Ecole Polytechnique Fédérale de Lausanne

Project Report

Mapping the Correspondence Network of Galileo's Last Disciple

---

Submitted By :

MERZOUK Khalil (237492)

Supervisor :

Simon Dumas Primbault

Professors :

Robert West

Jérôme Baudry

Abstract

In recent years, Network Analysis has become a classic method to understand the social construction of knowledge. With the increasing digitization of historical archives, it has become possible to pursue these efforts on a much larger scale. Thanks to digitization of large corpora of both passive and active correspondences, and to their interoperability, historians can now aggregate their metadata in order to understand wider tendencies: weak and strong tie, geographical and linguistic extension, differences in chronological periods etc.

In this report we examine Vincenzo Viviani's career from a different perspective: his surviving letters. we use it as a tool to build a database then a network to visualize his correspondence. A combination of close and distant reading allows us to understand and discover elements of his life and career and shows how efficient a quantitative analysis can be and how it can improve our historical knowledge.

Contents

1 Introduction

Vincenzo Viviani(1622 - 1703) was born and raised in Florence. He was known for his abilities in mathematics and science. At the age of 17, in 1639, he became Galileo's disciple in Arcetri. Viviani worked throughout his life to preserve and complete all of Galileo's work. His goal was to honour Galileo by publishing it. His primary interest was to study ancient mathematics and physics, which he did by building his own network of scientists and friends around Europe. Viviani became a mathematician of renown, which allowed him to get many job offers: a position at the "Académie Royale" by Louis XIV of France, for example. This grew his reputation even more and made him a more important figure in Italy. He ended up at the same position previously held by Galileo as First Mathematician of the Tuscan Court. During his career, Viviani published many books on mathematical and scientific subjects and successfully edited an edition of Galileo's collected works, as he wished.[1]

In this report, we try to analyse when and how fast he built his network, and the way the network grew in parallel with his career. Analysing his correspondence from his early years until the peak of his career can explain a lot about how he developed and grew his network from scratch. It can also explain how he managed to get his reputation by looking at his correspondents and their connections and says a lot about how he gained his reputation around Europe. We first start by a whole network analysis, then, since it contains hundreds of people, choose to focus on a single one more specifically. We choose to study the case of Robert Southwell. We try to combine a quantitative analysis (distant reading) with a more historical one (close reading).

"Robert Southwell (1635–1702), Anglo-Irish gentleman with political aspirations and an interest in natural philosophy [...] This was precisely Southwell's aim when he embarked for France, the Netherlands, and Italy in late 1659.[...] The itinerary of Southwell's tour consisted primarily of long sojourns in Montpelier, Florence, and Rome, where he mixed with dignitaries, philosophers, and writers.[...] They became friends on the basis of their shared interests in mathematics and experiments. Their correspondence over almost forty years reveals that their friendship proved mutually beneficial: Southwell publicized Viviani's work in England and Viviani urged his employers to recommend Southwell to King Charles II. In the early 1660s, they also hoped that their friendship would extend itself into collaboration between the Royal Society of London and the Accademia del Cimento in Florence. [...] we are seeing Southwell not only using his friendship with Viviani for his own benefit, but also exploiting his connections in England for the benefit of his friend.[...] as they kept in touch during nearly the rest of their lives, they continued to benefit from their correspondence and mutual admiration."[2].

The analysis that will be done is based on the correspondence of Viviani from 1640 to 1701. We do a quantitative analysis on the letters recovered that have been sent and received by Viviani during this time period. Since the meta data of the letters is not available as a Data Base, part of this project will be about recovering it first.

How did Viviani's network change throughout his career? Which impact did Southwell have on it? How good are the historical archives to visualize Viviani's career from a different perspective allowing us the reconsider elements of his life? What can a quantitative analysis add to complete the historical analysis? We will answer these questions at the end of the report.The work will be split as follows:

- Methods :
  1. Database creation : How the database is obtained and what use it can have.
  2. Data cleaning : How we make our data usable and how it can be improved.
  3. Use of Data : What we do with our data once it is clean and usable.
- Results and Discussion:
  1. Database: We explain and discuss the final version of the database.
  2. Histograms: We analyze and discuss the different histograms obtained analytically then historically.
  3. Networks: We explain and discuss the networks obtained analytically then historically.
- Conclusion:
  We summarize all the results obtained and answer the problematic while proposing improvements .

2 Methods

All of the code used in this project is available on GitHub. In the following subsection we will see all the different tasks achieved step by step.

2.1 Database Creation

The letters belong to the "National Central Library of Florence" (BNCF), a public national library in Florence. Its meta-data has been digitized by the "Biblioteca del Museo Galileo"[3], which is an online library that contains historical documents. This is where we find all the letters sent to or delivered by Viviani. The letters are classified in galleries, each with a specific title (example: Gal. 10 - I, Anteriori, 8 , Ricci Ostilio, Problemi geometrici.) Each gallery contains a specific number of letters. Each letter on this website has specific personal information: Online access, Sender, Main title, Notes, Gallery, Related names, subjects and places (provenance and destination, if known), as we can see in figure 1.

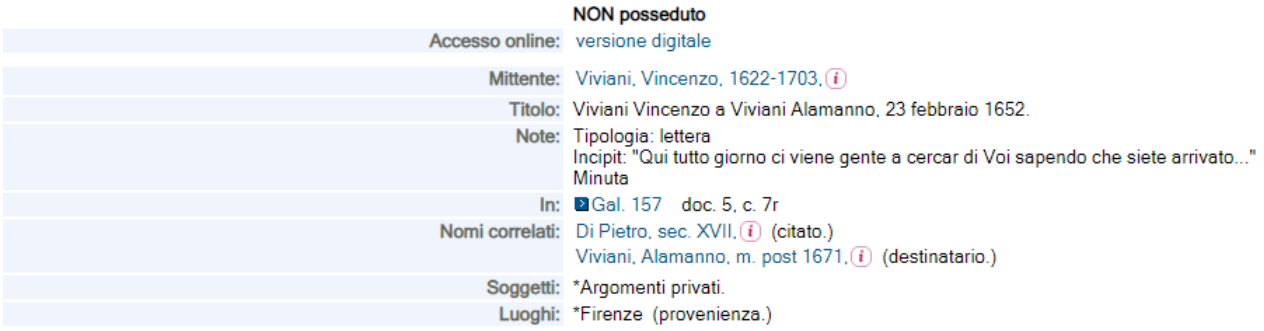

Figure 1: Meta-Data of the letters

What we are interested in for this project are Gal. 157 to 168, which are personal letters, Gal. 252 to 258, which are scientific letters and Gal. 275 to 286, which are letters of the Accademia Del Cimento, where we only keep the letters sent by or received by Viviani. Gal. 161 to 164 were unavailable outside of the BNCF's network, and most of the letters of Gal. 165 and 166 too, but the Museo Galileo gave us a special access to download all the necessary information for a limited time, thanks to Stefano Casati, who is responsible of the Biblioteca Digitale. The total number of letters is 3860 with 2871 sent or received by Viviani (1021 sent and 1850 received).

