Translating Networks: Assessing Correspondence Between Network Visualisation and Analytics

Martin Grandjean, Mathieu Jacomy

Translating Networks

Assessing correspondence between network visualisation and analytics

Digital Humanities 2019, Utrecht, Netherlands

Martin Grandjean (1) and Mathieu Jacomy (2)
1. University of Lausanne, Switzerland
2. Aalborg University, Denmark

Network interpretation is a widespread practice in the digital humanities, and its exercise is surprisingly flexible. While there is now a wide variety of uses in different fields from social network analysis (Ables et al., 2017) to the study of document circulation metadata (Grandjean, 2016) or literature and linguistic data (Maryl and Elder, 2017), many projects highlight the difficulty of bringing graph theory and their discipline into dialogue. Fortunately, the development of accessible software (Bastian et al., 2009), followed by new interfaces (Rosa Pérez et al., 2018; Wieneke et al., 2016), sometimes with an educational dimension (Beaulieu, 2017; Xanthos et al., 2016), has been accompanied in recent years by a critical reflection on our practices (Weingart, 2011; Kaufman et al., 2017), particularly with regard to visualisation. Yet, it often focuses on technical aspects.

In this paper, we propose to shift this emphasis and address the question of the researcher's interpretative journey from visualisation to metrics resulting from the network structure. Often addressed in relation to graphical representation, when it is not used only as an illustration, the subjectivity of translation is all the more important when it comes to interpreting structural metrics. But these two things are closely related. To separate metrics from visualisation would be to forget that two historical examples of network representation, Euler (1736) and Moreno (1934), are not limited to a graphic reading (the term "network" itself would only appear in 1954 in Barnes' work). In the first case, the demonstration was based on a degree centrality measurement whereas in the second case the author made the difference between "stars" and "unchosen" individuals while qualifying the edges as inbound and outbound relationships.

This is why this paper propose to examine the practice of visual reading and metrics-based analysis in a correspondence table that clarifies the subjectivity of the *translation* while presenting possible and generic interpretation scenarios.

Visual approach: making the global structure readable

The way we read networks has changed over time. Historically the question of network readability was asked in terms of aesthetic criteria. In the word of Jacob Moreno "the fewer the number of lines crossing, the better the sociogram". Even in the nineties, when giving birth to the modern layout algorithm, Früchterman and Reingold (1991) aimed at "minimizing edge crossings" and "reflecting inherent symmetry". However, these criteria do not seem so crucial to practices observed nowadays in digital humanities (and beyond).

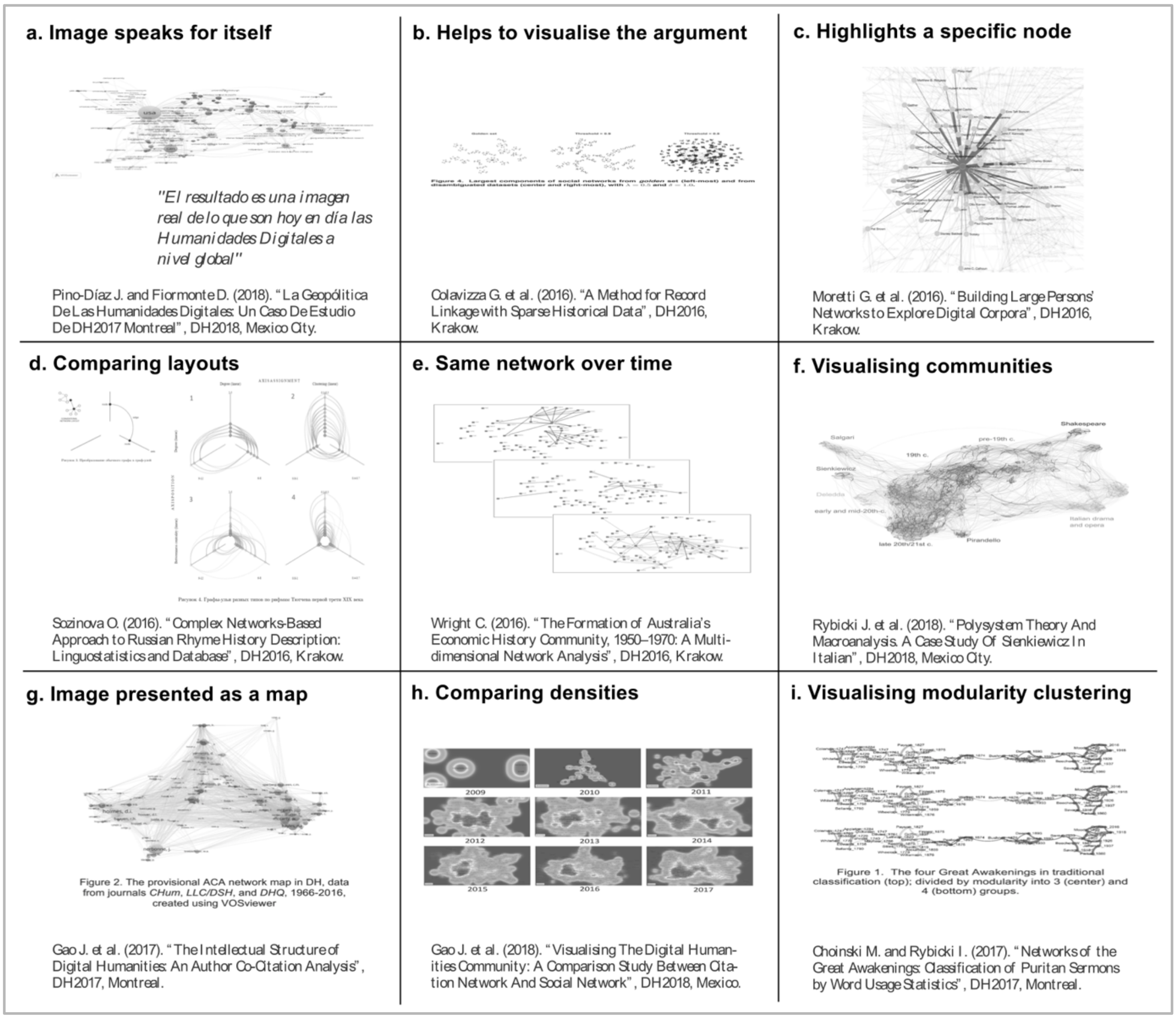


Fig. 1 Different contexts for network visualisation in DH2016, DH2017 and DH2018 abstracts.

