



'Double blind' graph data analysis: a pedagogical experiment to discuss the intersubjectivity of network interpretation

Mathieu Jacomy, Martin Grandjean

► To cite this version:

Mathieu Jacomy, Martin Grandjean. 'Double blind' graph data analysis: a pedagogical experiment to discuss the intersubjectivity of network interpretation. Digital Humanities, 2022, Tokyo, Japan. 10.5281/zenodo.6844790 . halshs-03753812

HAL Id: halshs-03753812

<https://shs.hal.science/halshs-03753812>


Submitted on 18 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

'Double blind' graph data analysis: a pedagogical experiment to discuss the intersubjectivity of network interpretation

Mathieu Jacomy, TANT Lab, Aalborg University, Denmark / [@jacomy](#)
Martin Grandjean, History Department, University of Lausanne, Switzerland / [@grandjeanmartin](#)

 Conference paper

Jacomy M. & Grandjean M. (2022). "[Double blind graph data analysis: a pedagogical experiment to discuss the intersubjectivity of network interpretation](#)". *Digital Humanities 2022*, Tokyo, Japan.

DOI [10.5281/zenodo.6844790](#)

Introduction

On this page we propose an experimental process of 'double blind' comparison which consists of giving the same dataset to two network analysts and to confront their methodological choices, results and interpretations. This 'double blind' approach (Jacomy and Grandjean, 2021) makes it possible to make the intersubjectivity of the process visible and provide empirical material to discuss how the eyes of specialists from different disciplines influence the interpretation of the data. We believe that producing and making explicit such interpretive pathways can help researchers, designers and students understand the diversity of possible approaches.

The subjectivity of network interpretation in the humanities

If network analysis is a method now very widely diffused in the field of digital humanities (Ahnert et al., 2020), often because the metaphor and lexical field of 'network' are very effective in describing the objects of study of the humanities, there is no definitive interpretation procedure. Indeed, describing a network entails the recontextualization of graph-theoretical elements (e.g., the "betweenness centrality" metric) into the language of the discipline (e.g., "knowledge brokers" or simply "bridges") through a methodological "translation" (Grandjean and Jacomy, 2019). This hermeneutic process is subjective by nature, especially in the humanities where visual network analysis is often privileged, which may seem paradoxical considering that the method is supposed to objectify historical or literary sources (this is more generally true for all digital methods). Yet as Popper argued, descriptive statements necessarily draw their validity from "intersubjective agreement" (Freeman, 1973). How much agreement does network analysis offer?

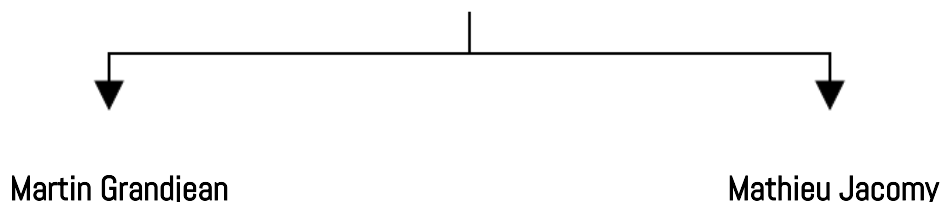
Experimental process

To ensure comparability, the phase of individual experimentation is framed by strict constraints in terms of time and tools. A network dataset is prepared upstream and duplicated: the two participants get a time budget of 2 hours to explore the data with Gephi (Bastian et al. 2009) and produce an annotated network visualization of their choice, possibly including other tables and visualizations, as well as a list of hypotheses on the structure of the graph. This phase is filmed and screencasted. Then the two people discover each other's work and compare their interpretations.

The result of these parallel interpretations is presented as two paths (below) and can be compared by means of an online document integrating the data set, interactive and/or composite visualizations, descriptive text and video sequences making it possible to follow the reasoning. In the end, the data set itself is only a pretext for an educational experiment, the aim of which is as much to confront us with the subjectivity of the interpretation as to show the interest of such an exercise of replication.


Data

The data set is a network of Wikipedia pages that link to each other. It has been harvested from the "[European cuisine](#)" category on the English Wikipedia. All its pages have been extracted as well as the pages listed in its subcategories using notebooks developed by Mathieu and available on his "[Mapping Controversies](#)" website. All the links between the pages of this list have been retrieved and all the pages of the main giant component that are linked at least twice have been kept. The network contains 5.800 nodes and 300.000 edges.



First visualization

Network Analysis 1/4 - Introduction and Fir...




First visualization

Euro cuisine network 1/4 - First visualization



Network exploration

Network Analysis 2/4 - Exploration




Network exploration

Euro cuisine network 2/4 - Network explora...




Final visualization

Network Analysis 3/4 - Final Visualization




Final visualization

Euro cuisine network 3/4 - Final visualization



Summary

Network Analysis 4/4 - Summary



Summary

Euro cuisine network 4/4 - Summary