Each letter has a URL of the form "https://opac.museogalileo.it/imss/resourceXsl?uri=ID" with a specific letter ID. Going from a letter to a next one on the same Gallery increments the ID by one. This is what is used to

automate the operation for all the galleries and all the letters. All that has to be done is find manually the ID of the first and last letter of each gallery.

To scrap the data for a letter we use the BeautifulSoup[4] library that allows us to access to the HTML content of a page. We can then, using filtering commands, find the values of each row of the table as in figure 1.

Each row has its specific keyword in the HTML content of each letter. Mittente for Sender, Titolo for the main title, Note for the notes, In for the gallery, Nomi correlati for the related names, soggeti for the subjects and luoghi for the places, which contains the destinations and the provenances if known. At this point we create the empty data base as in figure 2 and fill it as follows: For each letter we extract the 7 elements that we add to

Sender	Main title	Notes	in:	Related names	subjects	places

Figure 2: Empty Database

the respective columns as we see in figure 3 we repeat the same operations for all letter of all galleries to obtain

	Sender	Main title	Notes	in:	Related names	subjects	places
0	Viviani, Vincenzo, 1622-1703,\n\n	Viviani Vincenzo a Viviani Alamanno, 23 febbra...	Tipologia: letteraIncipit: "Qui tutto giorno c...	Gal. 157 doc. 5, c. 7r\n	Di Pietro, sec. XVII, (citato.)\nViviani, Ala...	*Argomenti privati.\n	*Firenze (provenienza.)\n

Figure 3: Database With one letter's metadata

the first version of the Database which needs a lot of cleaning. This operation is explained in the following subsection.

2.2 Data Cleaning

This first version of the database needs a lot of cleaning as it is not yet usable. We fix all the columns by removing all the unwanted characters e.g "\n\n", at the end of the Sender column in figure 3, we split the main title column into a main title and date columns, since every title contains the sender, the receiver and the date, if known. We split the places column into destination and provenance columns, if both are known, and fill the unknown information with 'NaN' values. Finally, we turn the dates into a numerical format. The result obtained is as follows:

We get in total 1398 different names cited throughout all the letters.

We can see that for the expediter and receiver columns, we do not always have the person's name which exchanged with Viviani. For the dates we sometimes have an empty value or just the year, as the exact date is not known. The subjects are not always defined too. these cells are being treated whenever it is necessary by deleting the row which contains the unusable information. e.g. when we need to filter the letters by years, whenever the date is unidentified the row is not taken into consideration and just deleted. When the provenance

	Expéditeur	Notes	References	Noms cités	Sujet	Titre	Date	Provenance	Destination	Destinataire
60	Southwell Robert 1635-1702	Tipologia: letteraIncipit: "Per salis communi ...	<td class="fieldValue"> <img border="0" src="ht...	Viviani, Vincenzo, 1622-1703, (destinatario.)	Ricette mediche	Southwell Robert a Viviani Vincenzo	15/09/1660	NaN	NaN	Viviani Vincenzo
258	Lebesque Marainville de sec. XVII	Tipologia: letteraIncipit: "Ho ricevuto l'ordi...	<td class="fieldValue"> <img border="0" src="ht...	Chapelain, Jean, 1595-1674, (citato.)Viviani, V...	Argomenti privati	Lebesque Marainville (de) a Viviani Vincenzo	20/11/1665	Parigi	Firenze	Viviani Vincenzo
320	Dati Carlo Roberto 1619-1676	Tipologia: letteraIncipit: "Carlo Dati manda a...	<td class="fieldValue"> <img border="0" src="ht...	Viviani, Vincenzo, 1622-1703, (destinatario.)	Argomenti privati	Dati Carlo Roberto a Viviani Vincenzo	28/06/1667]	Firenze	NaN	Viviani Vincenzo
342	Chapelain Jean 1595-1674	Tipologia: letteraIncipit: "Io vi scrivo di ra...	<td class="fieldValue"> <img border="0" src="ht...	Colbert, Jean Baptiste, 1619-1683, (citato.)Viv...	Notizie di libri	Chapelain Jean a Viviani Vincenzo	25/03/1668	Parigi	Firenze	Viviani Vincenzo
358	Chimentelli Valerio 1620-1668	Tipologia: letteraIncipit: "Mi dispiace della ...	<td class="fieldValue"> <img border="0" src="ht...	Viviani, Vincenzo, 1622-1703, (destinatario.)	Argomenti privati	Chimentelli Valerio a Viviani Vincenzo	19/09/1668	Firenze	NaN	Viviani Vincenzo
...	...	...	...	...	...	...	...	...	...	...

Figure 4: Database with minor fixes

or destination is not defined, the column value is replaced by 'unidentified'.

The column Notes is also corrected by removing the 'Incipit' word at the beginning, since it was common to all letters and not useful. The references have not been touched as it is not used at all for the purpose of this project, but could come useful for another one. The provenance and destinations column had sometimes the wrong values that had to be corrected manually, since the algorithm to find these values is automated for all the letters. For example, sometimes, the notations for the destination and provenance were written differently, which gave as a result for the specific letter an empty provenance and a destination of 'FirenzePisa' (see figure 5).

1843	Rinaldini Carlo	Incipit: "Servirò V.S. e quel signore insieme ...	<td class="fieldValue"> <img border="0" src="ht...	Della Rena, Cosimo, sec. XVII, (citato.)Galilei...	NaN	Rinaldini Carlo a Viviani Vincenzo	25/02/1658	NaN	FirenzePisa	Viviani Vincenzo	1658

Figure 5: Special case of destination and provenance

Since these columns were not redundant throughout the letters, a manual correction was easy to implement for each of these letters. We also fix all the names that contain a birth-date and death date by finding all the different patterns and fixing them one by one automatically. This step is important for the next subsections, as we do not want to have duplicates of the same names while creating histograms or networks. Once everything was corrected we could finally start using our data, as we will see in the next subsections.

