



Ca' Foscari
University
of Venice



Physics
at Ca' Foscari
University of Venice

www.unive.it

Venice Long Data: Networks and Trees from Archives

Networks and Data Analysis
to extract stories from the past



Ca' Foscari
University
of Venice

The Source of ASVe

www.unive.it

Venice and its state archive ASVe are the ideal place where to start the development of a new standard for long data exploration and comprehension due to its unique place and for the depth and length of its data and for the quality of the conservation.

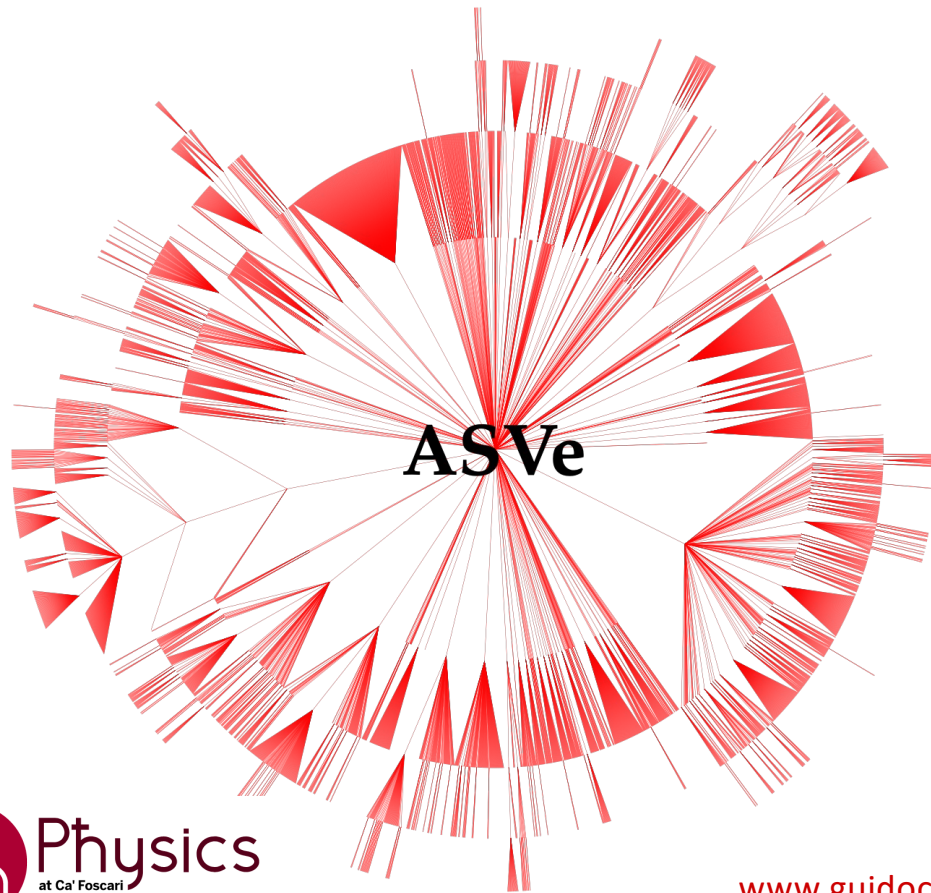




Ca' Foscari
University
of Venice

Structure of ASVe

www.unive.it



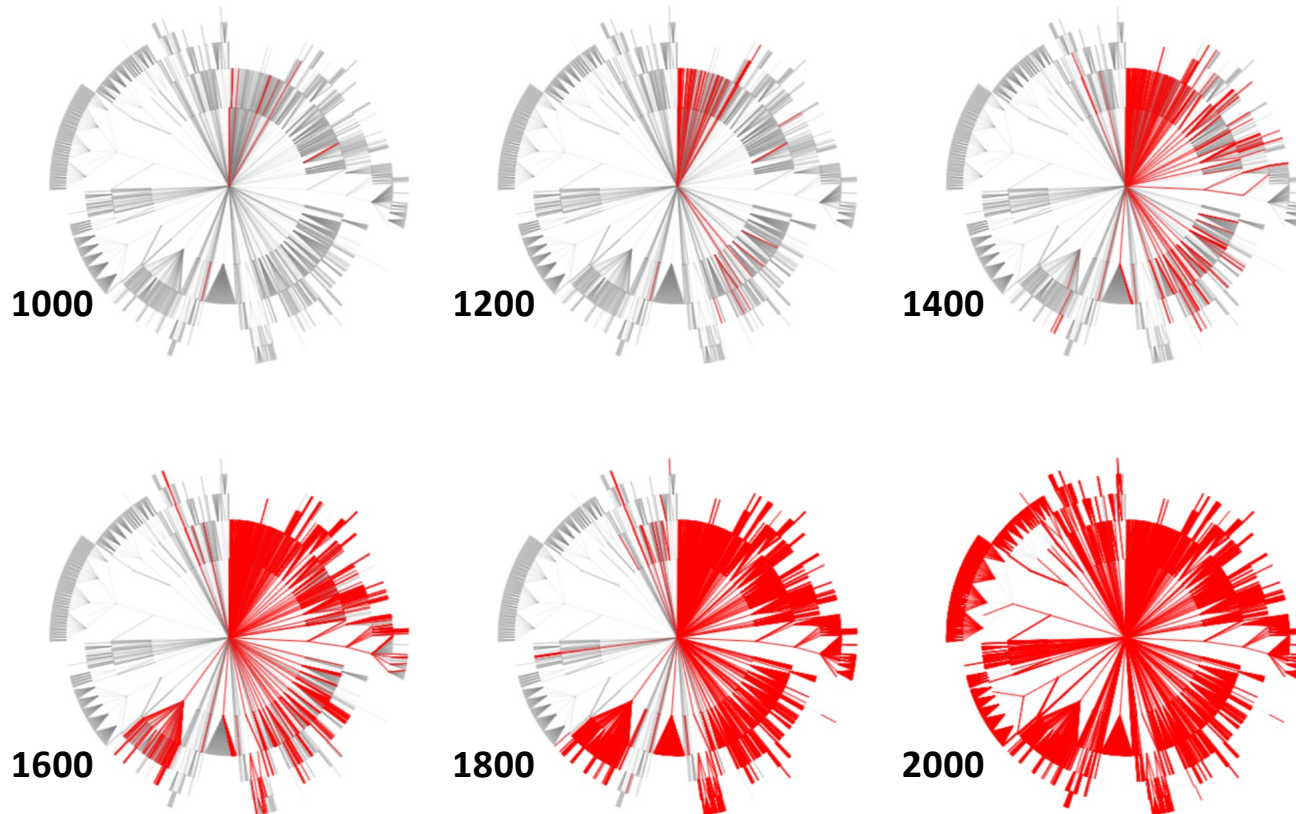
Venice Long Data is an interdisciplinary research project that aims to approach historical analysis from one point of view quantitative based on the historical database offered by the State Archives Veneto (ASVE).



