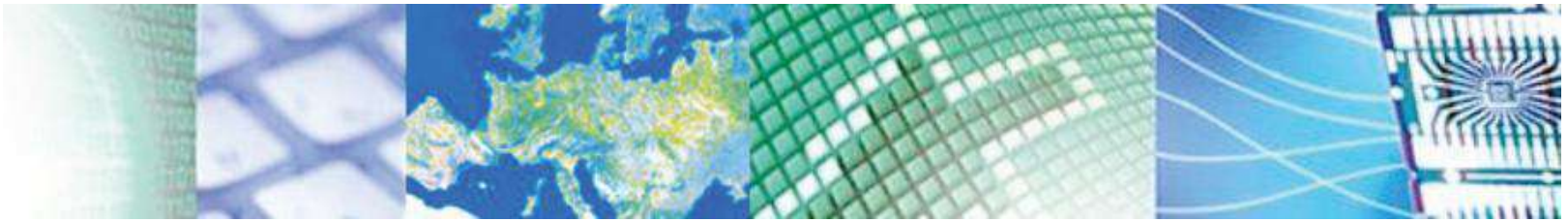


CHIST-ERA Projects Seminar 2019
Visual Analytics for Decision-Making under Uncertainty

Roberto Therón, PROVIDE DH

Dimitri Van de Ville, IVAN

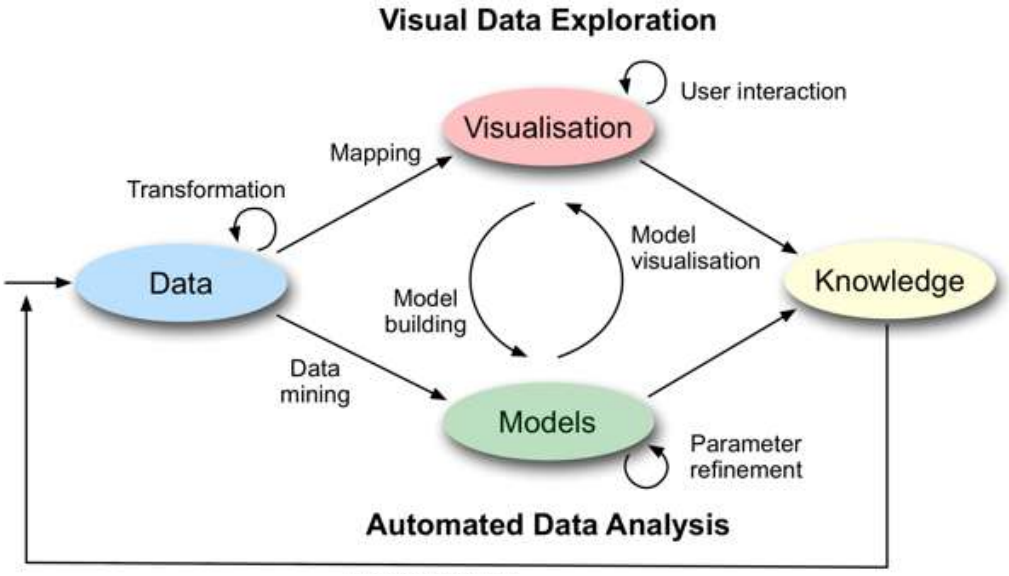
Bucharest, April 4, 2019

Programme co-funded by the EUROPEAN UNION

Visual Analytics
Decision-Making under Uncertainty

Big data analytics

- How to understand reliability, scope, and accuracy, and to communicate these aspects to users in an intuitive manner?
- How to support trusted and efficient decision making under uncertainty?
  - ✓ uncertainty analysis of ensemble data, sensitivity analysis of input-output models, etc.