Looking at recent papers in digital humanities, networks appear to have a wide range of usages. Their visualisations are either self-sufficient [fig. 1.a.] (Algee-Hewitt, 2018; Pino-Diaz and Fiormonte, 2018; Verhoeven et al., 2018; Marraccini, 2017), an optional help to understanding [fig 1.b.] (Colavizza et al., 2016) or strongly connected to the text. Some authors use them to highlight the position of a specific node [fig. 1.c.] (Moretti et al., 2016), to compare layouts [fig. 1.d.] (Sozinova, 2016) or the layout of the same graph in time [fig. 1.e.] (Wright, 2016). They may aim at visualising communities [fig. 1.f.] (Rybicki et al., 2018; Torres-Yepez and Zreik, 2018), mapping a general structure [fig. 1.g.] (Gao et al., 2017), tracking density patterns [fig. 1.h.] (Gao et al., 2018) or monitoring algorithms like modularity clustering [fig. 1.i.] (Choinski and Rybicki, 2017). These usages reveal a different perspective in network visualisation where we expect the visual to *translate* underlying relational structures.

It helps to give different names to these two different approaches. We call **diagrammatic** the perspective where the network is a diagram that we read by following paths. We do not want the edges to cross and we use aesthetic criteria to bring clarity. It was Moreno's perspective, and is still relevant to small networks and local exploration. Then we call **topological** the perspective where the network is a structure that we read by detecting patterns. We expect the visualisation to help us retrieve structural features like clustering or centralities. It is a common practice in digital humanities, more holistic and relevant to larger networks. Aside or in

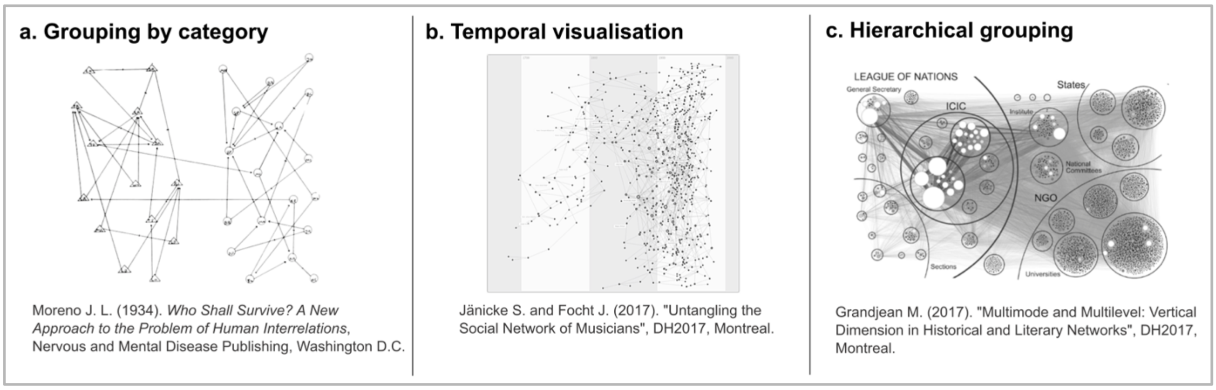


Fig. 2 Various layouts do not follow a force-driven algorithm to make non-relational dimensions of the data explicit.

complement, **classic data visualisation** is also employed to visualise non-relational structures (node attributes, etc.).

In the topological perspective, a standard procedure is to assign nodes a position using a force-driven algorithm. This family of algorithms is known for displaying clusters that match a widely used measure of community detection, modularity clustering (Noack, 2009). Its translation remains however difficult to interpret locally, as we can never give a simple explanation for a node's position. Classic data visualisation also translates non-relational structures, by itself or combined with a relational perspective. Different structural features may require different visualisations: the examples of fig. 2 shows curated visualisations using categories [fig. 2.a boys and girls, in the famous example of (Moreno, 1934)], temporality [fig. 2.b] (Jänicke and Focht, 2017) or hierarchy [fig. 2.c] (Grandjean, 2017). Though very different from force-driven placement, they display better certain structural features.

Objectifying the structure with metrics

Often opposed to visual interpretation, of which they would be a more objective and reliable representation, centrality measures have a history that goes back to more than half a century and shows that they are not immutable and require constant adaptation to usage Moreover, Freeman (1979) insists on the fact that the notion of "centrality" is the result of several intuitive conceptions. To remind that these metrics are based on "intuition" means to recognize that they have no meaning in themselves and that their interpretation must be rediscussed - and therefore translated - according to the context. This paper thus proposes to list and evaluate to which extent these metrics are applicable to humanities and social science data and can, if necessary, be "translated" into this language to complement visual analyses.

Global properties

Statistical analysis allows for comparing networks across multiple dimensions at once (Tank and Chen, 2017). For instance, comparing the **number of nodes and edges** of different graphs of the same type (Trilcke et al., 2016) can be a ranking tool that is directly translatable into natural language. In addition to that, studies suggest that **density** (the number of edges in relation to the number of nodes) is relevant to analyse character networks, especially when compared within a homogeneous collection (Evalyn and Gauch, 2018; Grandjean, 2015). This is also the case when measuring **average path length** (Trilcke et al., 2016).