Ca' Foscari
University
of Venice

Structure of ASVe

www.unive.it



Dendrogram maps of the ASVe – representing the hierarchical organization into

- fonds,
- series,
- sub-series

growth across one thousand year.
(Data starts from year 847
[testamento di Maru] to present).





Ca' Foscari
University
of Venice

Big Data/Long Data

www.unive.it

Big Data refer to a set of techniques and methods used in the analysis of large databases such as those that are made available by the advent of the internet, social media etc.

Long Data refers to the study, using these techniques, of databases that have a particular temporal depth (and therefore of historical relevance)

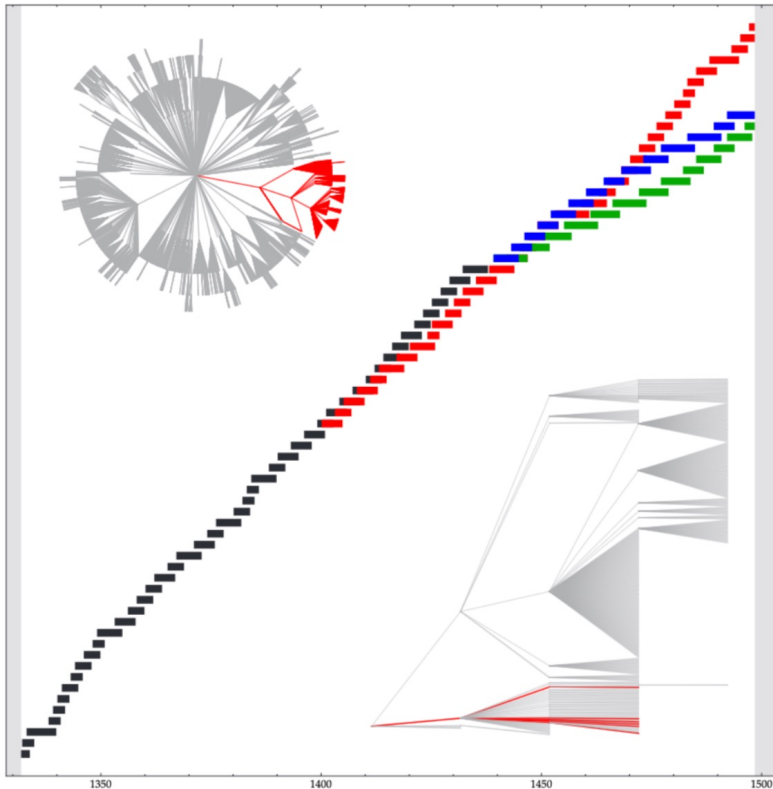




Ca' Foscari
University
of Venice

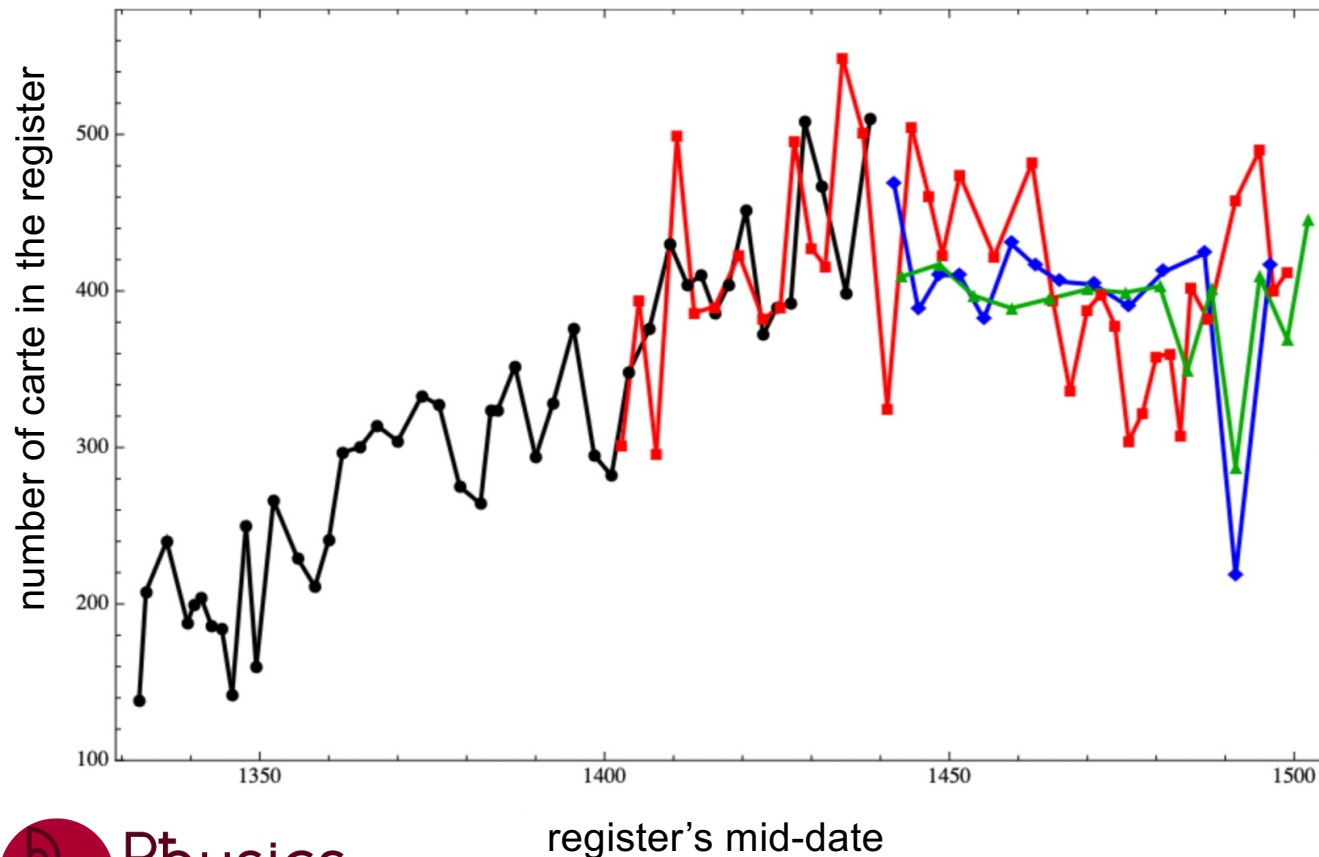
Senate Deliberations

www.unive.it



One important case of study and the one presented here is that of the **Deliberation of the Senate House.**

in the case of the Senato fond at the Venice State Archive (ASVe) only less than 1% of the documents has been transcribed