Visual Analytics – before addition of uncertainty

This is a multidisciplinary challenge that engages real-world applications

VADMU

Extending Visual Analytics to Encompass Uncertainty

When data has uncertainty and decisions should be made, ICT is not enough. Therefore the VADMU projects focus on the following:

- **Exploration, interaction, visualization, to understand the results of analyses and nature of uncertainty**
- **Low latency for algorithms to fit the human capabilities**
  - Fast algorithms, or approximate results, or progressive results
- **New lines or methods for data analysis at scale under uncertainty**

Introduction: Projects of the topic

Visual Analytics for Decision Making under Uncertainty

- **PROVIDEDH: PROgressive VIsual DEcision-making for Digital Humanities**

  Provide visual interactive tools that convey the degree of uncertainty of the datasets and computational models used behind, designed to progressively adapt the visualizations to incorporate the new, more complete or more accurate data.

- **IVAN: Interactive and Visual Analysis of Networks**

  Create a visual analysis system for the exploration of dynamic or time-dependent networks, in particular using modern clustering methods

PROVIDEDH: Project Focus

- **Two key scientific questions**

  Is it possible to **assess and track** the degree of uncertainty of DH datasets, and how they are affected when different computational models are applied to them?

  Is it possible to convey this evolution of uncertainty with **interactive multimodal interfaces that progressively adapt** to the moment of decision-making?

- **Key challenges**

  Categorise all **sources of uncertainty** that can affect DH

  Develop a set of **metrics for the degree of uncertainty** and propose a (software) framework that makes use it

  Clarify what **DH tasks** need decision-making and build a pipeline oriented towards the handling of uncertainty

PROVIDEDH: PROgressive VIsual DEcision-making for Digital Humanities

Partners

University of Salamanca (SPAIN)

- **Project Coordinator: Roberto Therón**

Austrian Academy of Sciences (AUSTRIA)

- **PI: Eveline Wandl-Vogt**

Trinity College Dublin (IRELAND)

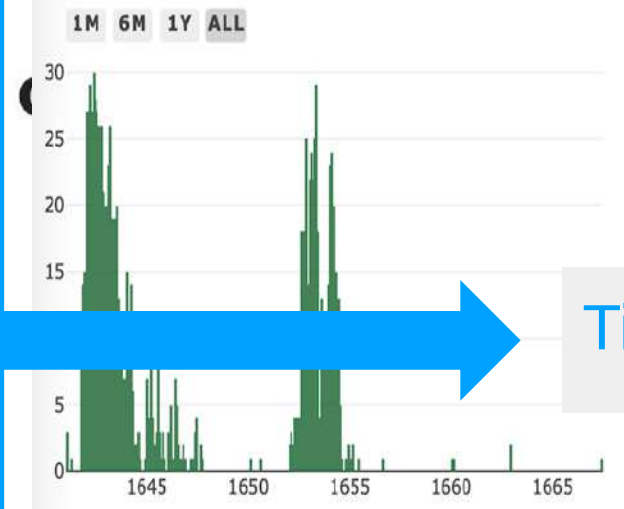
- **PI: Jennifer Edmond**

Poznan Supercomputing and Networking Center (POLAND)

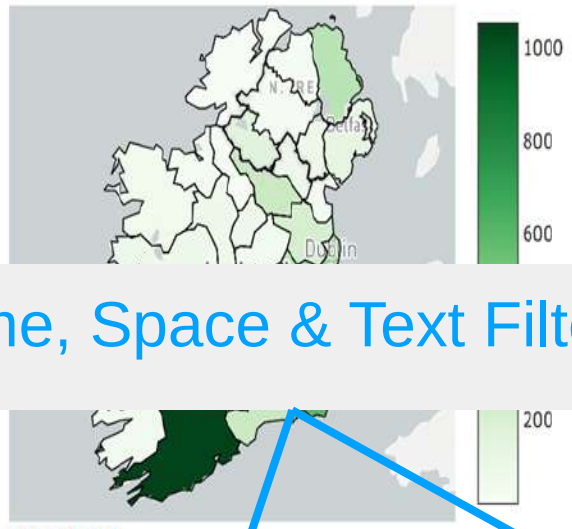
- **PI: Cezary Mazurek**

Project start: February 2018

The 1641 Depositions



6013 depositions in 32 counties (All counties) in period 1641-01-05 to 1667-06-19