2.3 Use of Data 1: Histograms

With the actual database we can already start the quantitative data analysis. It provides a visual representation of data distribution. It can also show the outliers, which are interesting for the analysis of Viviani's network. We can plot histograms of the number of letters sent and received by Viviani throughout the years, or the top cities he sent and received letters from. We can also find the top 9 correspondents, as we can see on figure 6. To draw the histograms we use the matplotlib[5] library.

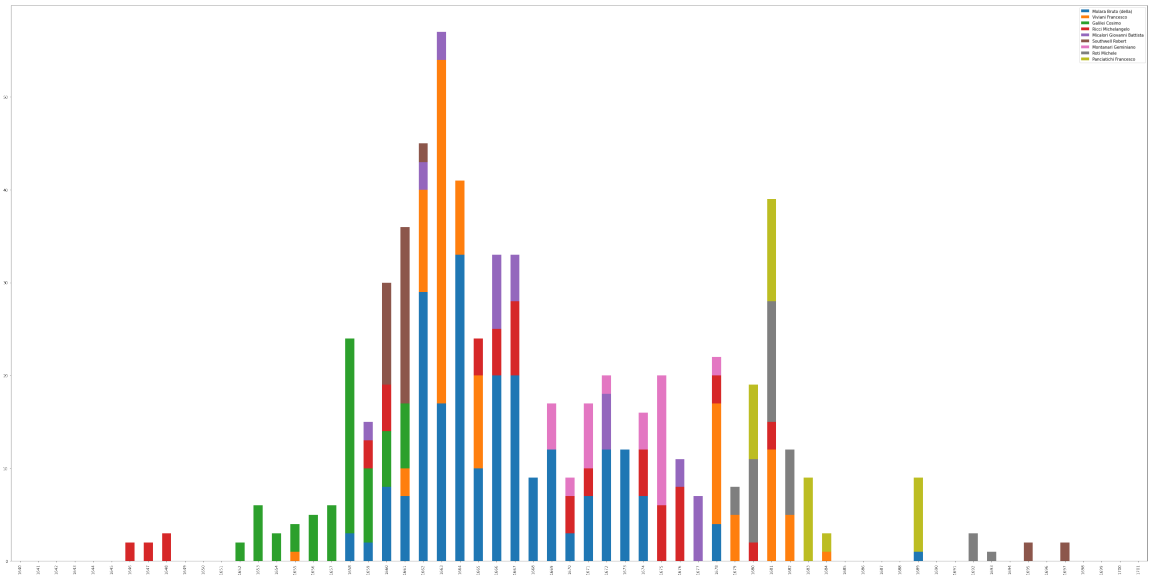

Figure 6: Top 9 correspondents throughout the years

For the sake of a yearly analysis, we also create a new column that contains the year on which the letter has been sent or received, whenever the information is available. As this makes plotting the histograms easier. The histograms can be filtered by years, sender, receiver, destination and provenance etc. The histograms already give us a better visualization of Viviani's correspondence throughout the years. but the main focus of this project is his network itself, which we will try to simulate as accurately as possible using the letters, in the following section.

2.4 Use of data 2: Networks

Networks are an effective visual tool because they present information quickly and easily. They allow us to do a distant reading on our literary data. We draw all the networks using the Networkx[6] library. First, let us explain how the network is built. We start by filtering the letters to only keep the ones that has a valid sender and receiver, as these two are the key elements of the network, since they represent the nodes. We drop any

letter that has an unidentified sender or receiver to avoid having unidentified nodes in the networks, that could potentially stack up and have higher weights. The key elements of the graph are:

- NODES: we go through all the senders and receivers, and for each one we add a red node to the graph. Each person that is only cited and not part of the correspondents is represented with a blue node, each person that is a correspondent and is cited in someone else's letter is represented with a green node.
- EDGES:For the edges, each edge between a cited person a participating person is represented with a blue edge, if a participant cited another participant, the edge is red and finally, if two participants both cited each others the edge is green.
- WEIGHTS: we use 3 different algorithms, one where any new edge that was non-existent is added with a weight of 1, and if it is already existent we increment the weight by one. second algorithm is whenever we add a new edge that was non-existent, we initialize it with a weight of 1 whenever one of its two nodes is a cited person only (blue node), and a weight of 5 whenever the edge is between two participants. (red or green nodes) we increment the edges weight by 1 whenever we add an edge that is already existent. the third one is the same as the second one but with higher weights initialization. A new edge between two participants is initialized with a weight of 10 and incremented with 1 whenever it is added again, and an edge between a participant and a cited person is initialized with 1 if it was non existent, and then incremented with 1 if re-added. These 3 algorithms become important when we try to find and draw the clusters, as we explain later. The weights influence the position of the nodes on the graph containing the clusters.

The default setup is not readable since the nodes are placed randomly around the center (Viviani, as he is present in all letters), as we will explain in the next section, which is why we use the spring layout. The spring layout is a way of positioning the nodes using the Fruchterman-Reigold force-directed algorithm. "Force-directed algorithms are among the most flexible methods for calculating layouts of simple undirected graphs. Also known as Spring Embedders, such algorithms calculate the layout of a graph using only information contained within the structure of the graph itself, rather than relying on domain-specific knowledge. Graphs drawn with these algorithms tend to be aesthetically pleasing, exhibit symmetries, and tend to produce crossing-free layouts for planar graphs." [7]

- CLUSTERS: Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group than those in other groups. It aims to segregate groups with similar traits and assign them into clusters. This is part