Number of carte (i.e. double sided pergamens) in each registro (archival unity) of the Senato-Deliberazioni series between 1332 and 1503. The color code indicates the sub-series:

- **Misti (black),**
- **Secreti (red),**
- **Terra (green)**
- **Mar (blue).**





Ca' Foscari
University
of Venice

Our analysis on archive

www.unive.it

Respecting the hierarchical organization of the ASVe (fonds, series, sub-series), we approach each archival unity (the smallest entity of information) individually and we apply the following steps:

- *digitalization* (mostly already available);
- *transcription* (using a combination of manual and automatic methods);
- *validation* and normalization of the transcribed sources by historians; *parse* to extract archival features (people, locations, summaries);
- *merging* of different archival unities to reconstruct the archival series.



Ca' Foscari
University
of Venice

Transkribus

www.unive.it



Transkribus

ScanTent

read&search

About

Resources

Plans & Pricing



Sign in

App

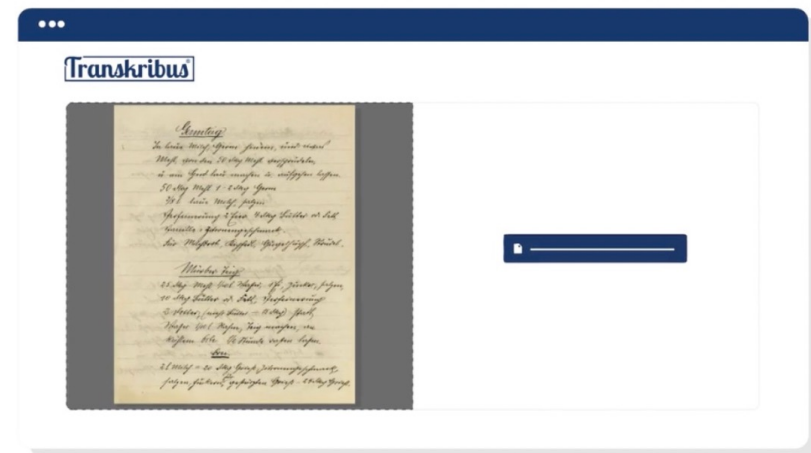
Transkribus

Unlock historical documents with AI

Transkribus is an **AI-powered** platform for **text recognition**, **transcription** and **searching of historical documents** – from any place, any time, and in any language.

Sign up for free

Watch Video



www.guidocaldarelli.com



@GuidoCaldarelli



Guido.Caldarelli@unive.it



Ca' Foscari
University
of Venice

Exploring data

www.unive.it

Application of our approach to Senato-Deliberazioni-Misti-Registri:

- 1. each registro is an archival unity;*
- 2. digitalization of individual pages;*
- 3. automatic transcription validated by historians;*
- 4. parse to segment and date **deliberations**,
extraction of **people** and **locations**;*
- 5. creation of a three-partite temporal network that
fully represents the archival unity (registro).*

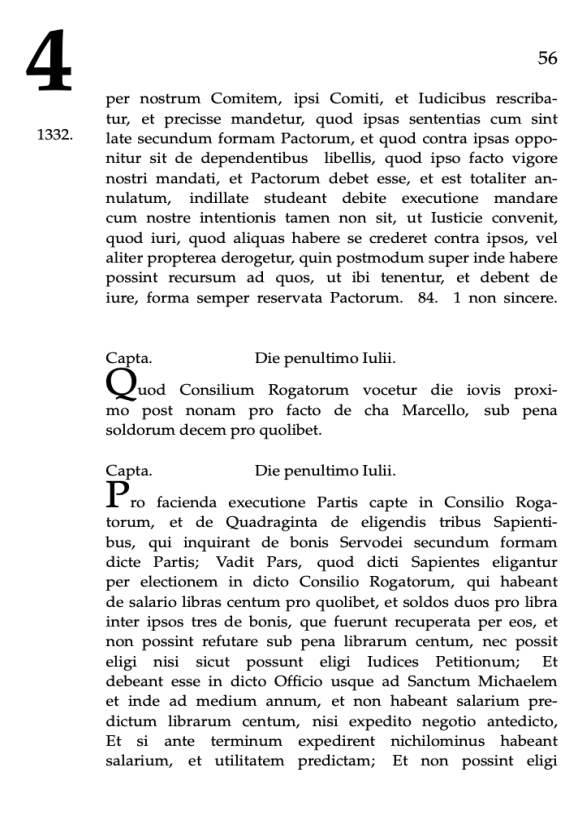
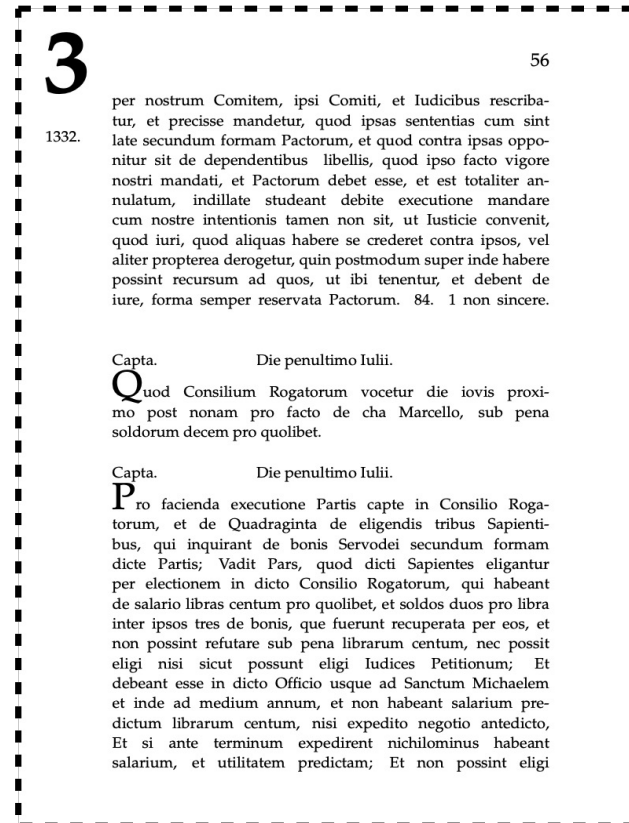
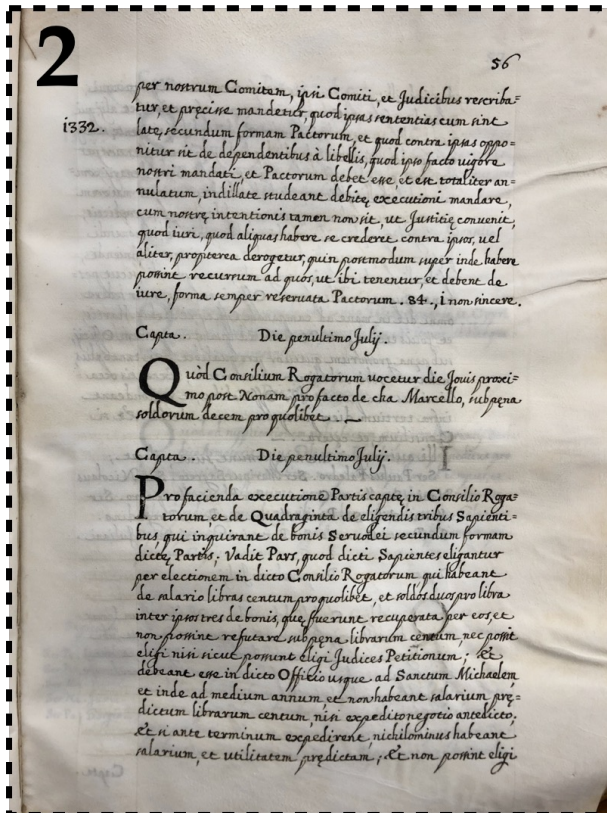