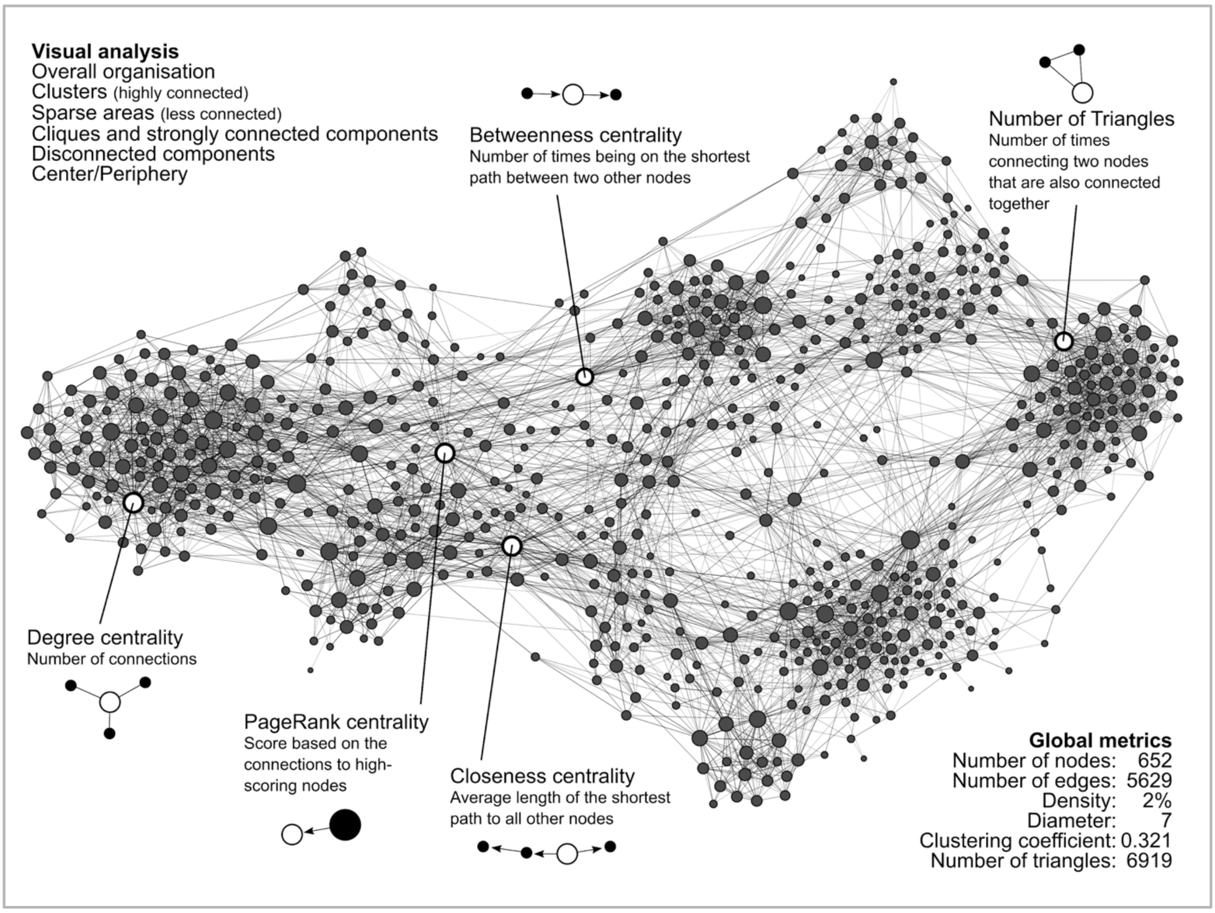


Fig. 3 Three levels of interpretation that can be articulated: visual analysis (examples top left), use of global metrics (examples bottom right) and use of local metrics (highlighted nodes).

Local properties

With regard to local measures, the **degree** (number of neighbouring nodes) is the simplest **centrality**, and the only one systematically used between the late 1950s and early 1970s, before the development of more diversified metrics (Freeman, 1979). Its simplicity allows for a transparent translation: in a literary network, for example, it counts the number of times one character speaks to another (Jannidis et al., 2016).

The notion of **betweenness centrality** disrupts the conception of what the "centre" of a network may consist of. Its ability to reveal structural elements bridging large, immediately visible clusters makes it popular in the social sciences since the emergence of Granovetter's concept of "weak ties" (Granovetter, 1973). Betweenness is very closely linked to the notion of circulation: it counts the shortest paths to detect intermediate "bridges" or "key passages" capable of opening or locking certain parts of the network (Tayler and Neugebauer, 2018). Depending on applications, these are therefore both positions of power and vulnerable places.

The **closeness centrality** allows to highlight the "geographical" middle of the graph. In networks of a certain density and when they are not divided into several distinct communities, the closeness is generally fairly evenly distributed and allows a good translation of the notions of "center" and "periphery".

For its part, the **eigenvector centrality** is quite complicated to translate since it works iteratively and is very much dependent on the structural context at short and medium range around a node. "Prestige" or "influence" centrality, named "power" centrality by its author (Bonacich, 1972), it qualifies a node's environment while operating in cascade: a well-connected node gives its neighbours a part of its authority capital, and so on. It is therefore particularly useful when trying to analyse the hierarchy of the nodes in a graph (Piper et al., 2017). The most well-known use of this measure is the backbone of the Google search engine: the PageRank algorithm (Brin and Page, 1998).

Towards mixed approaches

This contribution proposes a table of correspondence between the concepts of graph theory and the practice of visual network analysis in the social science and humanities. This effort must not be understood as a demarcationist attempt at telling the right method from the wrong. The "dictionary" is not exhaustive and only aims at helping to bridge two worlds that have more in common that what meets the eye. By focusing on *translating* methods, we want to stress that crossing points are real even though they do not come without issues, and thus require our methodological attention.