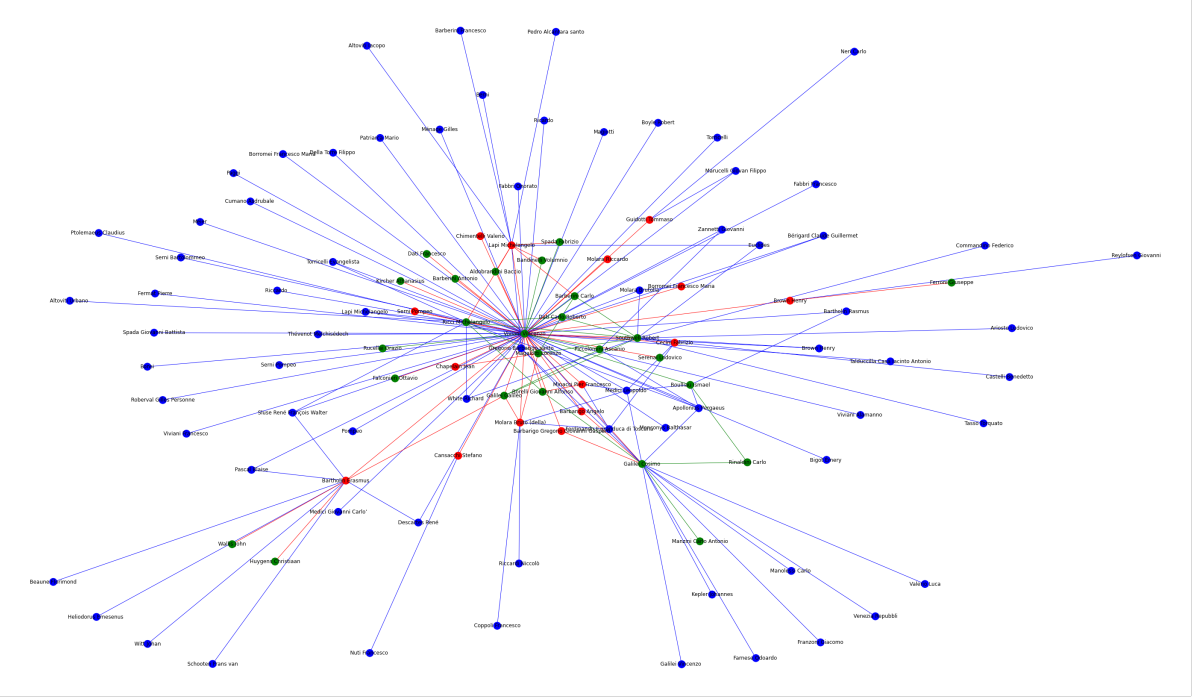

Figure 7: Network of Viviani's correspondence during 1660 with Viviani

of the distant reading that we do on our database. It allows us to group all the important persons that participated or have been cited in the letters to have a better visualization on Viviani's network. This, combined with a close reading allows us to see how his network really behaved during his career and life, and what kind of relationship he had with all these correspondents.

To find the clusters, we start by computing the partitions using the best_partition function of the community_louvain library which computes partitions in an undirected graph. This justifies why we use an undirected graph instead of a directed one for finding the clusters, as the function only works for undirected graphs. we then manually compute the positions of each node of the graph based on the cluster to which it belongs. This way we get a result that is much readable instead of the default one as we can see in figure X.

The algorithm works as follows: It computes the partition of the graph nodes which maximises the modularity using the Louvain heuristics. This is the partition of highest modularity, i.e the highest partition of the dendrogram generated by the Louvain algorithm. The Louvain algorithm[8] is an algorithm that detects communities in large networks. It maximizes a modularity score for each community where the modularity quantifies the quality of an assignment of nodes to communities. This means evaluating how

