In this case, we can geo resolve an annotation, we have associated specific geo coordinates

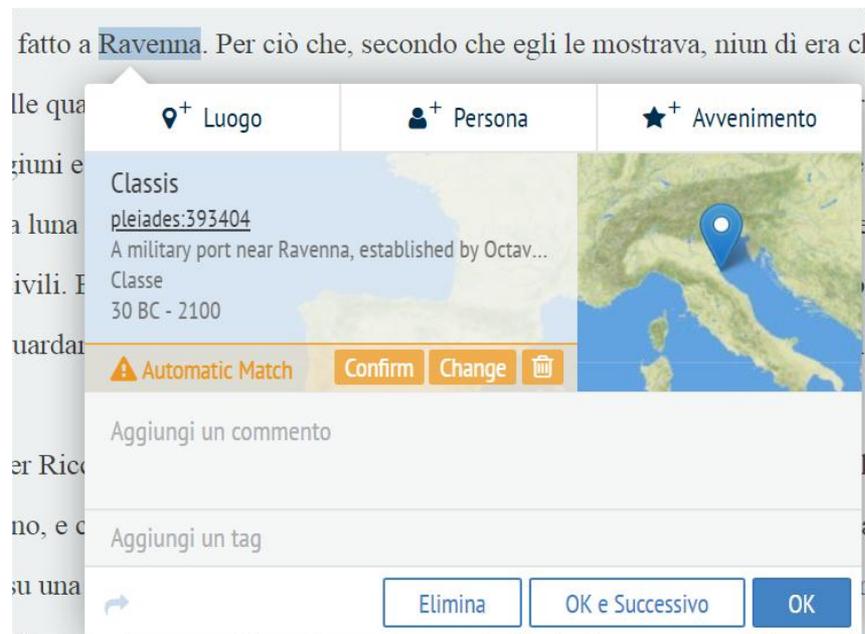
The screenshot displays the Recogito web interface. At the top, the user 'Tiziana @Tizja' is logged in. The document being annotated is 'Giornata_3.txt'. The interface shows a list of annotations for the word 'Milano'. A modal window is open over one of the annotations, showing the following details:

- Place** (selected): Marked as a Person
- Mediolan()**
- plades:583706**
- Now Milan. A Celtic settlement and crossroad, lat...**
- Milan, Milano, Mediolanum, Mediolanum, Mediol...**
- 550 BC - 2100**

The modal also includes a map showing the location of Milan, an 'Automatic Match' button, and options to 'Confirm', 'Change', or 'Add a comment...'. The background text from the document is visible but partially obscured by the modal.

Annotating Documents

- You can confirm the automatic guess or change it.
- If you don't confirm, the status of that geo-reference is set to “not verified”
- Some places are “unlocated”: Inghilterra
- Not all granularity levels are available (no streets or neighborhoods)