We also note that the analysis should not be limited to a catalogue of well-known methods (basic centralities, etc.) but that approaches combining several of those should be encouraged to obtain an optimal and innovative "translation". In this way, we could compare metrics (Escobar and Schauf, 2018) or combine them to establish rankings (Fischer et al., 2018; Grandjean, 2018: 328). Furthermore, the enrichment of the networks by means of categories that are not dependent on the structure, like the gender of individuals in a social network (Dunst and Hartel, 2017) or the discipline of projects in a scientometric analysis (Grandjean et al., 2017), allows to test translation and interpretation hypotheses by avoiding the blind approach of testing all possible graph metrics.

References

Ables M. et al. (2017). Using Archival Texts to Create Network Graphs of Musicians in Early Modern Venice, *Digital Humanities 2017*, Montreal.

Algee-Hewitt M. A. (2018). The Hidden Dictionary: Text Mining Eighteenth-Century Knowledge Networks, *Digital Humanities 2018*, Mexico City.

Barnes J. A. (1954). Class and Committees in a norwegian Island Parish, *Human Relations*, 7, 39-58.

Bastian M. et al. (2009). Gephi: an open source software for exploring and manipulating networks, *International AAAI Conference on Weblogs and Social Media*, 361-362.

Beaulieu M.-C. (2017). Perseids and Plokamos: Weaving pedagogy, data models and tools for social network annotation, *Digital Humanities 2017*, Montreal.

Bonacich P. (1972). Factoring and weighting approaches to status scores and clique identification, *The Journal of Mathematical Sociology*, 2, 1, 113-120.

Brin S. and Page L. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine, *Seventh International World-Wide Web Conference*, Brisbane.

Choinski M. and Rybicki I. (2017). Networks of the Great Awakenings: Classification of Puritan Sermons by Word Usage Statistics, *Digital Humanities 2017*, Montreal.

Colavizza G. et al. (2016). A Method for Record Linkage with Sparse Historical Data, *Digital Humanities 2016*, Krakow.

Dunst A. and Hartel R. (2017). Corpora and Complex Networks as Cultural Critique: Investigating Race and Gender Bias in Graphic Narratives, *Digital Humanities 2017*, Montreal.

Escobar Varela M. and Schauf A. (2018). Network Analysis Shows Previously Unreported Features Of Javanese Traditional Theatre, *Digital Humanities 2018*, Mexico City.

Euler L. (1736). Solutio Problematis ad Geometriam Situs, *Opera Omnia*, 7, 128-140.

Evalyn L. and Gauch S. (2018). Analyzing Social Networks Of XML Plays: Exploring Shakespeare's Genres, *Digital Humanities 2018*, Mexico City.

Fischer F. et al. (2018). To Catch A Protagonist: Quantitative Dominance Relations In German-Language Drama (1730–1930), *Digital Humanities 2018*, Mexico City.

Freeman L. C. (1979). Centrality in Social Networks: Conceptual Clarification, *Social Networks*, 1, 215-239.

Früchterman T. M. and Reingold E. M. (1991). Graph drawing by force-directed placement, *Software: Practice and Experience*, 21, 1129-1164.

Gao J. et al. (2017). The Intellectual Structure of Digital Humanities: An Author Co-Citation Analysis, *Digital Humanities 2017*, Montreal.

Gao J. et al. (2018). Visualising The Digital Humanities Community: A Comparison Study Between Citation Network And Social Network, *Digital Humanities 2018*, Mexico City.

Grandjean M. (2015). Network Visualization: Mapping Shakespeare's Tragedies, http://www.martingrandjean.ch/network-visualization-shakespeare.

Grandjean M. (2016). Archives Distant Reading: Mapping the Activity of the League of Nations' Intellectual Cooperation, *Digital Humanities 2016*, Krakow.

Grandjean M. (2017). Multimode and Multilevel: Vertical Dimension in Historical and Literary Networks, *Digital Humanities 2017*, Montreal.

Grandjean M. et al. (2017). Complex Network Visualisation for the History of Interdisciplinarity: Mapping Research Funding in Switzerland, *Digital Humanities 2017*, Montreal.

Grandjean M. (2018). *Les réseaux de la coopération intellectuelle. La Société des Nations comme actrice des relations scientifiques et culturelles dans l'entre-deux-guerres*, Lausanne.

Granovetter M. S. (1973). The Strength of Weak Ties, *American Journal of Sociology*, 78, 1360-1380.

Jänicke S. and Focht J. (2017). Untangling the Social Network of Musicians, *Digital Humanities 2017*, Montreal.

Jannidis F. et al. (2016). Comparison of Methods for the Identification of Main Characters in German Novels, *Digital Humanities 2016*, Krakow.

Kaufman M. et al. (2017). Visualizing Futures of Networks in Digital Humanities Research, *Digital Humanities 2017*, Montreal.

Marraccini M. (2017). The Victoria Press Circle, *Digital Humanities 2017*, Montreal.

Maryl M. and Eder M. (2017). Topic Patterns in an Academic Literary Journal: The Case Of Teksty Drugie, *Digital Humanities 2017*, Montreal.

Moreno J. L. (1934). *Who Shall Survive? A New Approach to the Problem of Human Interrelations*, Nervous and Mental Disease Publishing.

Moretti G. et al. (2016). Building Large Persons' Networks to Explore Digital Corpora, *Digital Humanities 2016*, Krakow.

Noack, A. (2009). Modularity clustering is force-directed layout. *Physical Review E*, 79(2), 026102.

Pino-Díaz J. and Fiormonte D. (2018). La Geopólitica De Las Humanidades Digitales: Un Caso De Estudio De DH2017 Montreal, *Digital Humanities 2018,* Mexico City.

Piper A. et al. (2017). Studying Literary Characters and Character Networks, *Digital Humanities 2017*, Montreal.

Rosa Pérez J. et al. (2018). Histonets, Turning Historical Maps Into Digital Networks, *Digital Humanities 2018,* Mexico City.