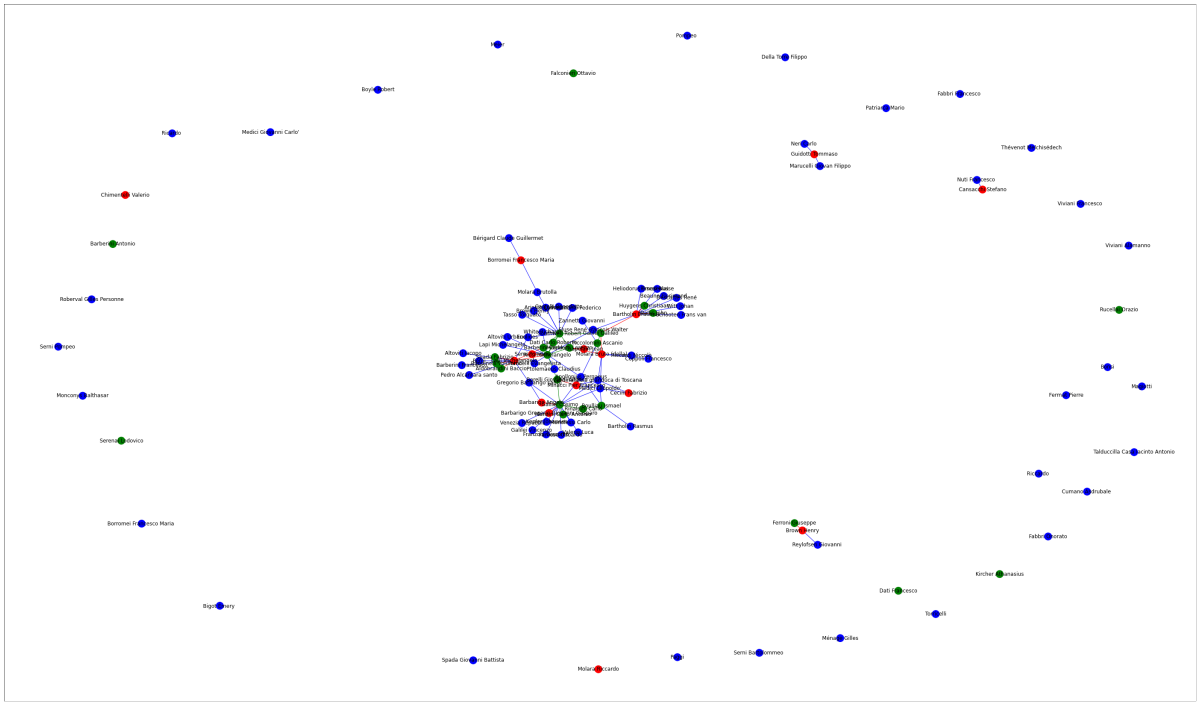

Figure 8: Network of Viviani's correspondence during 1660 without him

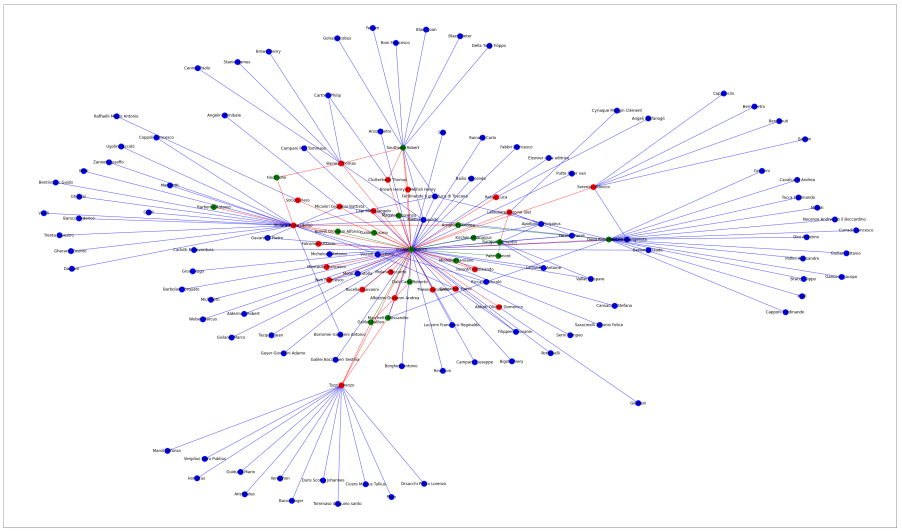

Figure 9: Viviani's network for the year 1662

much more densely connected the nodes within a community are compared to how connected they would be in a random network. It is a hierarchical clustering algorithm that recursively merges communities into a single node and executes the modularity clustering on the condensed graphs.

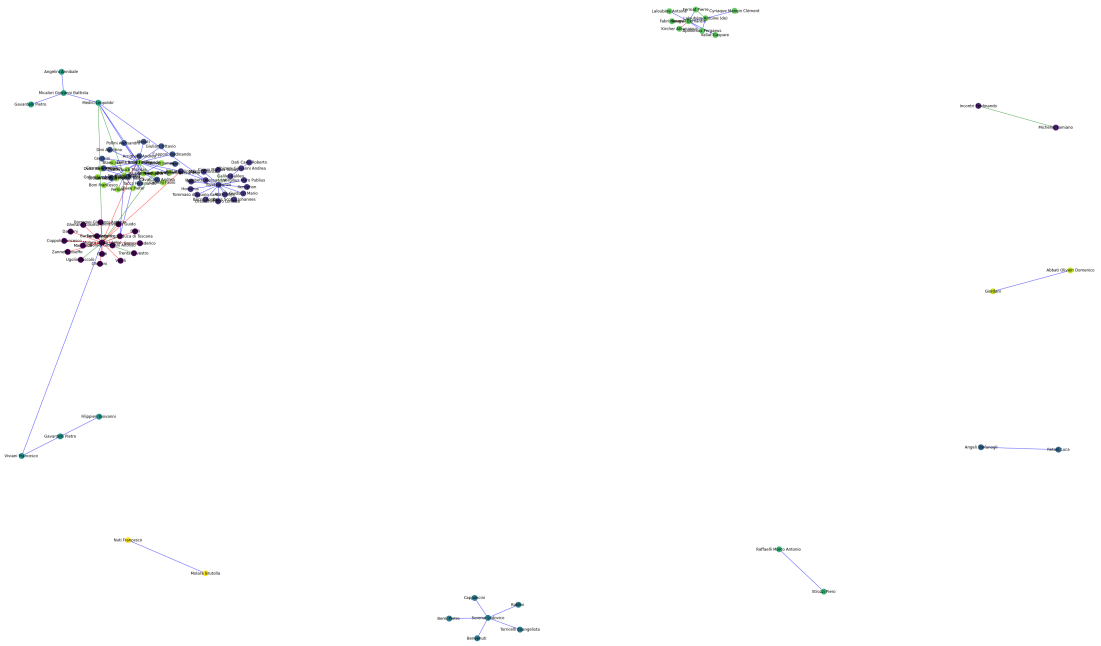

Figure 10: Clusters of the year 1662 for Viviani's network without him

To see how important is a correspondent to Viviani and his network, we find the clusters with and without this person to see how impactful he is and how he influences the connections. We can see this in figures 10 and 11.

In the following section we show and explain all the results obtained.

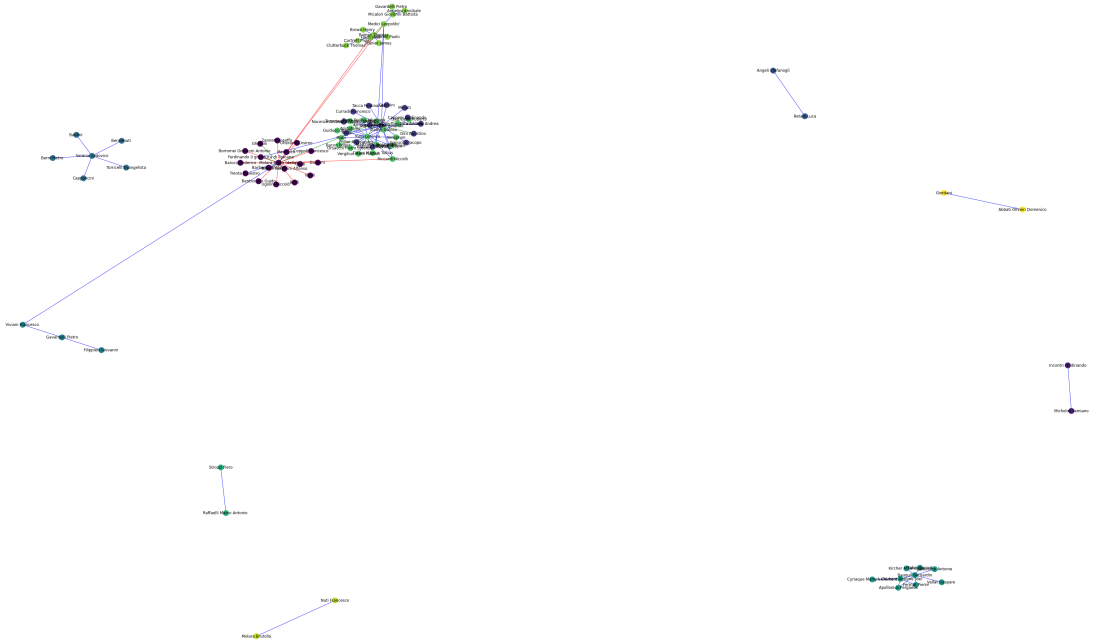

Figure 11: Clusters of the year 1662 for Viviani's network without him and Southwell