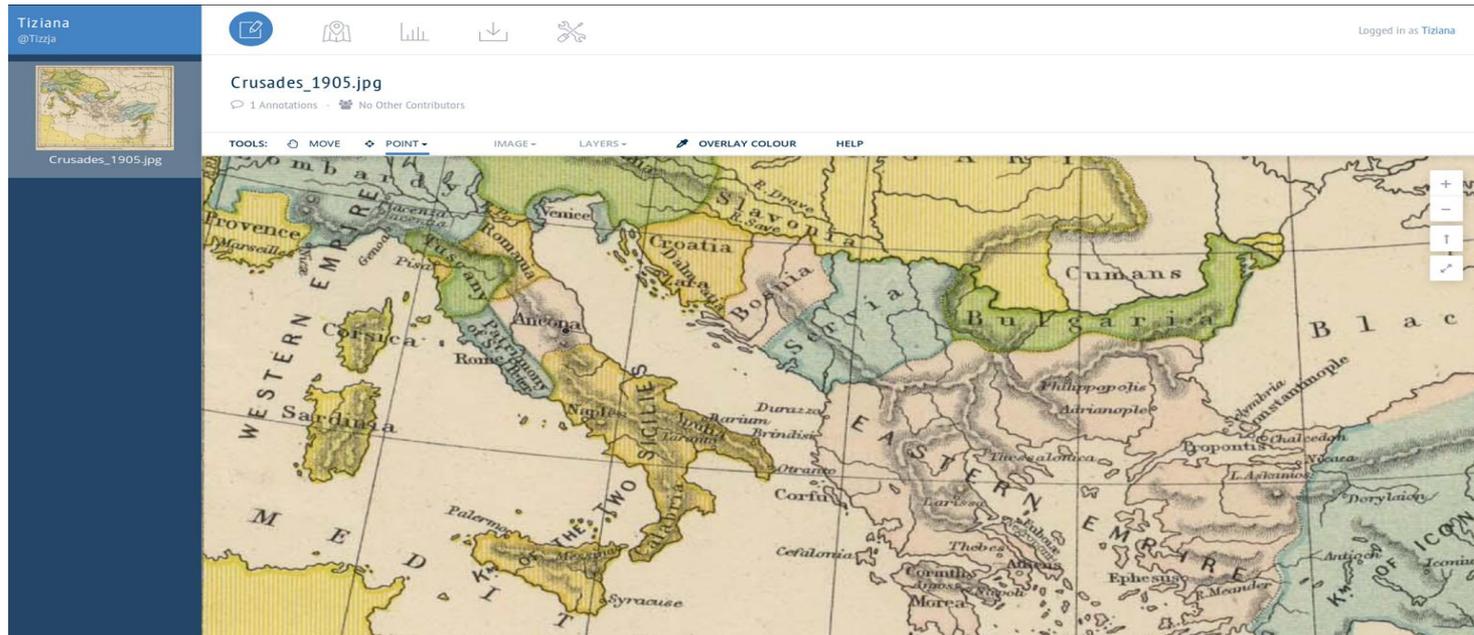


Not only places but people and events

- Very similar the way to annotate or even simpler. We can add **comments** and **tags**.
- What is important that information about “tags” - can be downloaded and used for our analysis.
- Different tags = different colors

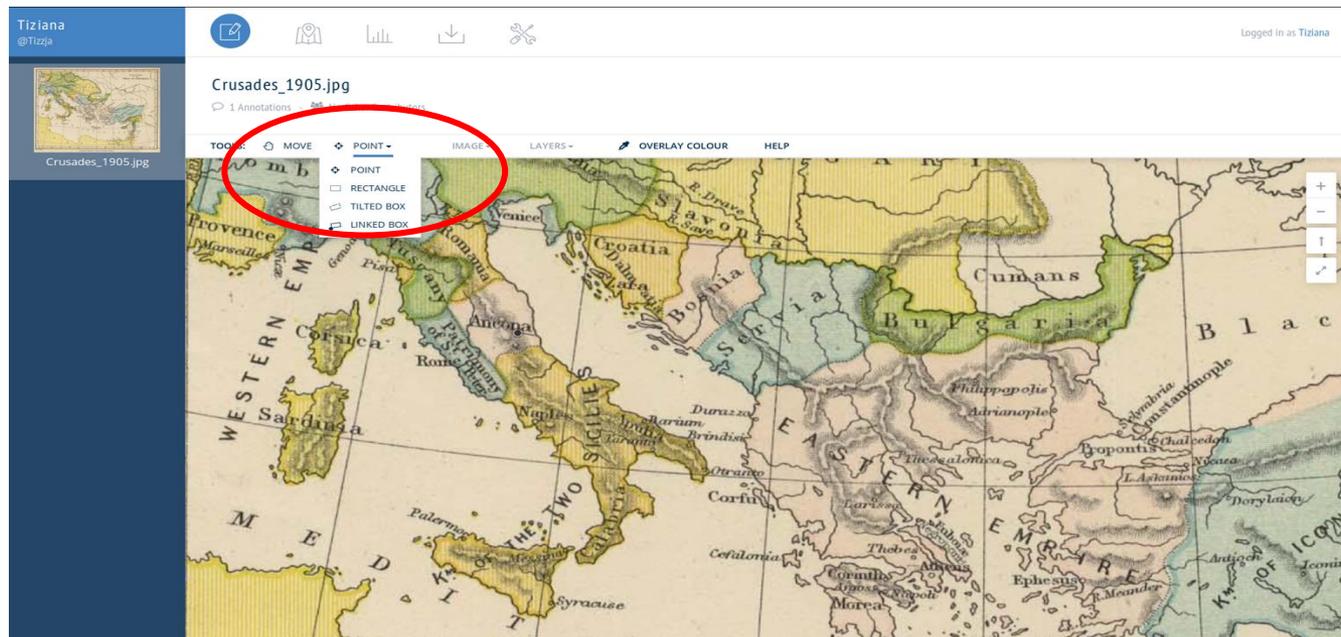
Images

- Uploading images is more or less the same as uploading text
- Point the place to the image



