chist-era

Interactive and Visual Analysis of Networks (IVAN)

CHIST-ERA Projects Seminar 2021 (Call 2016: VADMU)
April 13th, 2021
Christian Knoll (University of Vienna)

universität wien

EPFL

Inria

https://ivan.ai

chist-era

Content

- Scientific Background
- Key Challenges and Potential Impact
- Partners
- Scientific Results

chist-era

Scientific Background

- Dynamic Networks are used in many scientific domains, such as
  - neurobiology to model the evolution of the brain according to signals (EEG, fMRI)
  - social sciences to model the evolution of social ties according to documents or sites
- Clustering these networks is important to find regularities, invariants, better understand them, and achieve scalable analysis
- Novel clustering algorithms exist for dynamic networks, but they are black boxes, hard to understand, validate, control, adapt, steer
- The aim of the IVAN project is to make these novel algorithms accessible to a broad range of users and researchers to enable reliable and informed decisions based on the network analysis under uncertainty

chist-era

Key Challenges and Potential Impact

- Development of harmonic analysis for graphs
  - Spectral methods and graph wavelet transforms, hard to understand and to harness
- Directed and zoom-able analysis to nodes of interest
  - Currently not able to analyse parts of the network while watching global network organization
- Comparative network clustering analysis
  - Tool make sense of clusterings, assessing the quality and uncertainty of the process
- Clustering sets of networks and clustered networks
  - Maintain hundreds of clusterings and explore them to understand their variety
- Understanding the needs and user tasks across different communities
- Need for computationally efficient methods

chist-era

Partners

- Torsten Möller
  Visualization and Data Analysis research group, Faculty of Computer Science, University of Vienna
  - Expertise in exploring the issues of modeling under uncertainty and decision making
  - Successful in a collaboration with fisheries scientists and medical image segmentation

universität wien

- Dimitri Van De Ville
  Ecole Polytechnique Fédérale de Lausanne, Institute of Bioengineering
  - Expertise in advanced signal processing, such as dynamic brain networks, wavelet design to obtain tight frames for multiplex graphs, statistical inference and surrogate data generation on graphs

EPFL

- Jean-Daniel Fekete
  Inria Saclay - Ile-de-France - EPI Aviz
  - Expertise in HCI, network visualization, and evaluation methods for visualization
  - Collaborations with leading groups related to network visualization at scale

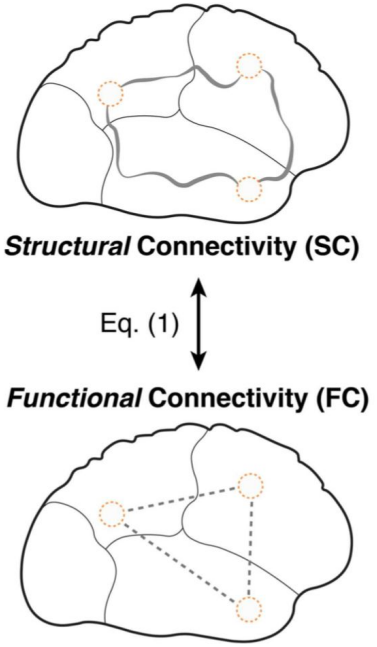
Inria

chist-era

EPFL - progress

Linking brain anatomy and function

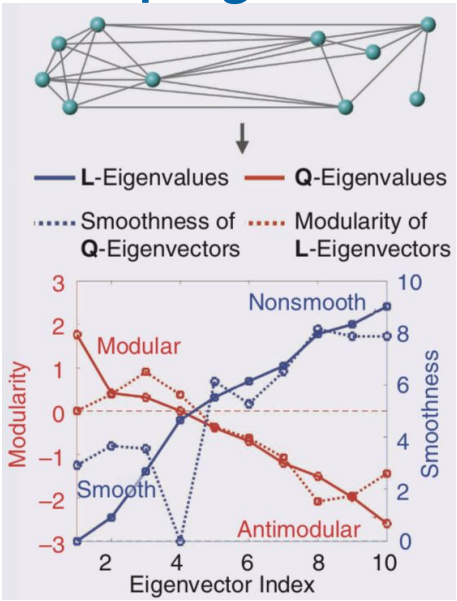
- Which *model* to represent this interplay?
- Are some brain functional *networks* more tightly connected to the underlying anatomy?
- What does the match between anatomy and function tell us abou brain *dynamics*?