821 828 845

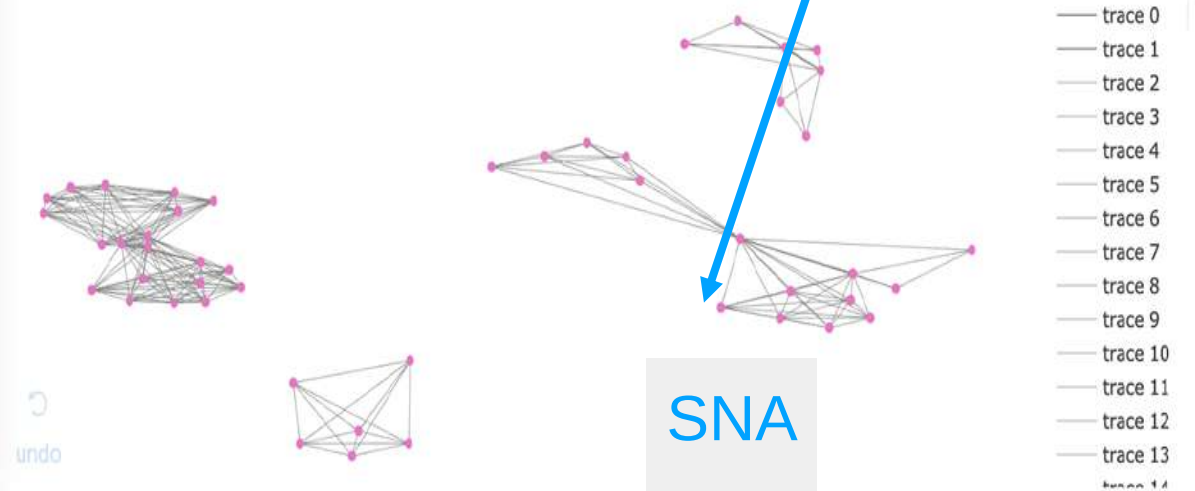
forename	surname	appearances
filter data...	filter data...	
john	Unknown	108
richard	butler	96
the	king	91
luke	toole	81
richard	condon	68
hugh	Unknown	64
brian	Unknown	62
william	stafford	61

PREVIOUS NEXT

Time, Space & Text Filters

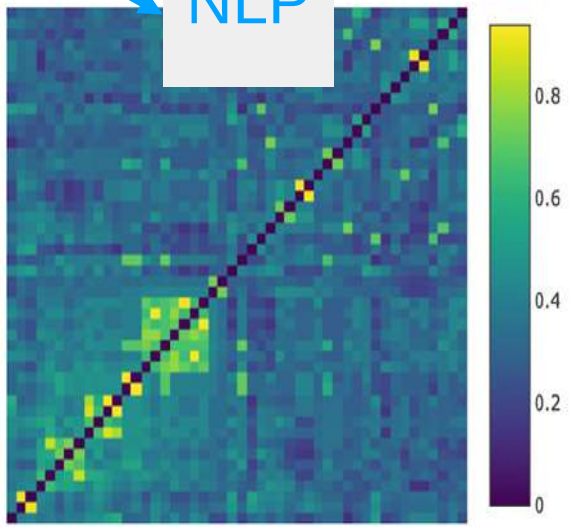
35364 distinct names referenced

GRAPH HEATMAP



SNA

NLP



PROVIDEDH: Significant Results

- Literature review within DH and interviews with humanities scholars
- Taxonomy of sources of uncertainty in DH
- A collaborative platform addressing the initial identified user stories
- A suitable case study and dataset (1641 Depositions) prepared and uploaded to the collaborative platform
- Several visualization techniques tested through micro-prototyping

PROVIDEDH: Dissemination to Date

Collaborative Platform (https://providedh.ehum.psnc.pl)

Journal Articles (OA)

- Santamaría, R., Therón, R., Durán, L., García, A., González, S., Sánchez, M., & Antequera, F. (2018). ***Genome-wide search of nucleosome patterns using visual analytics***. Bioinformatics†. DOI: https://doi.org/10.1093/bioinformatics/bty971
- Aouabed H, Santamaria R, Elloumi M (2018) ***Biclustering Impact in Biomedical Sciences via Literature Mining***. Int J Biomed Data Min 7: 134. DOI: 10.4172/2090-4924.1000134