Images

- We can define the place in different ways: point, rectangle, titled box, linked box
- You can transcribe the name of the map



Images and gazeeters

Keyboard Shortcuts

SPACE toggle between MOVE and current annotation tool

ALT + V toggle annotation visibility

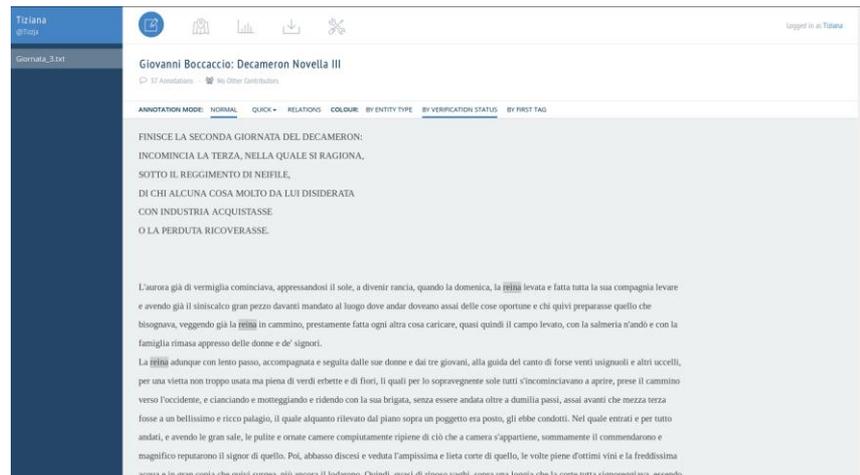
Map Navigation

Hold SHIFT to drag a box and zoom to the area

Hold ALT to rotate the map around the center of the screen

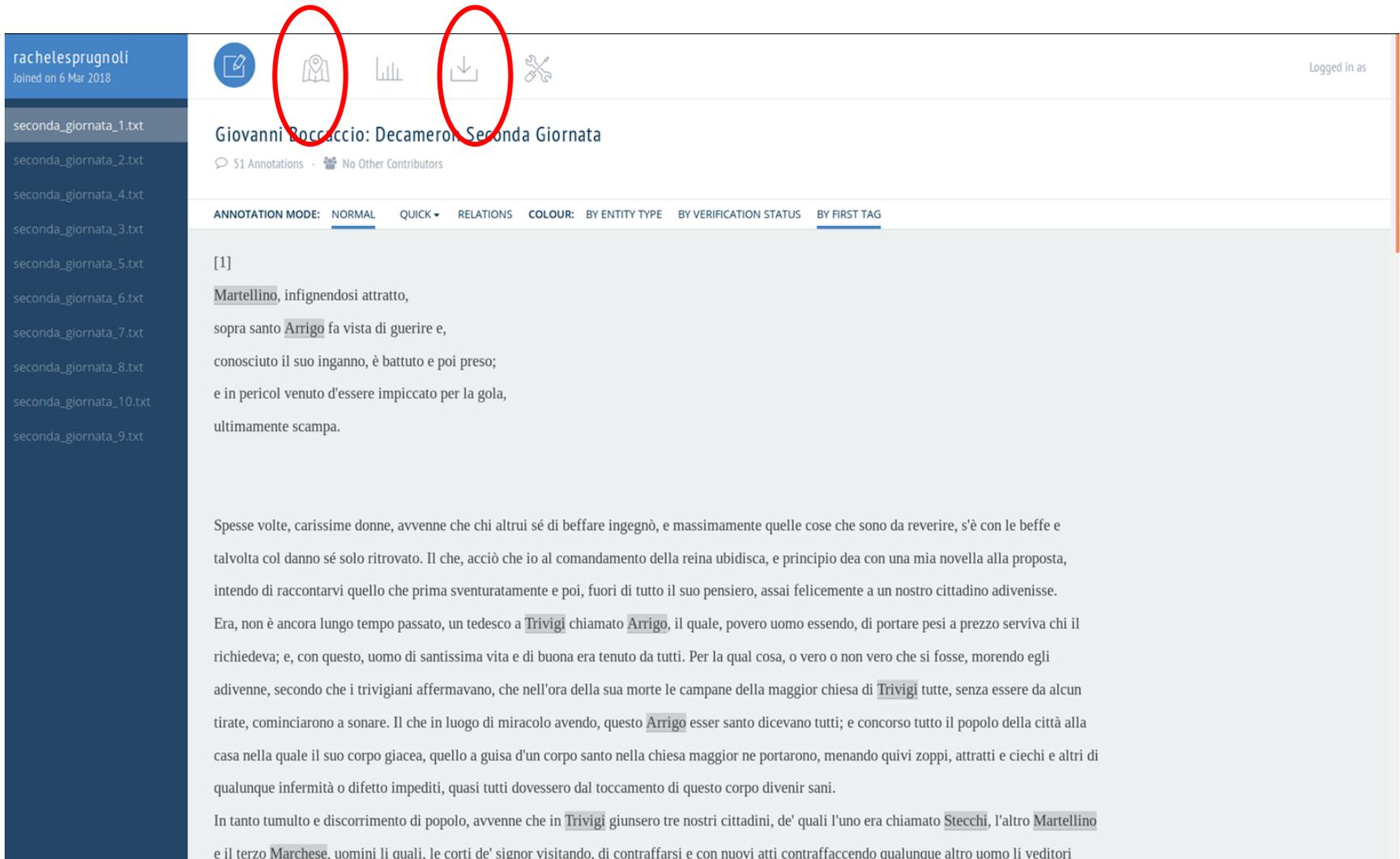
New Function: Relation Annotation

- Annotating relations among characters:



Tools, analysis and publishing

From Texts to Models to Visualizations



rachelesprugnoli
Joined on 6 Mar 2018

seconda_giornata_1.txt
seconda_giornata_2.txt
seconda_giornata_4.txt
seconda_giornata_3.txt
seconda_giornata_5.txt
seconda_giornata_6.txt
seconda_giornata_7.txt
seconda_giornata_8.txt
seconda_giornata_10.txt
seconda_giornata_9.txt

Logged in as

Giovanni Boccaccio: Decameron Seconda Giornata

51 Annotations · No Other Contributors

ANNOTATION MODE: NORMAL QUICK ▾ RELATIONS COLOUR: BY ENTITY TYPE BY VERIFICATION STATUS BY FIRST TAG

[1]
Martellino, infigendosi attratto,
sopra santo Arrigo fa vista di guerire e,
conosciuto il suo inganno, è battuto e poi preso;
e in pericol venuto d'essere impiccato per la gola,
ultimamente scampa.