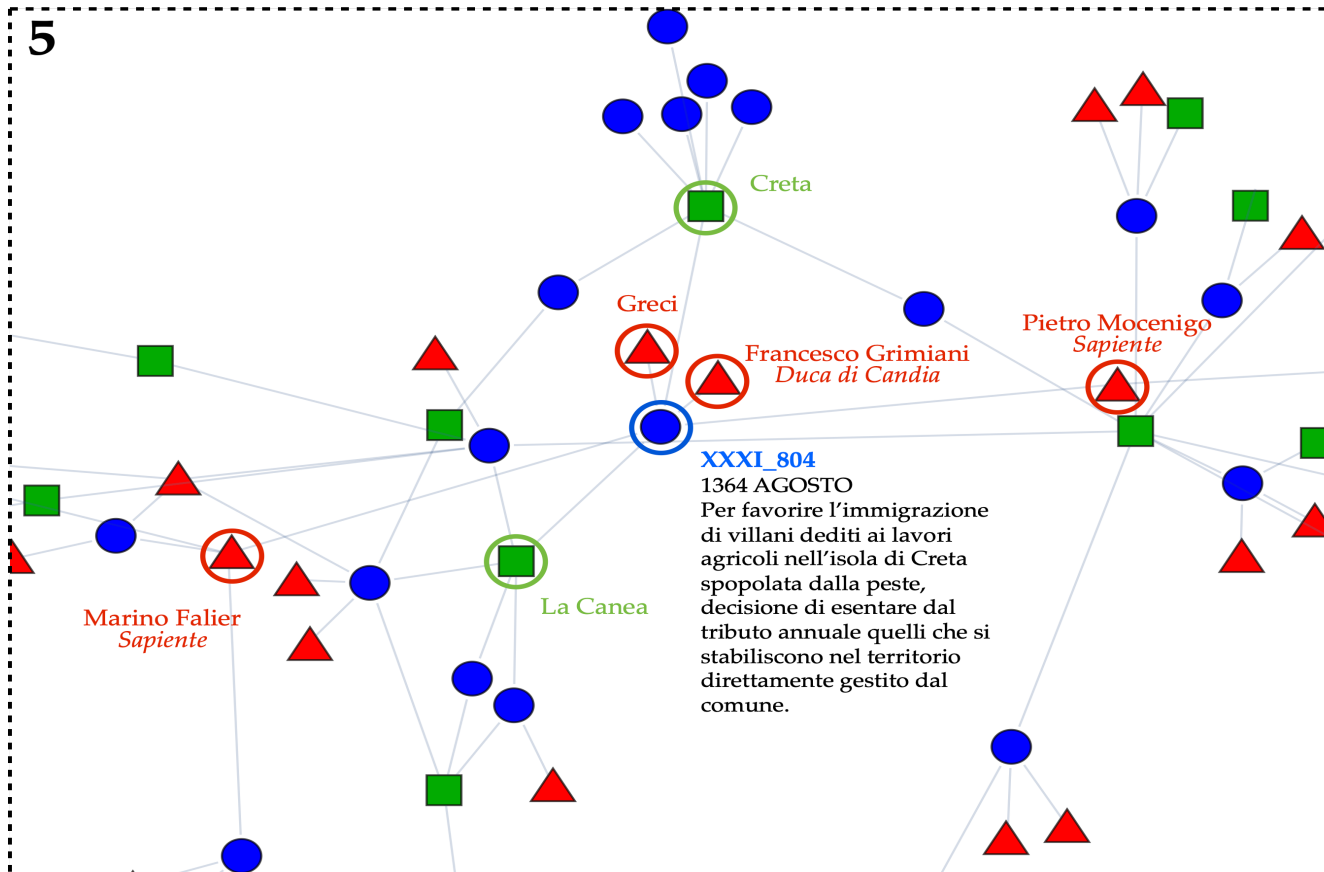


Ca' Foscari
University
of Venice

Transkribus

www.unive.it



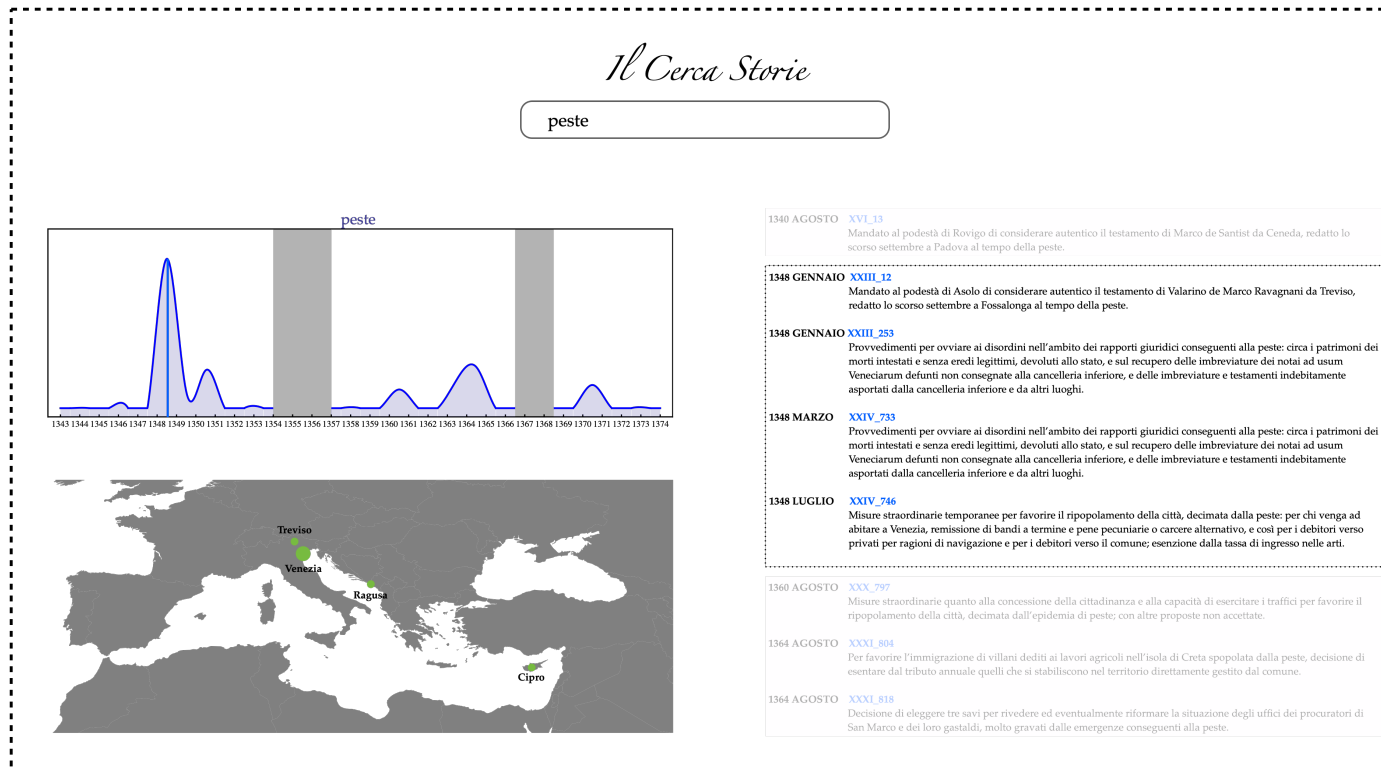




Ca' Foscari
University
of Venice

Network-based stories

www.unive.it



