

Lab 1

Descriptive Network Analysis – Collecting and Visualizing Data

Agenda

- Assignment context
- Collecting network data
- Visualizing network data
- Q&A

FORBES > INNOVATION > AI

Collective Intelligence Is About To Disrupt Your Strategy: Are You Ready?

Ganes Kesari Contributor 

Data science advisor, fractional CDAO, TEDx speaker, corporate trainer

Follow

Apr 28, 2021, 06:50am EDT

vs.

Harvard Business Review

Subscribe

Sign In

AI And Machine Learning

ChatGPT and How AI Disrupts Industries

by Ajay Agrawal, Joshua Gans, and Avi Goldfarb

December 12, 2022

Collective Intelligence



is the knowledge that arises when **people join forces, collaborate, and work together** to make decisions. It is shared or group knowledge.

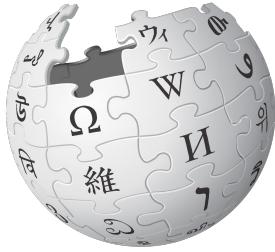
vs.

Artificial Intelligence



is the simulation of **human intelligence by machines or softwares**. Basically, machines learn how to learn, recognise speech patterns, process information etc. in the same way as the human brain.

Part I: Collecting Network Data



Quora

StackExchange 

 stackoverflow

vs.



ChatGPT

Step 1: Choose a topic for your text

You can decide on the topic based on personal interests, research interests, or popular topical areas, among others.

Example: Medici Family in Florence during the Renaissance



Step 2: Collect text from a collective intelligence source

For the collective intelligence sources (e.g., Wikipedia), you want to find a single webpage that contains an overview of your topic.

Example collective intelligence source:

Wikipedia → Florence (sections)

- Rise of the Medici
- Savonarola, Machiavelli, and the Medici popes

The screenshot shows the Wikipedia article for "Florence". The title is "Florence" with a red link to the disambiguation page. The main content discusses the city's history as the capital of Tuscany and its role in the Renaissance. It mentions the Medici family, Savonarola, Machiavelli, and the Uffizi Gallery. The page also highlights Florence's status as a World Heritage Site and a fashion capital. On the right side, there is a sidebar with sections for "Florence", "Comune", "Comune di Firenze", and images of the city skyline and famous landmarks like Palazzo Pitti, Santa Croce, and the Uffizi Gallery.

Step 3: Collect text from ChatGPT using similar prompts

For ChatGPT, you should query it to collect similar information.

Example ChatGPT prompts:

- Tell me about relationships built by the Medici family during the Renaissance in Florence.
- Tell me more about the links between the Medici and artists.

Default (GPT-3.5)

NJ Tell me about relationships built by the Medici family during the Renaissance in Florence.

The Medici family, one of the most influential and powerful dynasties during the Renaissance in Florence, played a pivotal role in shaping not only the political and economic landscape of the city but also the cultural and artistic developments of the era. Their relationships, both within Florence and across Europe, were instrumental in their rise to prominence.

1. **Political Alliances in Florence:** The Medici family established strong political relationships within Florence by strategically marrying into other influential families. Cosimo de' Medici, for example, married Contessina de' Bardi, which helped solidify his position in the city's ruling elite. The Medici often formed alliances with other influential families, both through marriage and patronage, to maintain their political power.

2. **Papal Connections:** The Medici family had close ties to the Catholic Church, which were essential in advancing their interests. Lorenzo de' Medici, known as Lorenzo the Magnificent, used his influence to

Send a message ➤

Step 4: Save your text in .txt or .rtf files

At the height of demographic expansion around 1325, the urban population may have been as great as 120,000, and the rural population around the city was probably close to 300,000.[21] The Black Death of 1348 reduced it by over half,[22][23] about 25,000 are said to have been supported by the city's wool industry: in 1345 Florence was the scene of an attempted strike by wool combers (ciompi), who in 1378 rose up in a brief revolt against oligarchic rule in the Revolt of the Ciompi. After their suppression, Florence came under the sway (1382–1434) of the Albizzi family, who became bitter rivals of the Medici.