Spesse volte, carissime donne, avvenne che chi altrui sé di beffare ingegnò, e massimamente quelle cose che sono da reverire, s'è con le beffe e talvolta col danno sé solo ritrovato. Il che, acciò che io al comandamento della reina ubidisca, e principio dea con una mia novella alla proposta, intendo di raccontarvi quello che prima sventuratamente e poi, fuori di tutto il suo pensiero, assai felicemente a un nostro cittadino adivenisse. Era, non è ancora lungo tempo passato, un tedesco a Trivigi chiamato Arrigo, il quale, povero uomo essendo, di portare pesi a prezzo serviva chi il richiedeva; e, con questo, uomo di santissima vita e di buona era tenuto da tutti. Per la qual cosa, o vero o non vero che si fosse, morendo egli adivenne, secondo che i trivigiani affermavano, che nell'ora della sua morte le campane della maggior chiesa di Trivigi tutte, senza essere da alcun tirate, cominciarono a sonare. Il che in luogo di miracolo avendo, questo Arrigo esser santo dicevano tutti; e concorso tutto il popolo della città alla casa nella quale il suo corpo giacea, quello a guisa d'un corpo santo nella chiesa maggior ne portarono, menando quivi zoppi, attratti e ciechi e altri di qualunque infermità o difetto impediti, quasi tutti dovessero dal tocco di questo corpo divenir sani.