Rybicki J. et al. (2018). Polysystem Theory And Macroanalysis. A Case Study Of Sienkiewicz In Italian, *Digital Humanities 2018*, Mexico City.

Sozinova O. (2016). Complex Networks-Based Approach to Russian Rhyme History Description: Linguostatistics and Database, *Digital Humanities 2016*, Krakow.

Tang M.-C. and Chen K.-H. (2017). A cross-language comparison of co-word networks in Digital Library and Museum of Buddhist Studies, *Digital Humanities 2017*, Montreal.

Tayler F. and Neugebauer T. (2018). Complex Networks Of Desire: Mapping Community In Visual Arts Magazines Fireweed, Fuse, And Border/Lines, *Digital Humanities 2018*, Mexico City.

Torres-Yepez L. and Zreik K. (2018). Estudio Exploratorio Sobre Los Territorios De La Biopirateria De Las Medicinas Tradicionales En Internet : El Caso De America Latina, *Digital Humanities 2018*, Mexico City.

Trilcke P. et al. (2016). Theatre Plays as 'Small Worlds'? Network Data on the History and Typology of German Drama, 1730–1930, *Digital Humanities 2016*, Krakow.

Verhoeven D. et al. (2018). Solving the Problem of the "Gender Offenders": Using Criminal Network Analysis to Optimize Openness in Male Dominated Collaborative Networks, *Digital Humanities 2018*, Mexico City.

Weingart S. B. (2011). Demystifying Networks, Parts I & II, *Journal of Digital Humanities*, 1, 1.

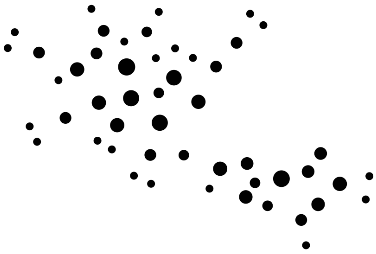
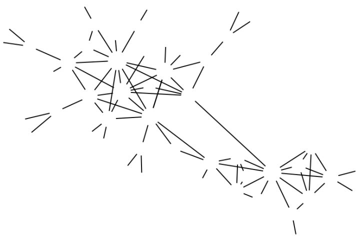
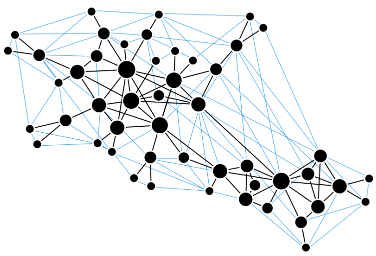
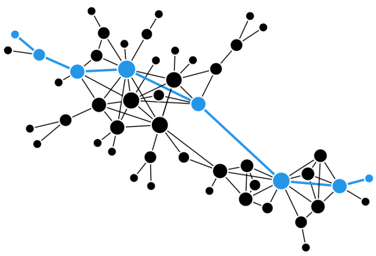
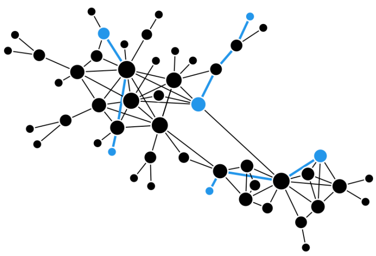
Wieneke, L. et al. (2016). Introducing HistoGraph 2: Exploration of Cultural Heritage Documents Based on Co-Occurrence Graphs, *Digital Humanities 2016*, Krakow.

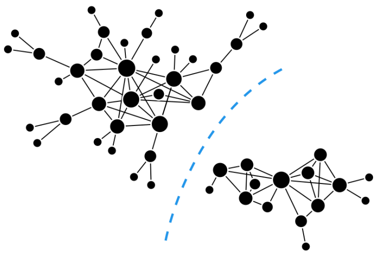
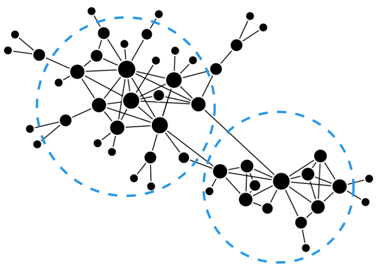
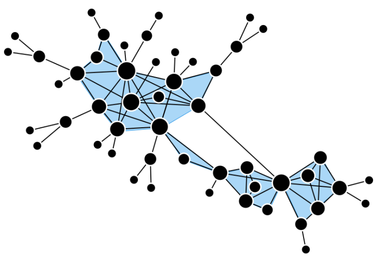
Wright C. (2016). The Formation of Australia's Economic History Community, 1950–1970: A Multidimensional Network Analysis, *Digital Humanities 2016*, Krakow.

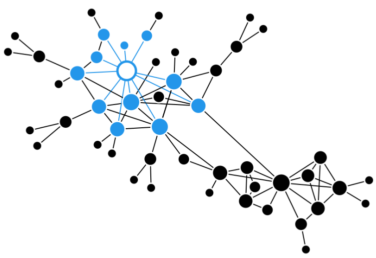
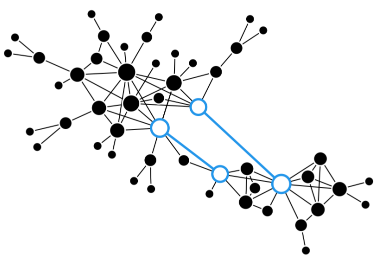
Xanthos A. et al. (2016). Visualising the Dynamics of Character Networks, *Digital Humanities 2016*, Krakow.