In the 15th century, Florence was among the largest cities in Europe, with a population of 60,000, and was considered rich and economically successful.[24] Cosimo de' Medici was the first Medici family member to essentially control the city from behind the scenes. Although the city was technically a democracy of sorts, his power came from a vast patronage network along with his alliance to the new immigrants, the gente nuova (new people). The fact that the Medici were hankers to the pope also contributed to their ascendancy. Cosimo was succeeded by his son, who soon after, succeeded by Cosimo's grandson, Lorenzo in 1469. Lorenzo was a great commissioner of works by Michelangelo, Leonardo da Vinci and Botticelli. Lorenzo was a poet and musician and brought composers and singers to Florence, including Alessandro Ghiselin, and Heinrich Isaac. By contemporary Florentines (and since) "Lorenzo the Magnificent" (Lorenzo il Magnifico).

Following Lorenzo de' Medici's death in 1492, he was succeeded by his son Piero II. King Charles VIII invaded northern Italy, Piero II chose to resist his army. Because of the size of the French army at the gates of Pisa, he had to accept the humiliating terms of the French king. These made the Florentines rebel, and they expelled Piero II. With the first period of Medici rule ended with the restoration of a republican government.

During this period, the Dominican friar Girolamo Savonarola had become prior of a monastery in 1490. He was famed for his penitential sermons, lambasting what he saw as immorality and attachment to material riches. He praised the exile of the Medici, punishing them for their decadence. He seized the opportunity to carry through

The Medici family played a significant role in the Renaissance in Florence, not only as patrons of the arts and sciences but also in their establishment of important political and social relationships. Through strategic marriages, alliances, and networks, the Medici family expanded their influence and solidified their power within Florence and beyond.

One of the most notable relationships built by the Medici family was their association with the papacy. Giovanni di Bicci de' Medici, the founder of the Medici Bank, was instrumental in forging close ties with the Catholic Church. His son, Cosimo de' Medici, continued this tradition and cultivated relationships with several popes, most notably Pope Nicholas V. The Medici popes, such as Leo X (Giovanni di Lorenzo de' Medici) and Clement VII (Giulio de' Medici), further strengthened the family's connection to the papacy and helped promote their interests.

The Medici family also formed alliances through marriage. Cosimo de' Medici's marriage to Contessina de' Bardi not only secured him a considerable dowry but also connected him to influential Florentine families. Similarly, Lorenzo de' Medici, known as Lorenzo the Magnificent, married Clarice Orsini, a member of the powerful Orsini family from Rome. These unions brought prestige, political alliances, and expanded the Medici family's reach in both Florence and Rome.

Furthermore, the Medici family developed relationships with prominent artists, intellectuals, and scholars of the Renaissance. Cosimo de' Medici was a great patron of the arts and attracted renowned artists like Donatello and Filippo Brunelleschi to Florence. Lorenzo the Magnificent continued this tradition, fostering relationships with prominent figures such as Sandro Botticelli, Leonardo da Vinci, and Michelangelo. These associations not only enhanced the Medici family's cultural influence but also strengthened their social and political standing.

The Medici family's relationships extended beyond Italy as well. They established banking connections throughout Europe, which helped solidify their economic power and political influence. Their banking activities and financial partnerships allowed them to cultivate relationships with influential families in countries like France, England, and the Holy Roman Empire. These connections also facilitated the Medici family's engagement in diplomatic affairs and international trade.