In tanto tumulto e discorrimento di popolo, avvenne che in Trivigi giunsero tre nostri cittadini, de' quali l'uno era chiamato Stecchi, l'altro Martellino e il terzo Marchese, uomini li quali, le corti de' signor visitando, di contraffarsi e con nuovi atti contraffacendo qualunque altro uomo li veditori

Analyzing my Data

- How to export data from Recogito in a variety of formats suitable for visualizing and analysis in other tools, such as spreadsheets, databases or GIS
- Visualising and contextualizing geographical information within documents can be an important step in reaching a deeper understanding of their content, potentially highlighting phenomena that would have been otherwise difficult to identify

Creating different output

Tizzja
Joined on 21 Aug 2018

Logged in as

Annotations

CSV
Download annotations as a data table for importing into spreadsheet software or a GIS.

RDF
Download annotations and document metadata as RDF, encoded using Open Annotation and Dublin Core.

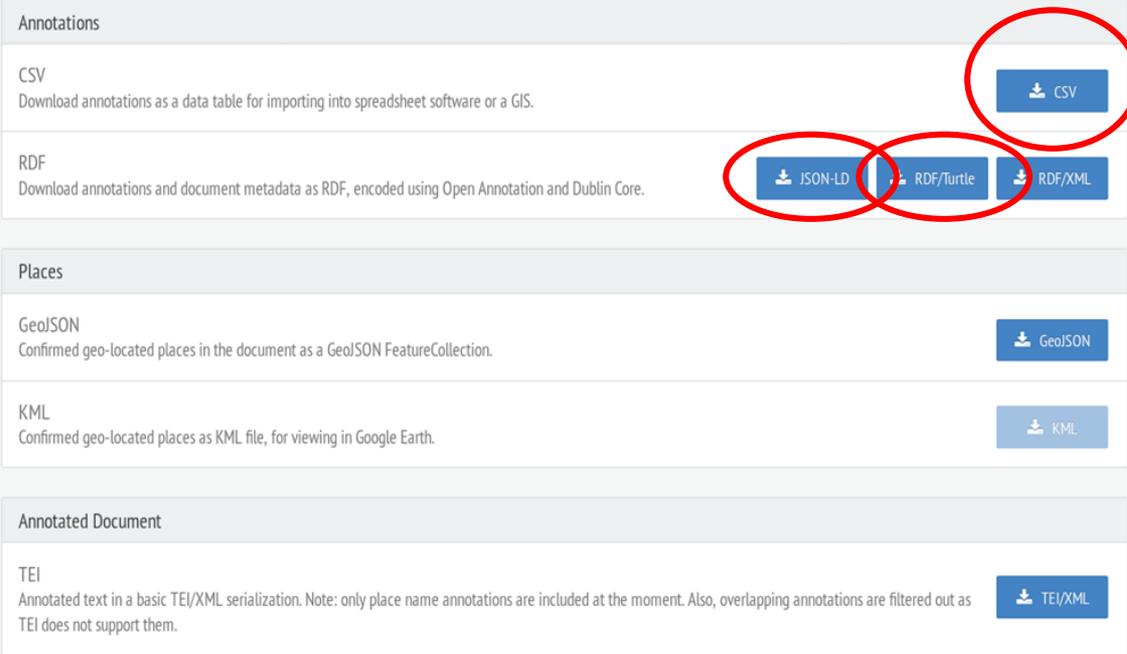
Places

GeoJSON
Confirmed geo-located places in the document as a GeoJSON FeatureCollection.

KML
Confirmed geo-located places as KML file, for viewing in Google Earth.