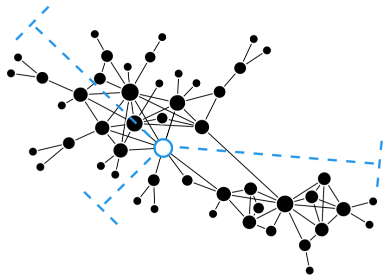
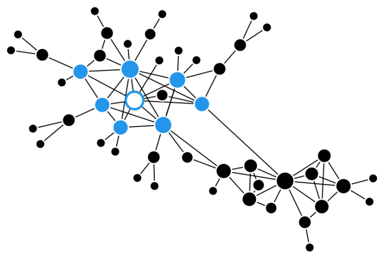
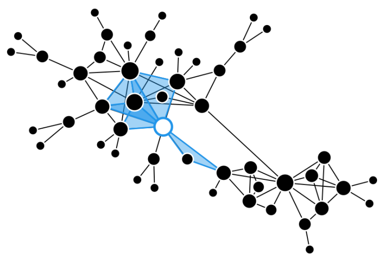
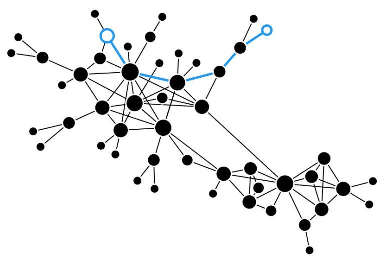
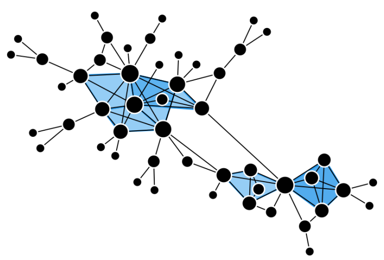
Network visual and topological patterns

This table of correspondence between network analysis concepts and interpretations or "translations" is a work in progress. The authors propose this document to open a discussion on the most relevant translation scenarios and examples/references from the different disciplines applying these methods.

	NOTION		VISUAL ANALYSIS		COMPUTATIONAL ANALYSIS		INTERPRETATIVE POTENTIAL
Global		**Graph size (nodes)** SIMPLE DEFINITION How many nodes are there in the graph?	VISUAL PATTERN There are few or many nodes. HERMENEUTIC CRITERIA Counting the nodes.	ISSUES Though it is easy to have an estimation of the total number of nodes, visualization decisions (for example, setting node sizes on a large scale) can make nodes with few connections difficult to see.	TOPOLOGICAL PATTERN Nodes count. COMPUTATIONAL CRITERIA Counting the nodes.	ISSUES None. Note that in graph theory the count of nodes is referred to as "graph order" while the "graph size" refers to the count of edges.	USEFULNESS Very basic information but useful when comparing networks. "TRANSLATION" The intuition of the **size** of a network is appropriate.
Global		**Graph size (edges)** SIMPLE DEFINITION How many edges are there in the graph?	VISUAL PATTERN There are few or many edges. HERMENEUTIC CRITERIA Counting the edges.	ISSUES The total number is hard to estimate as soon as the graph is not a simple diagram anymore. The distribution of edges and their weight has an influence on the visual estimation. The difficulty to count edges visually is a known issue, and probably impossible to overcome.	TOPOLOGICAL PATTERN Edges count. COMPUTATIONAL CRITERIA Counting the edges.	ISSUES None. Note that in graph theory the count of edges is referred to as "graph size" while the count of nodes is "graph order".	USEFULNESS Very basic information but useful when comparing networks. "TRANSLATION" Sometimes translated as **size** in natural language, but the number of edges is usually expressed in comparison to the number of nodes to indicate density or complexity, not for itself.
Global		**Density** SIMPLE DEFINITION How connected are the nodes overall?	VISUAL PATTERN The graph is more or less compact. If only certain parts are more compact, see "clusters" below. HERMENEUTIC CRITERIA Accumulation of edges, cluttered groups of nodes, "hairball". Easier to estimate in situation of comparison.	ISSUES The less edges there are, the easier to estimate. High densities are difficult to distinguish because the overall appearance of a graph with 60% edges will looks close to a graph with 100%. The visual aspect also depends on the layout algorithm used: some are more efficient at representing clusters.	TOPOLOGICAL PATTERN Network density. COMPUTATIONAL CRITERIA Divide the actual number of edges by the number of all potential edges.	ISSUES The formula of density slightly changes depending on the type of networks (directed or not, self-loops allowed or not...).	USEFULNESS A very important notion that allows to compare networks of different sizes if they are produced in the same way or on the basis of comparable data sets. "TRANSLATION" **Density**, **complexity**, **completeness**.
Global		**Diameter** SIMPLE DEFINITION How far are the most distant nodes?	VISUAL PATTERN The longest shortest path in the graph, or the two most visually distant nodes. HERMENEUTIC CRITERIA Following the series of edges from one node to another, trying to find the longest one. If impossible, the most visually distant nodes is acceptable.	ISSUES Generally hard or impossible to see except on small networks, but quite easy to estimate by following a few paths that goes from a side to another, or just looking at most distant nodes in the same connected component (visual distance approximately correlates with geodesic distance).	TOPOLOGICAL PATTERN Diameter. COMPUTATIONAL CRITERIA Maximal geodesic distance of all the pairs of nodes.	ISSUES Only relevant in a connected graph.	USEFULNESS Can be used to describe how the density is distributed: complex networks are often characterized by a small diameter while high diameter is frequent in geographical networks. "TRANSLATION" **Size**, **breadth**, **width.**
Global		**Average path length** SIMPLE DEFINITION On average, how close are nodes to each other?	VISUAL PATTERN The average distance between a couple of nodes. HERMENEUTIC CRITERIA Following the edges between every couples of nodes.	ISSUES Impossible to calculate visually since it is an average covering a very large number of values (already difficult to calculate). Loosely relates to density, which is easier to estimate.	TOPOLOGICAL PATTERN Average path length. COMPUTATIONAL CRITERIA Average number of steps along the shortest paths for all possible pairs of nodes	ISSUES Since it is an average, this value does not allow conclusions to be drawn at the individual level if the graph is strongly clustered. The average path length is more complicated for a directed graph than for an undirected graph.	USEFULNESS Could serve as a complement to diameter because the latter can be influenced by a few nodes that are very far from the main component of the graph. Can replace the diameter in case of unconnected graphs. "TRANSLATION" Can be used to describe the **size**, **breadth** or **width** of the network. But it can also be translated into an indicator of a **small world** situation.