Conference Papers

- Jennifer Edmond. (2018). ***Managing Uncertainty in the Humanities: Digital and Analogue Approaches***
- Roberto Theron, Antonio Losada, Alejandro Benito, & Rodrigo Santamaría. (2018). ***Toward supporting decision-making under uncertainty in digital humanities with progressive visualization***
- Roberto Theron & Eveline Wandl-Vogt. (2018). ***Uncertainty in Digital Humanities track summary***

International journal: Special Issue

- **Informatics** Open Access Journal on the topic of "**Uncertainty in Digital Humanities**"

PROVIDEDH: Dissemination to Date

Workshops to Date

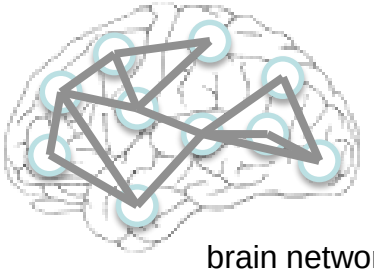
- TEEM 2018 (Salamanca, October 2018). Fifth edition of the International Conference Technological Ecosystems for Enhancing Multiculturality:
  - **Result**: a dedicated track to the research topics of PROVIDEDH: “Uncertainty in Digital Humanities ” was organized by the consortium. This session attracted international **researchers from Austria, Germany, Ireland, Spain, and Poland. 9 papers (with a total of 22 authors)** where published by ACM. The session triggered fruitful discussions and served as a means to create a community of researchers that can inform the PROVIDEDH activities and benefit from its results.
- ThePort’2018 (Geneva, October 2018):
  - **Result**: Design for an application, based on which the Austrian team creates a case study. The case study is presented at the LENS -conference 2019 (Designing sustainability for all; World Distributed Conference).
- EADH-2018 (Galway, December 2018): Workshop: Developing biographical data projects using open innovation methods and practices
  - **Result**: Introduction of OI and uncertainties in DH; User stories; experimenting with new interaction designs in DH.
- Uncertainties in Humanities Research Datasets
  - **Result**: a one day workshop co-hosted by the Humanities Research Institute at Maynooth University, the Trinity College Dublin Centre for Digital Humanities and the consortium members of the PROVIDEDH project, Maynooth, Ireland, 6 March 2019.

Future Workshops and Conferences:

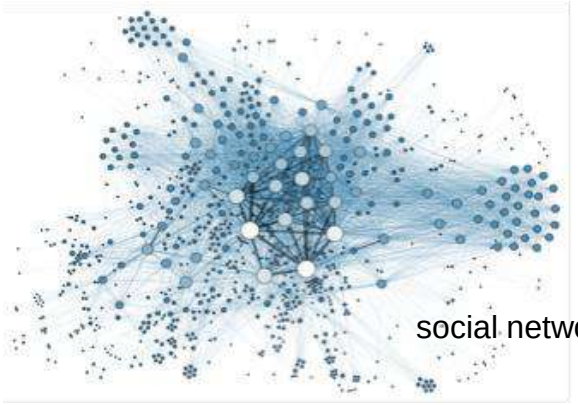
- DH2019 (Utrecht, July 2019) **Workshop on Complexity and Uncertainty in Digital Humanities Projects: A Co-Design Approach Around Data Visualisation**
  - https://providedh.eu/workshop-on-complexity-and-uncertainty-in-digital-humanities-projects-a-co-design-approach-around-data-visualisation/
- TEEM 2019 (León, October 2018). Fifth edition of the International Conference Technological Ecosystems for Enhancing Multiculturality: **Track: Uncertainties Digital Humanities**
  - https://2019.teemconference.eu/uncertainty-digital-humanities/

IVAN / Background

- Network models are increasingly important in many scientific domains



brain networks



social networks

high variety
dynamic networks
rich annotations
…

- State-of-the-art algorithms for network clustering/embedding have different parameters that are hard to tune
- **IVAN: make novel algorithms accessible to a broad range of users and researchers to enable reliable and informed decisions based on the network analysis under uncertainty**
  - *user-guided analysis, interactive visualization, dynamic information*