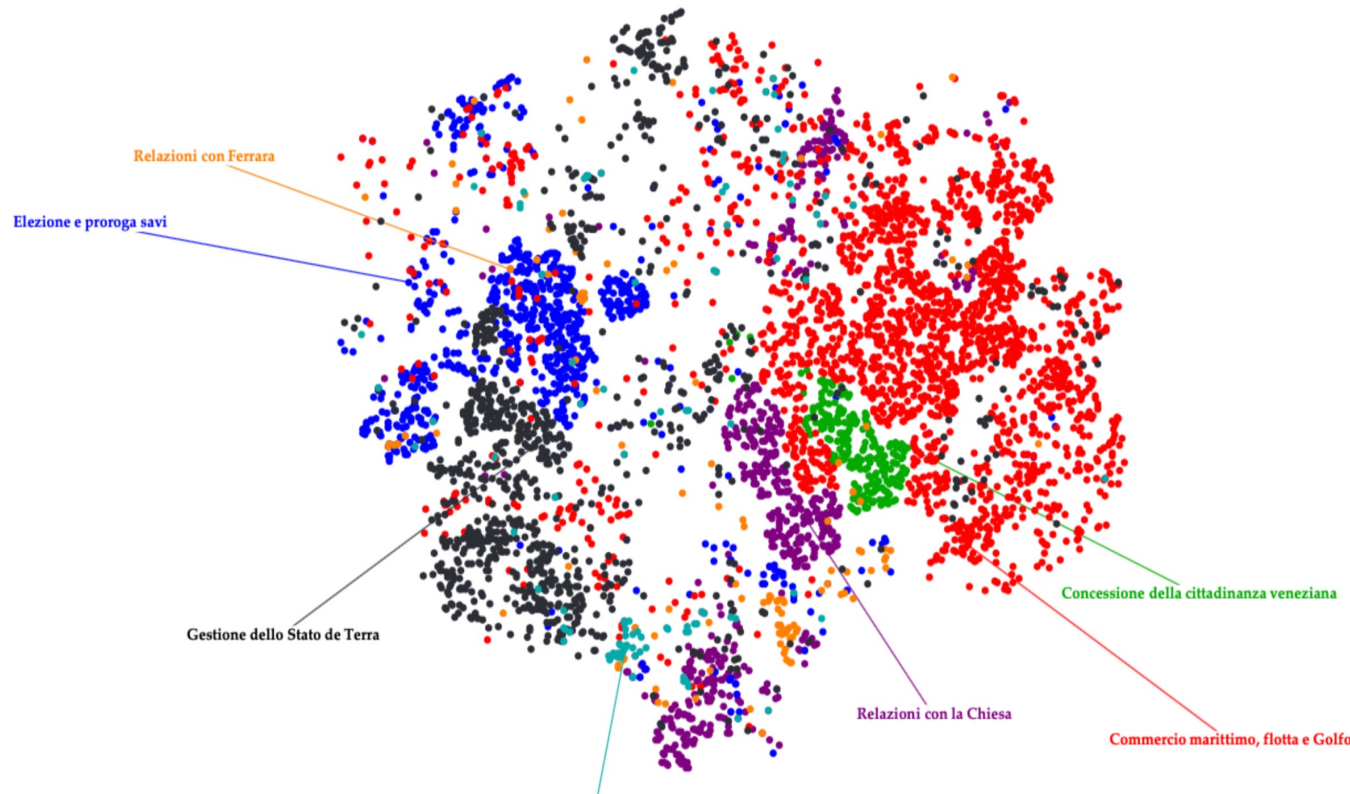
The archival series can now be queried and explored on the basis of **topic**, **people** or **places**, each time dynamically recreating the related original stories directly from the historical sources.