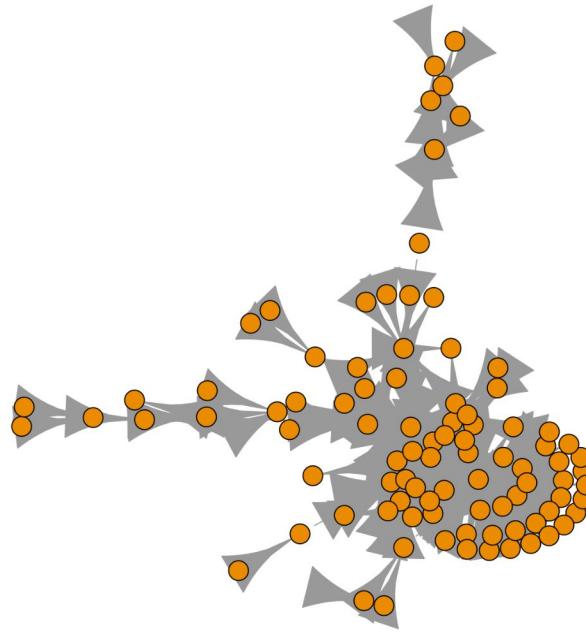
Step 5: Generate Semantic Networks

Some Details to Note:

- The raw text is initially “cleaned” by:
 - Removing **stop words** (words that are frequent but provide little information – e.g., “I”, “the”, “we’ll”, “it’s”, etc.)
 - **Lemmatization** (grouping together different inflected forms of the same word – e.g., the lemma of “ran”, “runs”, and “running” is simply “run”)
- The network ties represent words that co-occur in the text **within 10 words** of each other, excluding **rare word co-occurrences** (those that occur less than 3 times in the full body of text)
- **LDA** is used to group the words into 3 topics

Part II: Visualizing Network Data

Network Visualization



Network Visualization

Please check out this page: <https://kateto.net/network-visualization>

The screenshot shows a dark-themed website for Katya Ognyanova. At the top, there is a large, abstract network visualization consisting of numerous overlapping circles in shades of red, teal, and grey. Below this, the name "Katya Ognyanova" is displayed in a white serif font. A horizontal navigation bar follows, containing links for "Home", "Blog", "Bio", "Research", "Publications", "Teaching", "Tutorials", "Media", and a search icon represented by a magnifying glass. The "Research" link is currently highlighted in white.

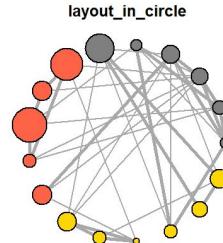
Static and dynamic network visualization with R

June 14th, 2017 | Tags: animation, centrality, D3, dataset, graph, igraph, JavaScript, map, ndtv, network, network analysis, network science, network tutorial, network visualization, networkD3, node, plot, R, reciprocity, RStudio, sna, Statnet, threejs, transitivity, visNetwork, visualization

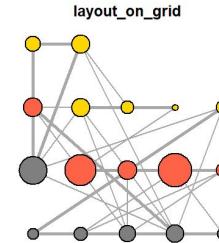
Comparing Layout Algorithms

Different stochastic algorithms attempt to arrange nodes and links in a “useful” way.

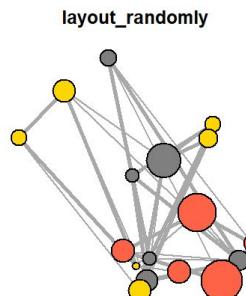
Results vary based on your network ... and how many times you run the algorithm.



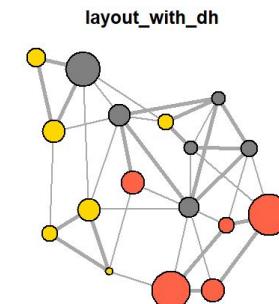
Nodes in a circle
(evenly spaced)



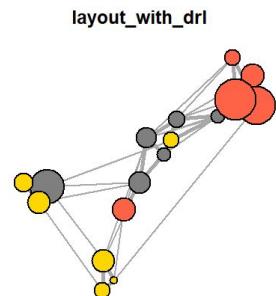
Nodes spaced out
on a uniform grid



Nodes placed
completely at
random



Davidson-Harel
Algorithm: Stochastic
optimization of edges



Distributive Recursive
Layout Algorithm: Kind of
like “springs”