3 Results and Discussion

At the beginning of the project, we started from scratch. All we had was the "Biblioteca del Museo Galileo" website that contained the letters and its meta-data. Now we have a database ready to be used to create more histograms and Networks. In this section, we explain all the results and discuss it. We split our results in 3 categories as follows:

3.1 Database

First, let us take a look at our final database in figure 12. We have a total number of 2807 letters with known

	Expéditeur	Notes	References	Noms cités	Sujet	Titre	Date	Provenance	Destination	Destinataire	Year
0	Viviani Vincenzo	Tipologia: letteraIncipit: "Come posso non agg...	<td class="fieldValue"> <img border="0" src="ht...	[Torricelli Evangelista, Ricci Michelangelo]	NaN	Viviani Vincenzo a Ricci Michelangelo	20/05/1646	Firenze	Roma	Ricci Michelangelo	1646
1	Viviani Vincenzo	Tipologia: letteraIncipit: "Reputo a mio gran ...	<td class="fieldValue"> <img border="0" src="ht...	[Torricelli Evangelista, Renieri Vincenzo]	NaN	Viviani Vincenzo a Renieri Vincenzo	13/04/1647	Firenze	Pisa	Renieri Vincenzo	1647
2	Viviani Vincenzo	Tipologia: letteraIncipit: "L'esser io per spa...	<td class="fieldValue"> <img border="0" src="ht...	[Torricelli Evangelista, Ricci Michelangelo]	NaN	Viviani Vincenzo a Ricci Michelangelo	24/11/1648	Firenze	Roma	Ricci Michelangelo	1648
3	Viviani Vincenzo	Tipologia: letteraIncipit: "Haverei necessità ...	<td class="fieldValue"> <img border="0" src="ht...	[Cecchi Andrea , Arrighetti Andrea, Manetti Br...	NaN	Viviani Vincenzo a Manetti Braccio	18/10/1651	NaN	Firenze	Manetti Braccio	1651
4	Viviani Vincenzo	Tipologia: letteraIncipit: "Qui tutto giorno c...	<td class="fieldValue"> <img border="0" src="ht...	[Di Pietro , Viviani Alamanno]	Argomenti privati	Viviani Vincenzo a Viviani Alamanno	23/02/1652	Firenze	NaN	Viviani Alamanno	1652

Figure 12: Final Database

sender and receiver that can be used for quantitative treatment. We could have more but we dropped all the letters with unknown sender or receiver since the main goal of this project is to build and analyse Viviani's network. It did not make sense to keep these letters. We also have the year for most of the letters (2742 letters). 952 of these are sent by Viviani and 1847 are received. This is expected since we only work with the letters that has been conserved. Viviani has 508 unique correspondents in total from the 2807 letters that we analyse, between 1640 and 1701. 2122 of the letters has been sent or received from Florence, 551 from Roma, 208 from Paris and 327 from Pisa. This may indicate that most of his correspondents were from these cities.
The Database can also be combined with an external historical world map to have a different visualization of Viviani's network and its evolution and growth throughout the years. This has not been treated in this project as the main focus was the network itself.

This Database can be greatly increased by adding more galleries, as we only use Gal. 157 to 168, Gal. 252 to 258 and Gal. 275 to 286. There is actually in total 307 galleries while we only used the 28 that we are interested in for this project, the other galleries contain different letters about many different subjects, many are about pure mathematics, others about geometry, architecture, scientific letters, astronomy etc. To add other galleries the code that automatises the cleaning should be revisited and modified to handle all the new added content, as there are many special columns that are expected to need correction.
Adding more letters to our database allows us to do a better distant reading, as we have more information about Viviani's network and make a better representation of it. With a database this rich one can filter all the letters depending on the subject of the letter, the content or the correspondent and still expect to get a decent result. It is also possible to go the other way by using the database to find letters that cite certain persons, or subjects, using the reference column. The Database makes studying and analysis the historical archives quantitatively a lot more efficient.
In the next subsection we show and discuss the histograms obtained.

3.2 Histograms

Another good result that we get from this database is the histograms. As we can see on figure 6, we can easily find the top 9 correspondents of Viviani. The histogram shows that some of Viviani's correspondent were part of his network for all of his career, while others kind of disappeared, or were just present for a few years.

On this project we focus on Robert Southwell, one of his top correspondents. Looking at figure 6, from 1660 to 1662, he was his best correspondent, then they stopped exchanging letters until 1695, and 1697. This could mean that Southwell was not part of Viviani's network anymore, or that the letters between these years have not been secured. But this makes sense. These dates coincide with Southwell's arrival in Italy where he met Viviani. They spent some time together in Florence before he went to Rome. This is when they both started exchanging lots of letters. *"Viviani and Southwell seemed to develop a genuinely warm friendship. Between October 1660 and April 1663, during Southwell's stay in Rome, and then during his return voyage to England, they exchanged letters regularly on a variety of topics: medicine, the flooding of the rivers Tiber and Arno, and news of common acquaintances and new books [...] Viviani insisted on his wish for Southwell to write him at every stopover. Part of Viviani's motivation for making this insistence was to receive regular reports on the distribution of his book [...] After Southwell's arrival in England, the correspondence slowed down, but the two men kept in touch.[...] By the time he wrote to Viviani in 1665, he had married into another noble Protestant*

Irish family, increasing his wealth considerably and earning him a clerkship to the Privy Council.[...] Their correspondence slowed down again during the 1670s and stopped completely during the 1680s, as Southwell was elected to parliament.[...]Then in December 1690 he was elected President of the Royal Society, a position he kept until 1695. At this point he and Viviani re-started their correspondence, which they maintained until 1698, mostly exchanging news and information about new books, and discussing Southwell's sponsorship of Viviani for his election as a foreign fellow to the Royal Society in April 1696."[2] The histogram on figure 6 shows that the correspondence stops definitively at 1697.

Using the histogram alone can lead to erroneous conclusions. Combining the visual information with the historical context can give a better understanding of Viviani's correspondence with Southwell. For this purpose, the more letters we have in the Database, the more accurate the analysis becomes.