Ca' Foscari
University
of Venice

Big Data/Long Data

www.unive.it



15K deliberations between 1333 and 1375 are clustered semantically.

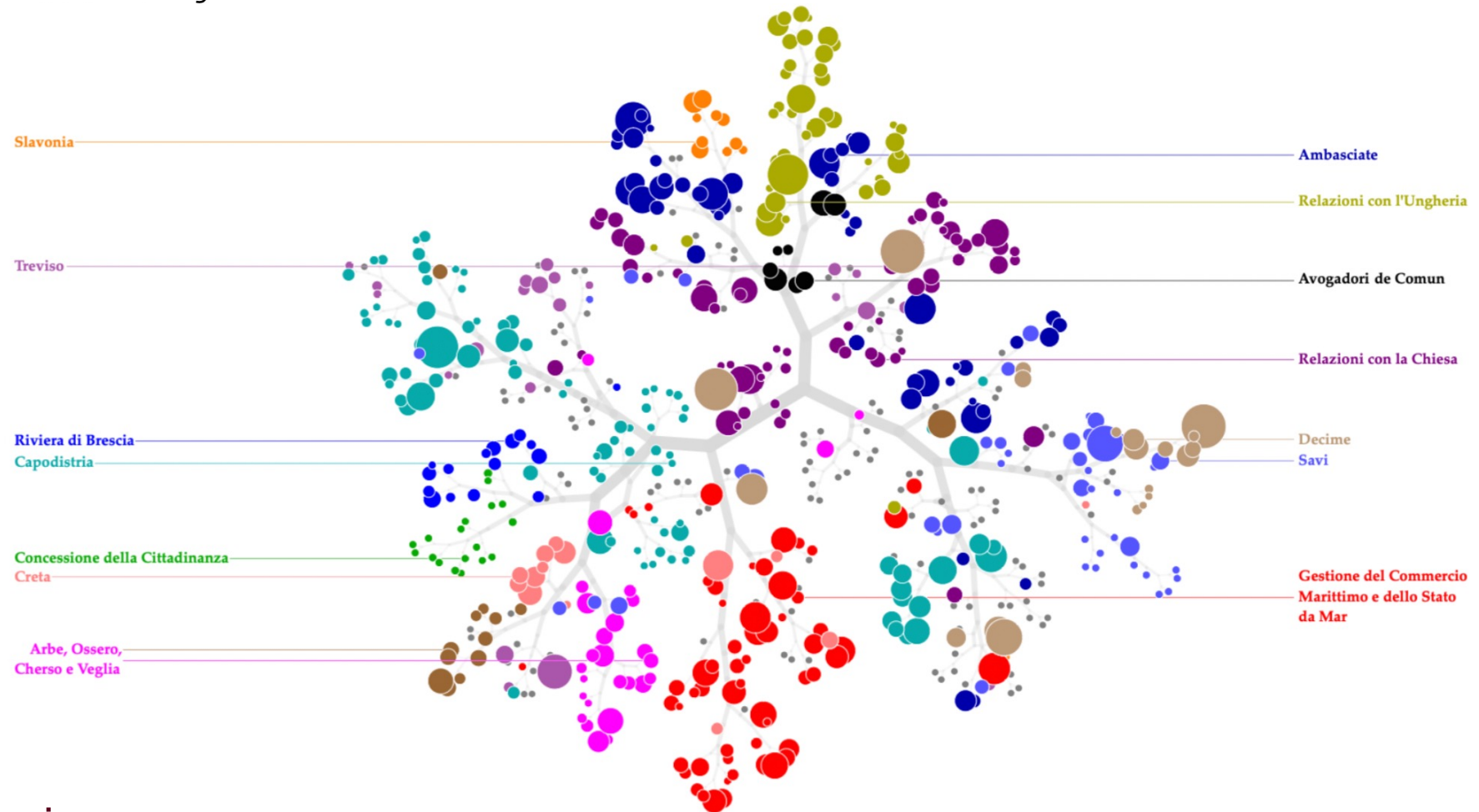
- **Red deliberations related to the management of the fleet**
- **Green citizenships**
- **Black is management of the Stato da Terra (orders to podestà);**
- **Purple relationship with church**
- **Cyan and Orange other states**