IVAN: Interactive Visual Analysis of Networks

chist-era

universität wien

Torsten Möller, Stefan Starflinger

Long history in exploring the issues of modeling under uncertainty and decision making

Inria

Jean-Daniel Fekete, Paola Valdivia

Well-established reputation in HCI, network visualization, and evaluation methods for visualization

EPFL

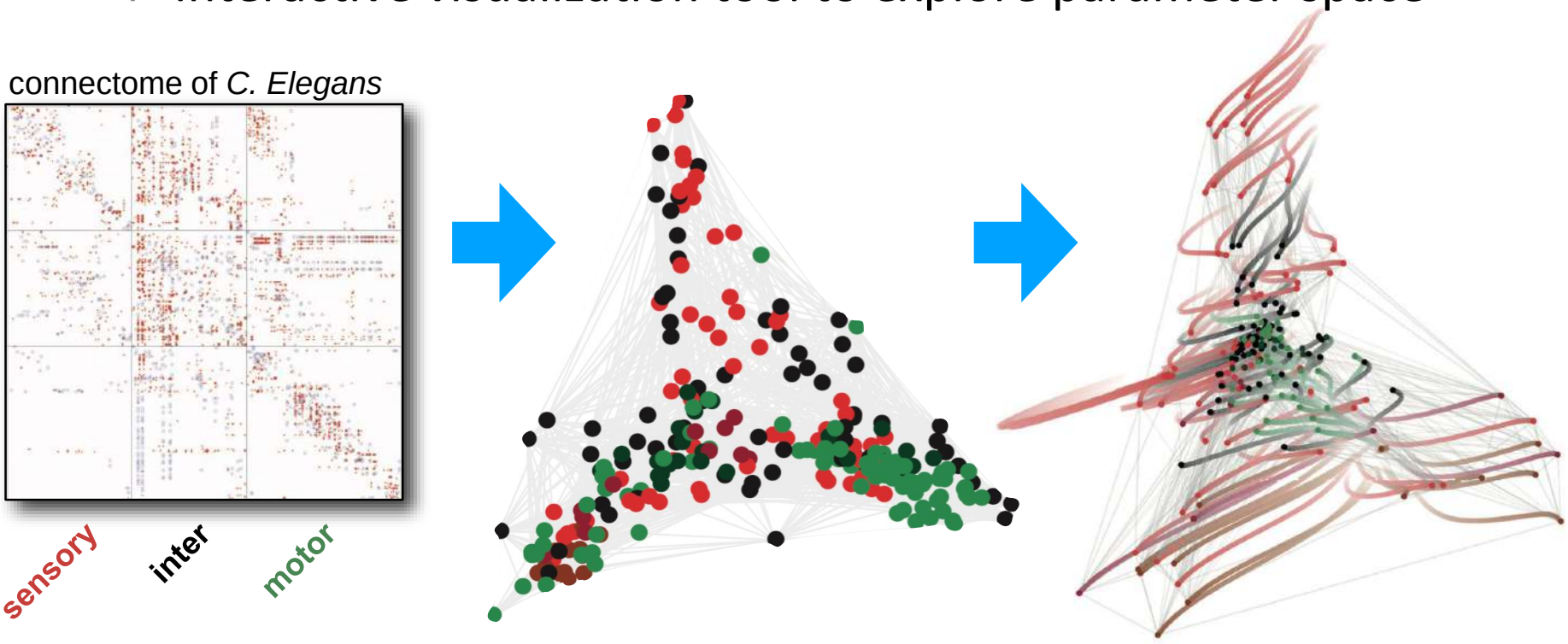
Dimitri Van De Ville, Raphaël Liégeois

Graph signal processing expertise including spectral/wavelet approaches with application to brain networks