There could be other uses of the histograms depending on the information that we wish to visualize. We can see that from 1640 to 1655, he only exchanges a few letters per year. This is because he is only 18 years old in 1640. It coincides with the fact that in 1639 he moved to Galileo's home in Arcetri to assist him in his studies. He was not known yet and just started his career. They were joined by Evangelista Torricelli in 1641. After his death in 1647, Viviani was appointed to fill his position at the Accademia Dell'Arte del Disegno in Florence. A decade later, Viviani became one of the first members of the Grand Duke's experimental academy, the Accademia del Cimento, when it was created. This is when he started to have more responsibilities and his career flourished.[1] This is also projected on the number of letters he started to send and received throughout these years. We can see that he has many periods of high correspondence followed by periods with a lot less. This also makes sense as he needed years to develop his ideas then share it with his network.
Another analysis could be done with histograms for the letters sent and received from certain cities (Roma, Florence, Paris etc.). We could also filter the letters depending on the subject to have a better visualization on the topics discussed in the letters, to better understand his priorities with his network, to know which subjects he discussed most during the years, and most importantly with who. This combined with the Database and an interactive map can lead to a more in depth analysis of Viviani's network.

3.3 Networks

In this subsection, we explain the results obtained with our networks and discuss how this method can be used. As explained before, we use 3 different weight initialization for the edges of our networks. The influence of the

weight initialization changes the look of the network obtained. With a smaller weight initialization all the nodes are considered equal, even the ones for the persons that has only been cited and did not participate. The result was not satisfying as many persons were only cited or participated once. An initialization with a higher weight for the participating persons made more sense, as these persons are more important to us since they sent or received the letters. With an initialization of 5 and 10, we can see more red and green nodes around Viviani. If a blue node is also close then it means that this person has been cited a lot and may be important too to Viviani's network, life and career. We finally decide to keep a weight initialization of 5 for the clusters as we explain later.

We can see that many of the closest persons are also connected with red and green edges. This means that these persons are part of the same network since they are all connected. If we delete Viviani's node from the graph we can see that there are many single nodes remaining, as these have been directly cited by him in letters. The nodes that remain connected even after deleting him are part of the networks. The number of connected components obviously changes at this point since everyone is connected to Viviani, because of the letters he received. What is interesting is what happens when we do the same with Southwell Robert. If the number of connected components changes, then the new components are not part of the network, if it does not change then the persons who were connected to him and himself are part of Viviani's network and are important. We can see that in the year 1662, when they exchanged the most, he is connected to a small cluster of persons (Blaeu Joan[9], Della Torre Filippo[10], Boni Francesco etc.) who all have flourished careers, one is a cartographer from the Netherlands, one an archaeologist from Italy, one a lawyer, publisher and politician from the Netherlands. This coincides with the fact that in 1662, he was still traveling in Europe to develop his own network and his educational goals while on the tour. Since these people were not necessarily connected to Viviani, they necessarily become single nodes when we delete Southwell from the graph. We can see that the network alone is not enough to fully understand the importance of people in Viviani's network and the role that they play, but combined with the historical context, we can get a far better understanding.

About the clusters, we can see that by removing Viviani, we get many different clusters. We can finally justify the choice of the initialization of the weights since by taking an initialization of 10, the graphs are too far and the it becomes harder to read. The choice of 5 seems to be optimal from a practical perspective and allows a good reading.

We can see from figure 10 that we have many clusters after deleting Viviani, which is expected. Some of the cluster have blue edges, meaning that the members of the clusters are a correspondent and a cited-only person (or many). We can also see that we sometimes have clusters with red or green edges, meaning that this cluster is itself a small network of Viviani. After deleting Southwell from the graph and redrawing the clusters, we can see that for the year 1662, all the nodes that were cited by him disappear, meaning that he is not part of any

network. This again makes sense as in 1662 he already left Florence and continued his travel in Europe[2] as explained before, so he was not part of any of Viviani's existing networks.

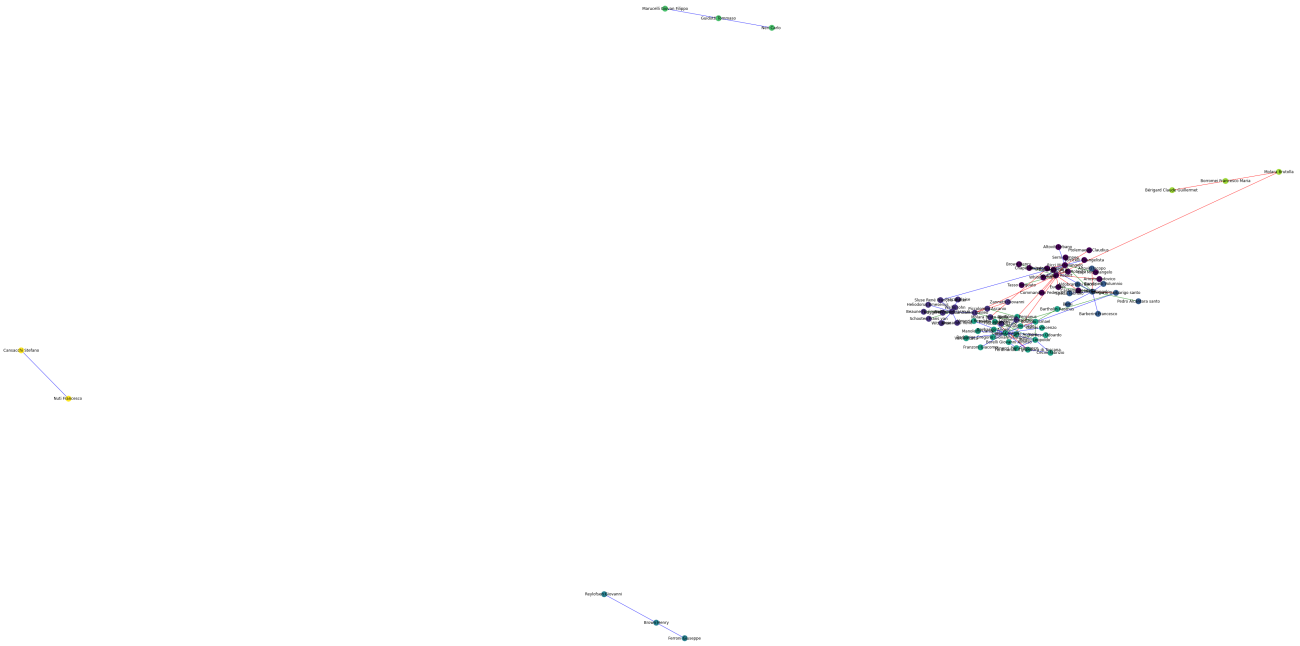

Figure 13: Clusters of the year 1660 with Southwell and Without Viviani

However, in the year 1660, we can see that he is part of a cluster containing many other correspondents of Viviani, as we can see in figure 13. He has many red edges in his cluster which means that they were all part of the same network at this time. This coincides with the time period when he was in Italy working with Viviani[2]. It makes sense and it is exactly what we expect. We can also see green edges in the same cluster which means that both persons cited each others in letters from or to Viviani. This again confirms the fact that during 1660, Robert Southwell was a close person to Viviani. We can see from figure 14 that when we delete Southwell from the graph, We have a new cluster containing Molara Brutolla, Borromei Francesco Maria and Bérigard Claude Guillermet. These are all correspondents to Viviani. This means that he was part of two of Viviani's networks at this time and proves his importance. We also see no blue edge connected to Southwell which confirms what we said before.