Ca' Foscari
University

Decision tree

re.it

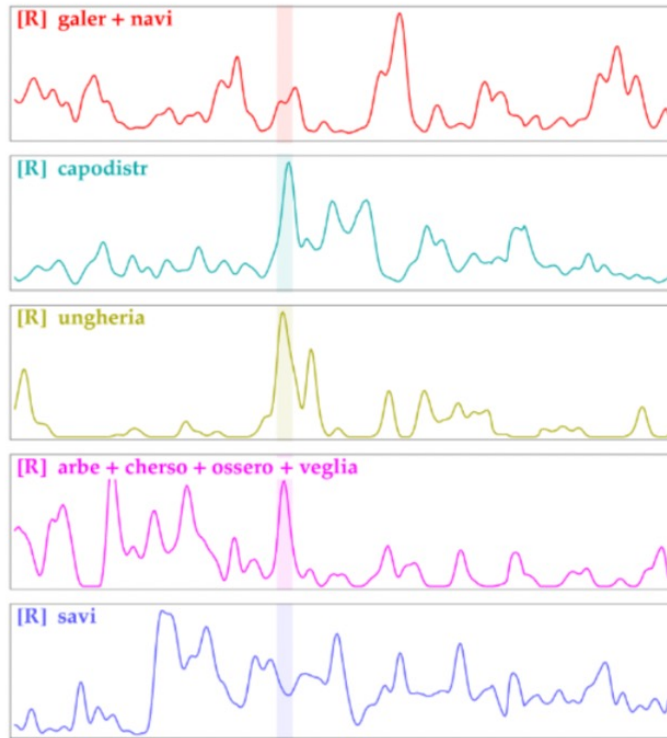
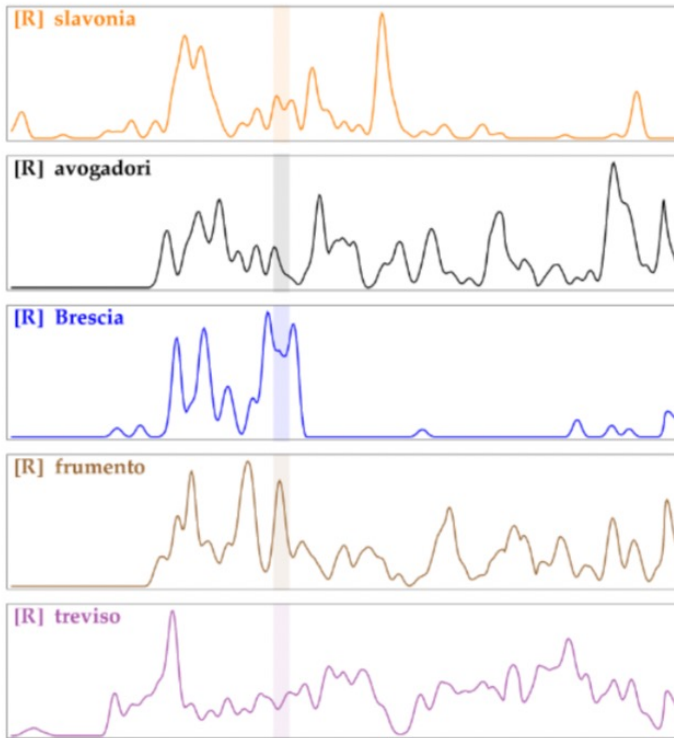




Ca' Foscari
University
of Venice

Co-occurrence

www.unive.it





Ca' Foscari
University
of Venice

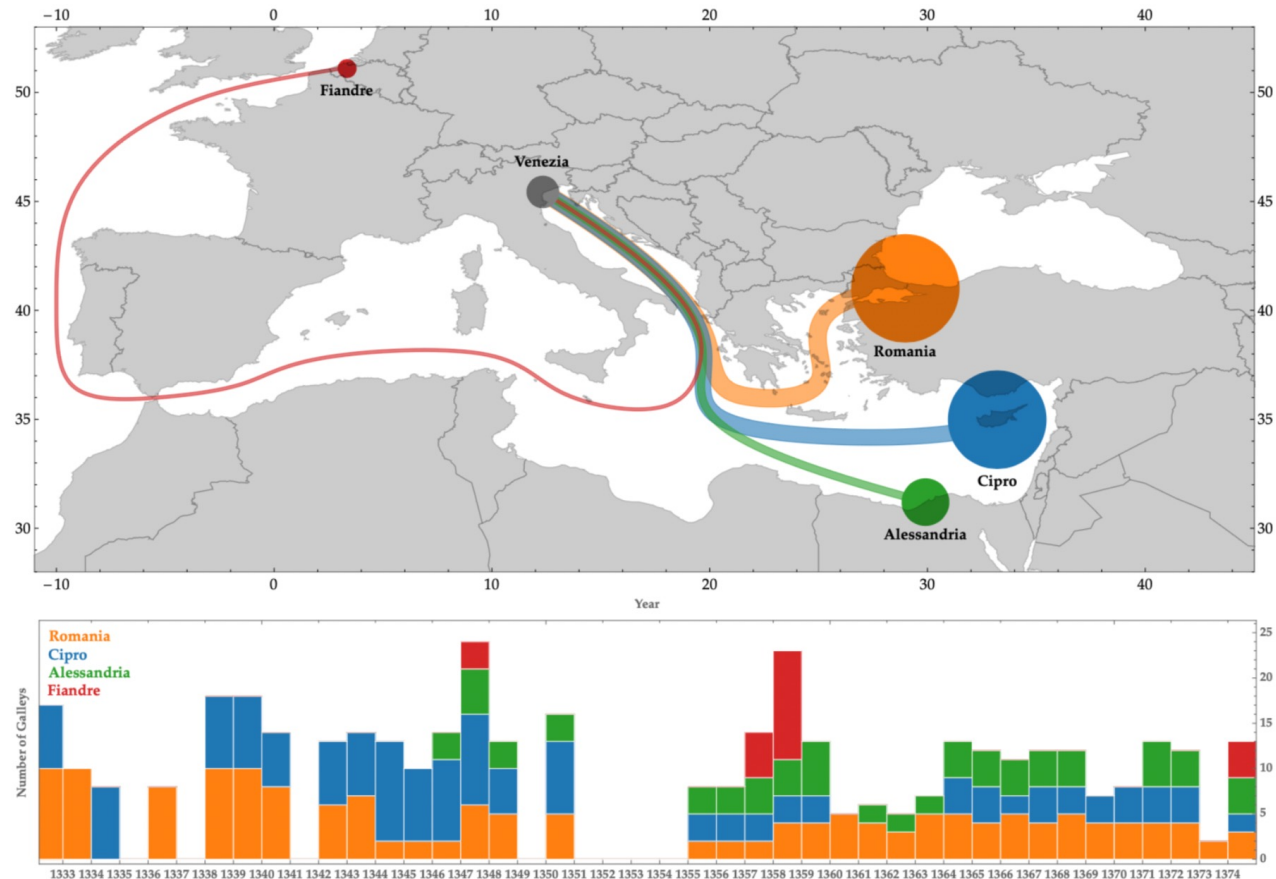
Trade evolution

www.unive.it

For each year and trade route we have automatically extracted from the relative incanto deliberation the number of ships (i.e. galere) that sailed toward the four major destinations:

- * Cyprus,
- * Romania (i.e. the Byzantium and Black Sea),
- * Alexandria
- * Flanders.