	NOTION		VISUAL ANALYSIS		COMPUTATIONAL ANALYSIS		INTERPRETATIVE POTENTIAL
Global		**Connectedness** SIMPLE DEFINITION Is the graph a connected system where there is a path between every nodes?	VISUAL PATTERN There must be only one component, not several groups of nodes disconnected from each other. HERMENEUTIC CRITERIA Looking at empty areas (structural holes). Groups count as disconnected only if there are no edges between them.	ISSUES Depending on the layout, it is possible that "islands" hide in dense groups of nodes, but with a properly set force-directed layout, the risk is marginal. Possibly the easiest property to observe visually.	TOPOLOGICAL PATTERN In a connected graph, the number of connected components must be one. COMPUTATIONAL CRITERIA There must be a path between each pair of nodes.	ISSUES The notion of connectedness is more complicated for a directed graph than for an undirected graph.	USEFULNESS The absence of edges between components is more remarkable if they contain many nodes. In many applied cases, connected graphs are artificially created by removing solitary nodes (frequent in messy extracted data). "TRANSLATION" The network is a **continent**, or, on the contrary, an **archipelago**.
Global		**Clusters/ Communities** SIMPLE DEFINITION What are the groups where nodes are more connected to each other?	VISUAL PATTERN Clusters (uneven distribution of nodes) HERMENEUTIC CRITERIA Looking for groups of nodes, as visually dense and separated as possible.	ISSUES Force-directed placement algorithms are known to represent clustering very well, if properly set.	TOPOLOGICAL PATTERN Max modularity (modularity of the partition with the highest modularity). COMPUTATIONAL CRITERIA Running a modularity clustering detection algorithm and looking at the obtained modularity.	ISSUES Maximal modularity is too hard to compute, so we rely on an estimation. Modularity is a measure of a graph partitioning, so it is necessary to partition the graph first. Different algorithms exist (Louvain, Leiden...). Other metrics than modularity also exist, though they are less widely used.	USEFULNESS Useful for exploration. It is tempting to take the result of a cluster calculation as a given. In some cases, it is interesting to compare these clusters with previously known groups (categories that do not depend on the structure obtained). "TRANSLATION" The term **cluster** has become part of the common language, but we also like to talk about **groups**, **communities** or **hubs**. This notion of community is very directly related to the way in which the social sciences and humanities use the metaphor of the "network".
Global		**Global or average clustering coefficient** SIMPLE DEFINITION General indication of the graph's tendency to be organized into clusters	VISUAL PATTERN Clusters. HERMENEUTIC CRITERIA Looking for triangles (groups of three nodes with three connexions between them).	ISSUES Triangles are easy to count visually in a small network, but the ratio between this result and the total number of potential triangles is impossible to calculate directly. Very difficult to count in dense graphs. Graphs with clusters and/or visually dense tend to have a higher clustering coefficient.	TOPOLOGICAL PATTERN Number of closed triplets. COMPUTATIONAL CRITERIA Dividing the number of closed triplets by the number of triplets in the graph.	ISSUES The global clustering coefficient is obtained by dividing the closed triplets by the number of all possible groups of three nodes. The average clustering coefficient is quite different but serve a relatively close purpose: it is the average of the local clustering coefficient of all the nodes.	USEFULNESS A global measure that complements density well and, like the latter, is useful for comparing similar networks with each other. "TRANSLATION" Gives an idea of the **entanglement** / **intrication** and the presence of a more localized **density**.

Local		**Connectivity (degree)** SIMPLE DEFINITION How well connected is a node / how many links it has / how many neighbors	VISUAL PATTERN There are many links to the node. HERMENEUTIC CRITERIA Counting the edges converging to that node.	ISSUES In a dense image, it is not always obvious which edges converge to the node or just happen to pass through it visually. Sometimes, counting is also impractical.	TOPOLOGICAL PATTERN Degree. COMPUTATIONAL CRITERIA Degree.	ISSUES In a directed network, we also distinguish indegree (inbound link) and outdegree (outbound links). In that case, the degree is the sum of those and hence the number of links, not the number of neighbors.	USEFULNESS The simplest form of centrality. In most cases, the degree shows information that is already known as part of the basic data and not dependent on the structure. This is why we often focus on the degree distribution. "TRANSLATION" The basic intuition of the number of **neighbors**. In directed networks, interpretation varies greatly between in- and out- degrees: indegree is often the primary way to look at a hierarchy of nodes, because being "cited" is often a good proxy for **authority/notoriety**.
Local		**Betweenness** SIMPLE DEFINITION Being a bridge, connecting otherwise separated groups of nodes. Removing that node would break many shortest paths.	VISUAL PATTERN A bridge between clusters. HERMENEUTIC CRITERIA Looking for an edge appearing through a (mostly) empty area between large groups of nodes.	ISSUES Many bridges look as expected, they connect over empty spaces. But sometimes, bridges are hidden in the complicated structure of the image. It is generally easier to see the bridging edges than the bridging nodes (however, most of the studies using betweenness centrality focus on bridging nodes).	TOPOLOGICAL PATTERN Betweenness centrality. COMPUTATIONAL CRITERIA The score of betweenness centrality represents the number of shortest paths through a given node or edge.	ISSUES Note that the undirected version of the algorithm is often used even for directed graphs. It is the most used metric for detecting bridges, but it does not exactly meet the intuition.	USEFULNESS The definition of bridge implemented by betweenness centrality meets both intuition of a bridge and of a center. Indeed both a bridge and the center of a star are things that, if removed, disconnect parts of the network. In that sense betweenness is also a "centrality". "TRANSLATION" The notion of **bridge** (but also **link**, **gateway**, **broker** or **key passage**) is very often used when applying network analysis to social or circulation issues. In some cases, it can represent a form of **social capital** because it describes a structural position of **power** (or **vulnerability**).