Project start: April 2018

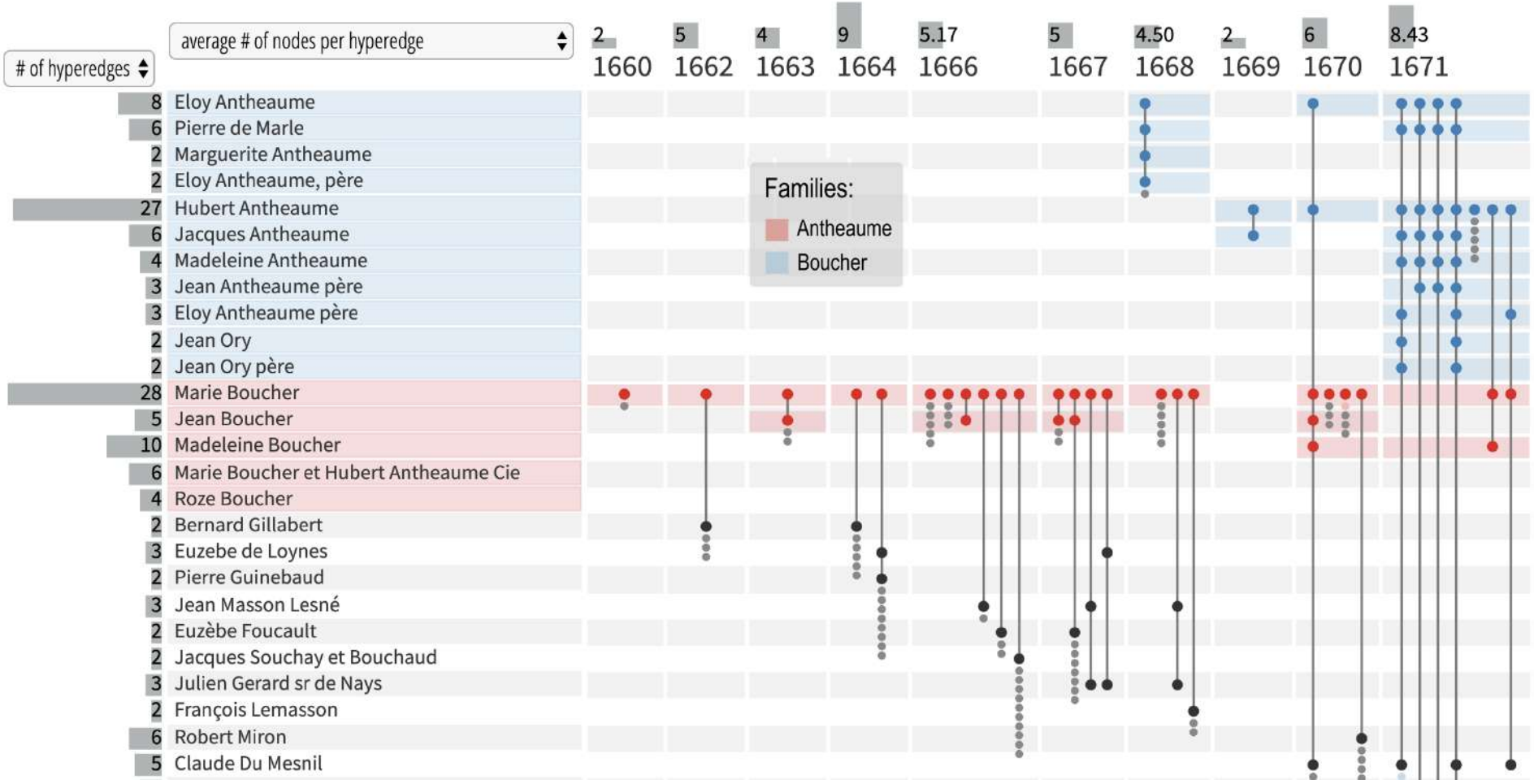
IVAN: Interactive Visual Analysis of Networks

- **Zooming into network to reveal "deep" organization**
  - New algorithm based on concept of graph Slepians that generalizes Laplacian embedding
  - Interactive visualization tool to explore parameter space