The resulting flux is a precise representation of the state of trade along the major Venetian trade routes during the decades **1330 to 1370**. We have found a total of 419 auctioned galleys (an average of ten each year); the histogram reports the breakdown of this total per-year and per-trade route

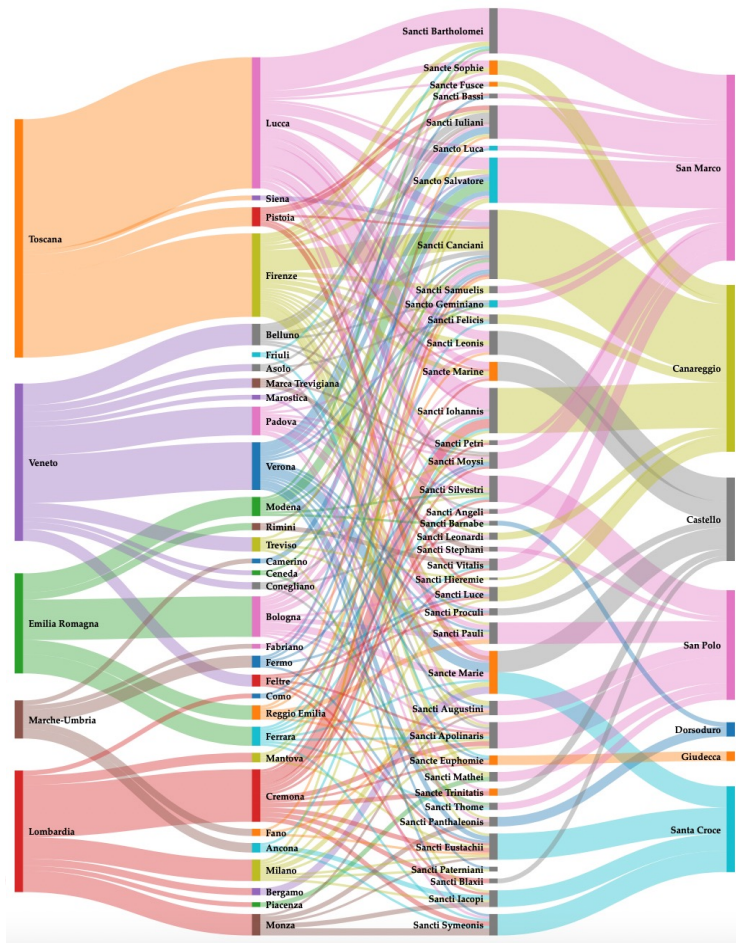




Ca' Foscari
University
of Venice

Citizenship

www.unive.it



The immigration flux towards Venice

The diagram displays

- region/city of emigration (left)
- contrada/sestiere of settlement within Venice (right)

as extracted from over 800 deliberations granting Venetian citizenship's found in our database. Only locations with two or more related deliberations have been included for clarity

www.guidocaldarelli.com



@GuidoCaldarelli



Guido.Caldarelli@unive.it



Ca' Foscari
University
of Venice

Conclusions and Perspectives

www.unive.it

Especially for the ones that started their activity hundreds of years ago (Long Data),
It is very difficult to have a clear and complete view of all the information stored.
Links must be created amongst different archival units irrespective of the place where they are
stored, their nature or media.

Scientific and technological advancements makes now possible to construct elaborate structures as a
temporal multi layer network that describes the evolution of the Venetian Society across the years



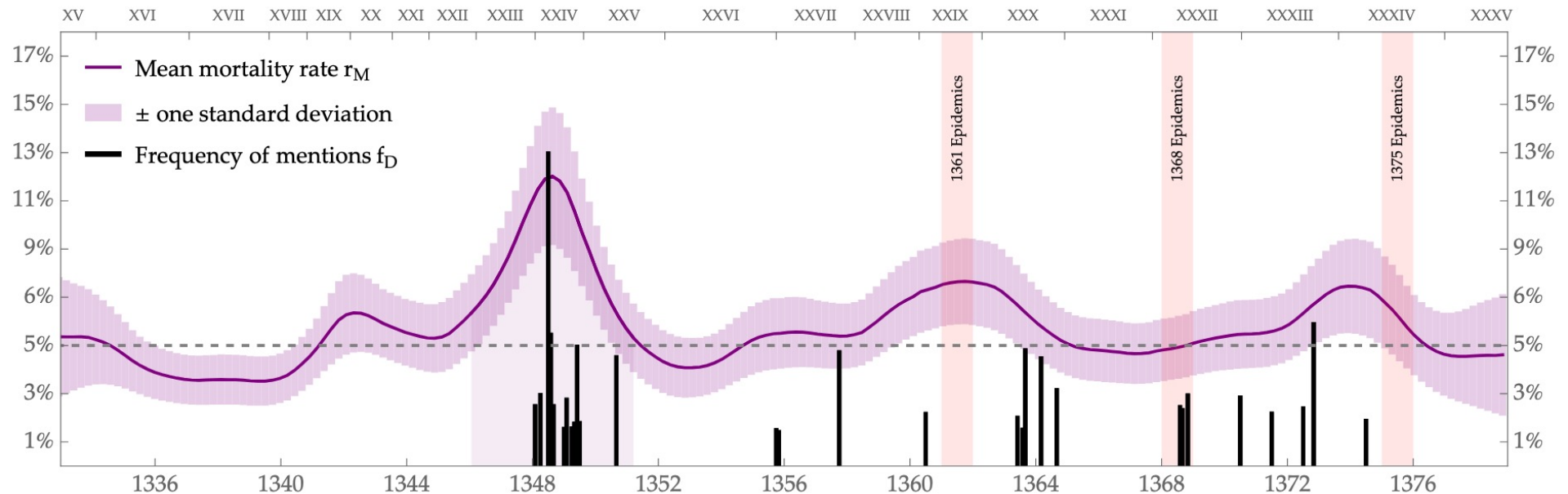


Fig. 2. Probabilistic estimate of the mortality rate curve $p(r|D)$ from the statistics of names citations in the *Senato Misti* dataset. In purple is the point-wise mean and in light-purple one standard deviation around it. The black bars are the mentions of "pestem" and other words in the Venetian Latin language meaning an epidemic plague (e.g. "epydimia", "epidimia", "epydemia", "pestem", "mortalitatis"). The shaded area around the 1348 pick is the one used in total mortality estimates. The light-red lines represent the plague outbreaks found in the London study (4).





Ca' Foscari
University
of Venice

Coworkers

www.unive.it



Raffaella Burioni
Un. Parma



Alessandro Codello
Ca' Foscari



Raffaele Santoro
Archivi di Stato



Renata Codello
Fondazione CINI



Chiara Casarin
Fondazione CINI



www.guidocaldarelli.com

 [@GuidoCaldarelli](https://twitter.com/GuidoCaldarelli)

 Guido.Caldarelli@unive.it