Annotated Document

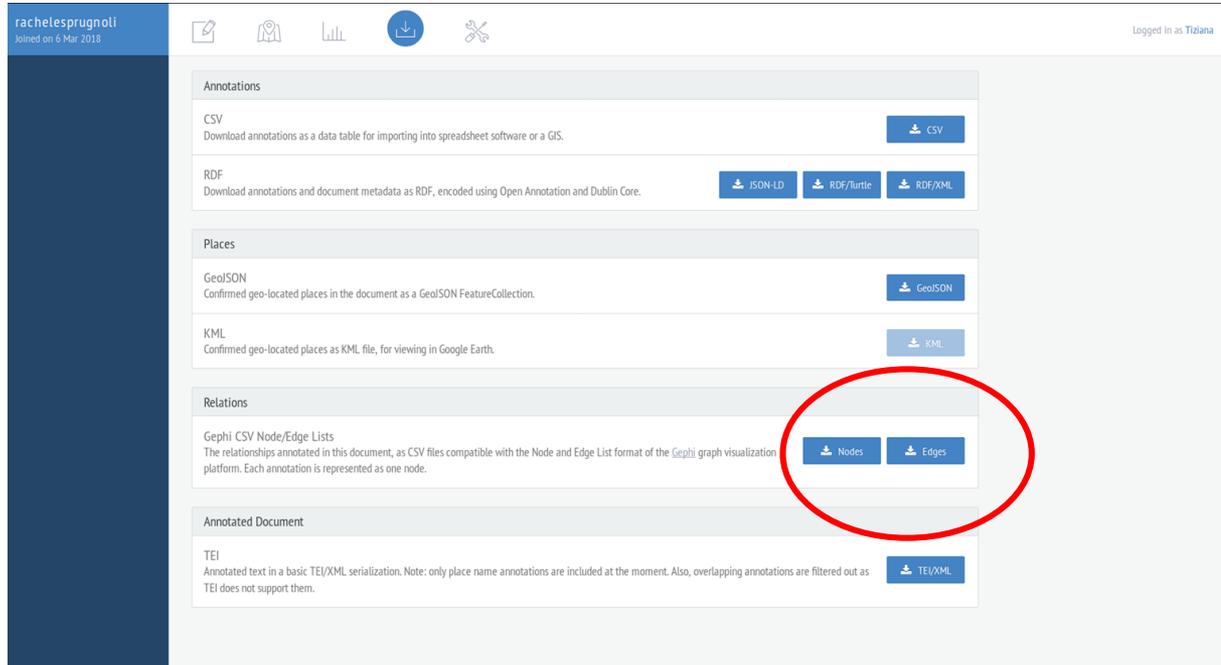
TEI
Annotated text in a basic TEI/XML serialization. Note: only place name annotations are included at the moment. Also, overlapping annotations are filtered out as TEI does not support them.



The screenshot shows a web interface with a dark blue sidebar on the left containing the user's name 'Tizzja' and the date 'Joined on 21 Aug 2018'. The main content area is light gray and contains several sections. The 'Annotations' section has three rows: 'CSV' with a 'Download CSV' button, 'RDF' with 'Download JSON-LD', 'Download RDF/Turtle', and 'Download RDF/XML' buttons. The 'Places' section has two rows: 'GeoJSON' with a 'Download GeoJSON' button and 'KML' with a 'Download KML' button. The 'Annotated Document' section has one row: 'TEI' with a 'Download TEI/XML' button. Red circles are drawn around the 'Download CSV', 'Download JSON-LD', 'Download RDF/Turtle', and 'Download RDF/XML' buttons.

Gephi

- Download the two files “edges” and “Nodes”



The screenshot shows the Gephi web interface. The user is logged in as Tiziana. The interface is divided into several sections: Annotations, Places, Relations, and Annotated Document. The 'Relations' section is highlighted with a red circle, showing two download buttons: 'Nodes' and 'Edges'. The 'Annotations' section has a 'CSV' button. The 'Places' section has 'GeoJSON' and 'KML' buttons. The 'Annotated Document' section has a 'TEI/XML' button.

Open Gephi and import:

- Edges
- Nodes

N.B. Post-processing needed to create correct networks...

Conclusion

- Allows to collaborate on the same project
- It is a very useful tool for teaching
- Allows to compare different sources
- People with different level of technical expertise to work with it

Methodological open issues

- How to deal with:
 - Ambiguity: e.g. places that are also people (and vice versa)
 - Imaginary places
 - Imaginary people or like the name Madonna
 - Places that you don't know if are existed or ...
- How to choose the right gazetteer?
- Do we annotate only named entities?
“Fu già nella nostra città” = Firenze

Hands-on Session

Boccaccio

Decameron collection of tales by Giovanni Boccaccio, probably composed between 1349 and 1353.