IVAN: Interactive Visual Analysis of Networks

- **Taking into account fluctuations of networks (dynamics)**



IVAN: Significant Results and Dissemination to Date

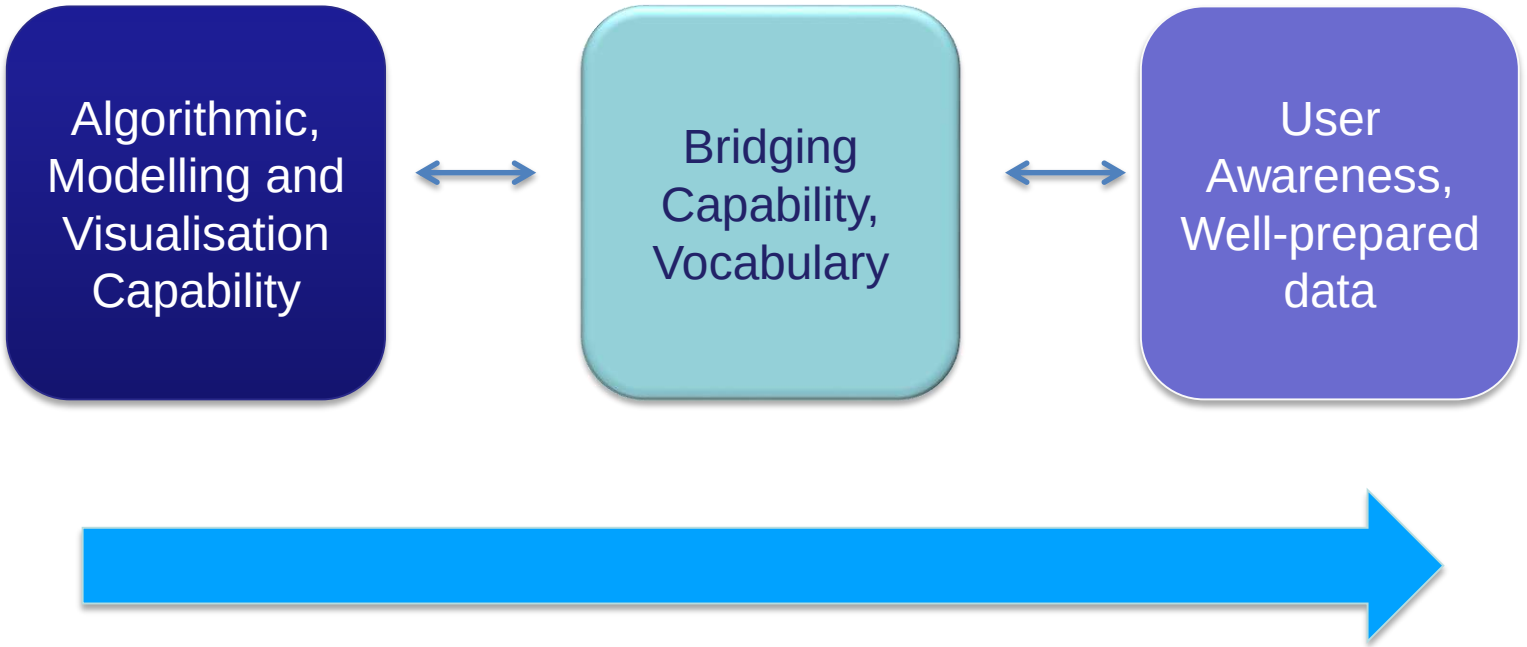
- **Dissemination**
  - Prototype code on git
  - Interactions and collaborations with user groups from neuroscience and social sciences
- **Academic output**
  - Journal paper “Guided Graph Spectral Embedding” accepted to *Network Neuroscience* (Gold OA)

Topic Challenges and Needs

- **Scalability, Efficiency**
- **Interpretability**
  - Lack of ‘ground truth’
- **Evaluation**
  - Limited value of existing benchmarks
- **Role and needs of the human agents**
  - Challenges of new user communities
- **Using the language of the domain**

Possible Roadmap

Three areas of interlinked challenges and opportunities that need to develop in concert



Role of the CHIST-ERA Support

- **Most helpful aspects**
  - Focus of the thematic call
  - Multidisciplinary Collaboration between the partners
  - Transnational dimension
- **Inherent challenges**
  - Can be tricky to align four national projects under one umbrella
    - **Would be good to keep additional national overhead to a minimum (such as additional reporting, proposal phase, etc.)**

Questions

Questions ?