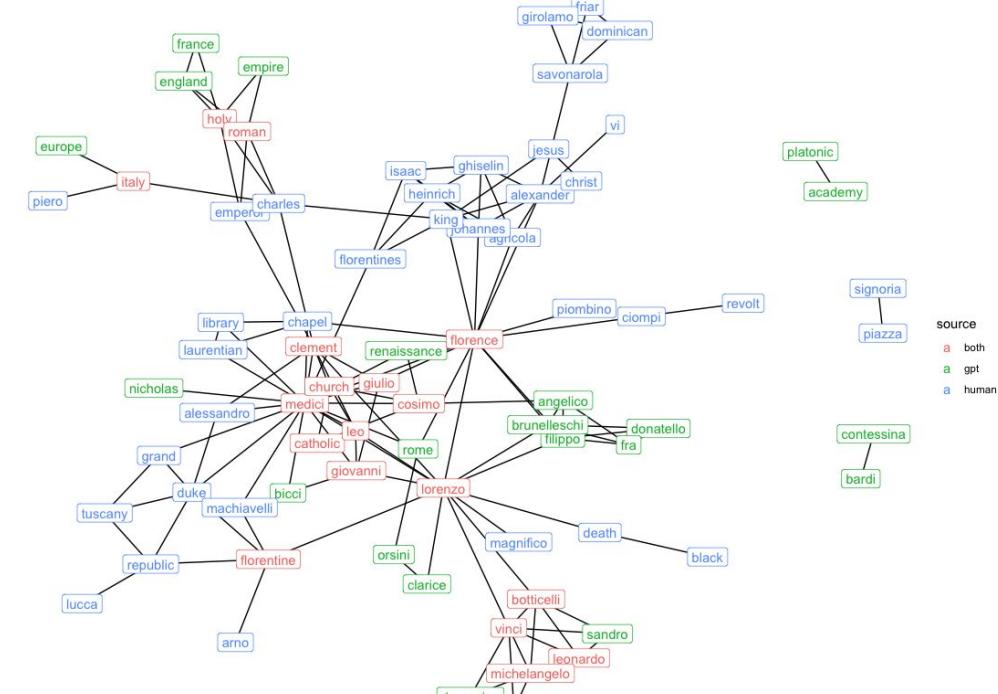
Layouts using igraph

Tutorial by Katherine Ognyanova

<https://kateto.net/network-visualization>

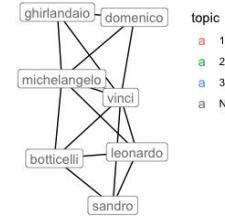
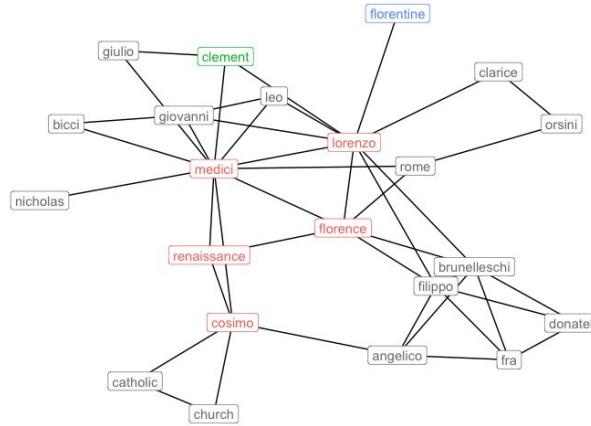
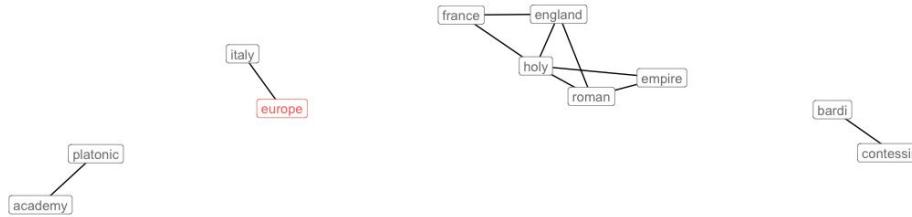
Interpreting the Network Visualization