	NOTION		VISUAL ANALYSIS		COMPUTATIONAL ANALYSIS		INTERPRETATIVE POTENTIAL
Local		**Closeness** SIMPLE DEFINITION Being in the middle of the network.	VISUAL PATTERN The geographical center of the graph. HERMENEUTIC CRITERIA Finding the barycenter (the center of the "land masses") of the graph.	ISSUES The visual estimation of centrality is considered acceptable, but it remains an evaluation. It is harder to find in very sparse graphs.	TOPOLOGICAL PATTERN Closeness centrality. COMPUTATIONAL CRITERIA The score of closeness centrality is the average length of the geodesic distances to all the other nodes.	ISSUES The undirected version is often used even for directed graphs.	USEFULNESS There is no single implementation of centrality, but closeness centrality is the most aligned with the notion of a middle, a point that is in proximity to all the others. It looks at the structural distance to other nodes, and can be interpreted as such. "TRANSLATION" Excellent to describe the **centre** or the **middle** of a network, especially when the latter is described in topographical terms. Low values of this metric are very appropriate for the use of concepts which are the opposite of the center: the **periphery**, the **margins**, etc.
Local		**Prestige (eigenvector)** SIMPLE DEFINITION Being connected to well connected nodes without necessarily having a large number of neighbors itself.	VISUAL PATTERN Proximity to well connected nodes (often inside a cluster). HERMENEUTIC CRITERIA None, except if the size of the nodes is visually proportional to the degree centrality, which helps to identify nodes in the hubs' surroundings.	ISSUES This centrality is hard to see, though it correlates with other forms of centrality that point to well connected nodes at the center of the graph.	TOPOLOGICAL PATTERN Eigenvector centrality or Page Rank. COMPUTATIONAL CRITERIA This is a score computed recursively. It flows from each node to its neighbors (following the direction of edges in a directed graph).	ISSUES None.	USEFULNESS This form of centrality is notably adapted to directed networks, and can be related to the functioning of a search engine (the Page Rank principle was used in the first Google search engine) or a system where information flows. "TRANSLATION" The iterative nature of this notion makes it difficult to translate (and difficult to use in some contexts). It is confused with the notions of **prestige**, **authority**, **influence** and, sometimes, **power** and **elites**. This measure distinguishes nodes that are **"well"** connected (and not just "a lot"). Relates to the notion of **assortativity**.
Local		**Local clustering coefficient** SIMPLE DEFINITION Are the neighbours of a node also connected together?	VISUAL PATTERN Nodes are inside a cluster. HERMENEUTIC CRITERIA Looking for nodes that have many edges to their cluster (where the other nodes are also connected together). Bridges have a low clustering coefficient.	ISSUES Clustering coefficient is generally hard to see and visual interpretation is considered unreliable. Exceptions are small networks, nodes that have only a few neighbors that we see well, and nodes that are only connected to a very dense cluster.	TOPOLOGICAL PATTERN Clustering coefficient (local). COMPUTATIONAL CRITERIA Calculate density of the subgraph of neighbors (how close from complete is the graph formed by the node and its neighbors).	ISSUES None.	USEFULNESS Meets a notion of redundancy in the local connections, comparable to centrality but at a very local scale. Tells if a node is in a clustered environment. Complex networks are often characterized by a high average clustering coefficient. "TRANSLATION" This local measure makes it possible to analyse relationships at the collective level: it can be translated as an indicator of **participation in a group** (or, on the opposite, **loneliness**, **solitude**). It opposes the notion of a bridge.
Local		**Shortest path** SIMPLE DEFINITION Two nodes are connected by a path	VISUAL PATTERN Presence of a path between the two nodes whose relationship is to be analyzed. HERMENEUTIC CRITERIA Following the series of edges from one node to another to find the shortest.	ISSUES Requires that we can follow the links in practice, which is possible only for small (undirected) networks and depends on the graphic settings. Finding a path can be difficult, and ensuring that this path is the shortest can be too difficult. However the visual distance is a loose approximation of the shortest path length.	TOPOLOGICAL PATTERN Shortest path(s). COMPUTATIONAL CRITERIA Geodesic distance, algorithms for shortest path detection.	ISSUES Can be computationally costly on big networks. Note that multiple shortest paths can exist.	USEFULNESS Very adapted to the use of the graph as a research interface to test the relation of couples of nodes. Very close to the qualitative approach of the humanities, which are often focusing on a few individuals in the network. "TRANSLATION" Corresponds to the intuitive notion of **distance** in the graph structure. Note that this translation does not take into account the fact that nodes are not always aware of the steps between them and that the perceived distance is not always the shortest path.
Local		**Cliques** SIMPLE DEFINITION Groups of nodes where all possible edges exist between them.	VISUAL PATTERN Very dense clusters. HERMENEUTIC CRITERIA Looking for groups of nodes where each of them is connected to all the others.	ISSUES Possible for small or sparse networks, especially if the focus is on cliques that are 4+ in size. Visually impossible for complex networks where cliques are very frequent.	TOPOLOGICAL PATTERN Clique (group of nodes with density of 1). COMPUTATIONAL CRITERIA A group of nodes has a density of 1. Clique detection algorithms exist.	ISSUES The number of cliques is often very high. Quasi-cliques are almost as important as cliques. But multiple algorithms deal with this structure (eg. clique percolation).	USEFULNESS The number of cliques, their size and distribution are metrics that are complementary to the clustering coefficients (local and global). They can be used as a more strict community detection algoritm. "TRANSLATION" The term **clique** itself refers to social groups of individuals who knows each other. They can be translated as **communities**, **neighborhoods**, closed **societies**.