This kind of interpretation and analysis can be done for any of Viviani's correspondents as long as the historical context is known, as it justifies the changes that can happen whenever we delete a certain node from the graph.

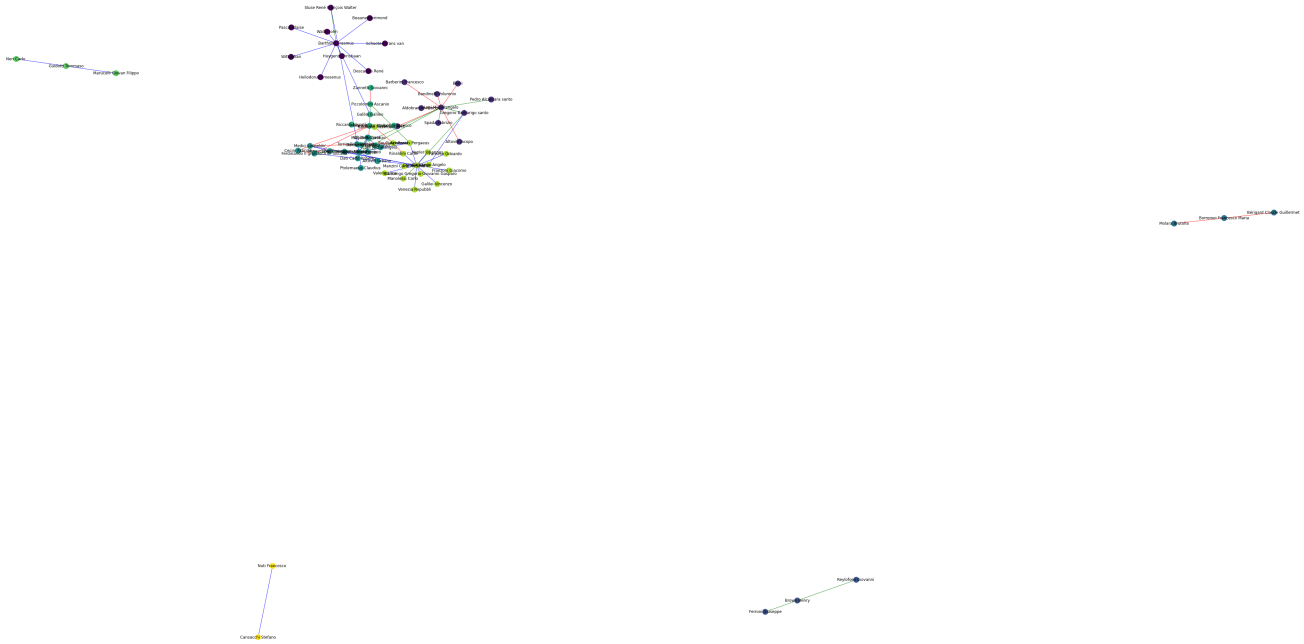

Figure 14: Clusters of the year 1660 without Southwell and Viviani

4 Conclusion

Viviani developed his network and reputation brilliantly. He started as a pupil of Galileo and ended up years later at his same position[1]. He did it step by step, starting from his close friends and colleagues then extended even outside of Europe. We saw from our results that Southwell had a non-negligible impact on this since he transported Viviani's work outside of Italy[2] and participated in his growth as a renown mathematician and thinker. The graphs and clusters show that he was definitely part of Viviani's network and also showed the other members of the different networks throughout the years.

The Database can be used in many different ways to get more statistics, histograms and graphs that explain the evolution of Viviani's network. An upgrade to this project could be done by making an interactive map using the database to follow the evolution of the clusters and networks throughout the years.

This analysis shows that the historical archives allows us to visualize Viviani's career from a different perspective. Combined with the historical context we can reconsider elements of his life. With even more letters from the "Biblioteca Del Museo Galileo" we can even push the analysis further by implementing a neural network that would try to predict the unknown correspondents of letters based on the subject, the content, the provenance/destination etc. We could try to complete the meta-data whenever it is not available. This could greatly help the analysis of Viviani's network and life in a different way by literally adding new information instead of the classical interpretation of what is already existing and known, as it may give more information that is currently unavailable in the history books and archives. In this context, the quantitative analysis could, combined with a good historical knowledge, help discover even more secrets about Viviani's life.

This project proves that Network Analysis is important and can add new layers to our actual historical knowledge.

References

[1] J J O'Connor and E F Robertson. Vincenzo viviani. *MacTutor History of Mathematics Archive*, 2009.

[2] Luciano Boschiero. "robert southwell and vincenzio viviani: Their friendship and an attempt at italian–english scientific collaboration.". *Parergon*, 26(2):87–108, 2009.

[3] Biblioteca del museo galileo. `https://bibdig.museogalileo.it/Teca/Viewer?an=000000018000`.

[4] Beautifulsoup documentation. `https://beautiful-soup-4.readthedocs.io/en/latest/`.

[5] Matplotlib documentation. `https://matplotlib.org/`.

[6] Networkx documentation. `https://networkx.org/`.

[7] Stephen G. Kobourov. "spring embedders and force-directed graph drawing algorithms". 2012.

[8] Jean-Loup; Lambiotte Renaud; Lefebvre Etienne Blondel, Vincent D; Guillaume. "fast unfolding of communities in large networks.". *Journal of Statistical Mechanics: Theory and Experiment.*, 10, 9 October 2008.

[9] Joan blaeu. `https://en.wikipedia.org/wiki/Joan_Blaeu`.

[10] Filippo del torre. `https://fr.wikipedia.org/wiki/Filippo_Del_Torre`.