Combined Network



Interpreting the Network Visualization

GPT Network

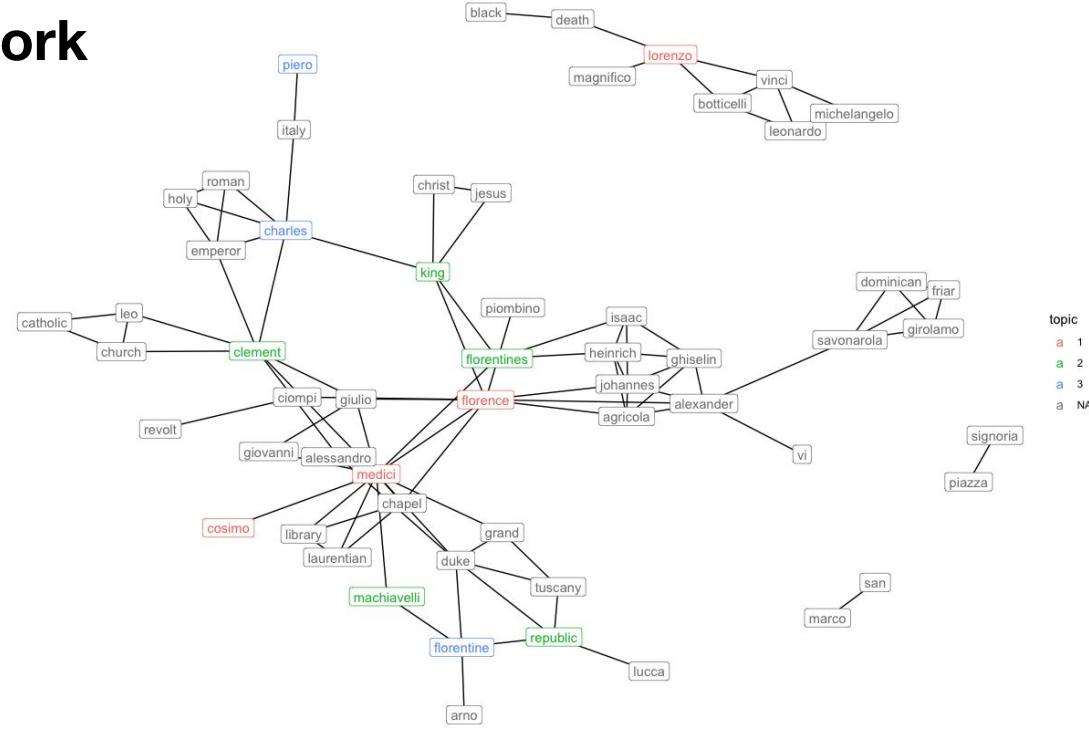


topic

- a 1
- a 2
- a 3
- a NA

Interpreting the Network Visualization

Wikipedia Network



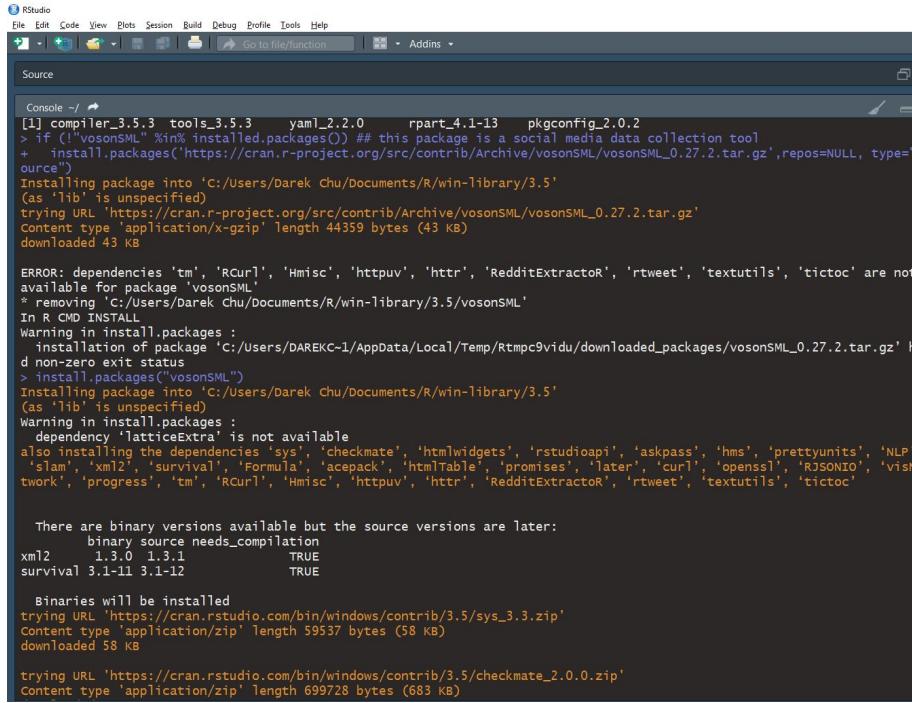
Some Tips

General Tips

- Collect data as soon as possible (**0 points for not collecting data**)
- Make sure to run the code and confirm that it works
 - Execute the code chunk by chunk - Don't execute the entire script!
 - For Windows users, Ctrl+Enter
 - For Mac users, Cmd+Return
- Once you collect data, save it as an RData file
- Focus on interpreting results

When Asking Questions

- Please include screenshots:



```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function | Addins

Source
Console ~/ ~
[1] compiler_3.5.3 tools_3.5.3 yaml_2.2.0 rpart_4.1-13 pkgconfig_2.0.2
> if (!"vosonSML" %in% installed.packages()) ## this package is a social media data collection tool
+ install.packages("https://cran.r-project.org/src/contrib/Archive/vosonSML/vosonSML_0.27.2.tar.gz", repos=NULL, type="source")
Installing package into 'C:/Users/Darek Chu/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
trying URL 'https://cran.r-project.org/src/contrib/Archive/vosonSML/vosonSML_0.27.2.tar.gz'
Content type 'application/x-gzip' length 44359 bytes (43 KB)
downloaded 43 KB

ERROR: dependencies 'tm', 'RCurl', 'Rmisc', 'httpuv', 'httr', 'RedditExtractor', 'rtweet', 'textutils', 'tictoc' are not
available for package 'vosonSML'
* removing 'C:/Users/darek Chu/Documents/R/win-library/3.5/vosonSML'
In R CMD INSTALL
Warning in install.packages :
  installation of package 'c:/Users/DAREK-1/AppData/Local/Temp/Rtmpc9vidu/downloaded_packages/vosonSML_0.27.2.tar.gz' had
non-zero exit status
> install.packages("vosonSML")
Installing package into 'C:/Users/Darek Chu/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
Warning in install.packages :
  dependency 'latticeExtra' is not available
also installing the dependencies 'sys', 'checkmate', 'htmlwidgets', 'rstudioapi', 'askpass', 'hms', 'prettyunits', 'NLP',
'slam', 'xml2', 'survival', 'Formula', 'acepack', 'htmlTable', 'promises', 'later', 'curl', 'openssl', 'RJSONIO', 'visNetwork',
'progress', 'tm', 'RCurl', 'Rmisc', 'httpuv', 'httr', 'RedditExtractor', 'rtweet', 'textutils', 'tictoc'

There are binary versions available but the source versions are later:
  binary source needs_compilation
xml2      1.3.0  1.3.1      TRUE
survival   3.1-11 3.1-12      TRUE

Binaries will be installed
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/sys_3.3.zip'
Content type 'application/zip' length 59537 bytes (58 KB)
downloaded 58 KB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/checkmate_2.0.0.zip'
Content type 'application/zip' length 699728 bytes (683 KB)
```

Common Error

- If you're running into this error:

```
Error in separate_wider_delim(unnest_tokens(pivot_longer(text, cols = c(gpt, :  
could not find function "separate_wider_delim"
```

- Run **packageVersion("tidyverse")**
 - If you have anything less than **version 2.0.0**, please run **install.packages("tidyverse")**, then restart R, and run the code again from the beginning

Thank You!