Liégeois et al. 2020. Network Neuroscience 4(4), 1235–1251.

chist-era

EPFL - progress



Petrovic, Liégeois* et al. 2020. IEEE Signal Processing Magazine 37 (6).*

Making Graph Signal Processing (GSP) *community*-aware

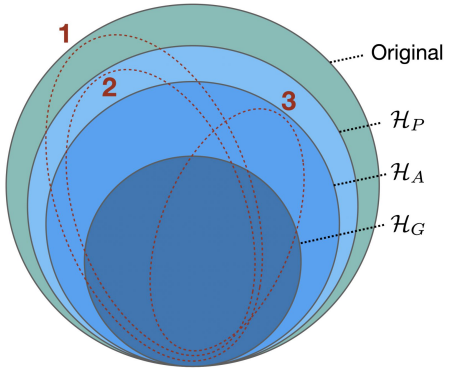
- Classical GSP is based on the graph *Laplacian*
- Instead, we propose to focus on the graph *modularity* matrix in order to account for *communities*
- Modularity-based GSP extracts information to which Laplacian-based GSP is blind (e.g., in the brain)

chist-era

EPFL - progress

How to interpret temporal fluctuations in time series?

- We review the rich literature of time-series *null models*
- Each model preserves a *subset* of statistical properties from the original data: H_G, H_A, H_P
- Null-model testing allows us to *finely characterize* time series properties giving rise to observed fluctuations



Liégeois et al. 2021. Submitted.

chist-era

Uni Vienna - progress

- Requirement analysis: needs and user tasks
- Multiple full-day Creative Visualization-Opportunities (CVO) Workshops*

Eight DH Researchers

Historic trading networks

Most common interests:

- the grouping (clustering) of nodes and edges in the network
- dynamic changes of the networks over time and their effect on specific parts
- following individual paths of nodes as well as
- finding patterns and comparing different dynamic networks.

* *A Framework for Creative Visualization-Opportunities Workshops*
E. Kerzner, S. Goodwin, J. Dykes, S. Jones and M. Meyer
in IEEE Transactions on Visualization and Computer Graphics, vol. 25, no. 1, pp. 748-758, Jan. 2019, doi: 10.1109/TVCG.2018.2865241.

chist-era

Uni Vienna - progress

- Requirement analysis: needs and user tasks
- Multiple full-day Creative Visualization-Opportunities (CVO) Workshops*

Eight DH Researchers

Historic trading networks

Eight Medical Researchers

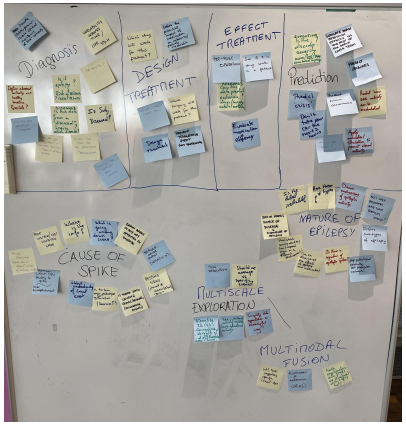
EEG data of Epilepsy patients

* *A Framework for Creative Visualization-Opportunities Workshops*
E. Kerzner, S. Goodwin, J. Dykes, S. Jones and M. Meyer
in IEEE Transactions on Visualization and Computer Graphics, vol. 25, no. 1, pp. 748-758, Jan. 2019, doi: 10.1109/TVCG.2018.2865241.

CVO Workshop in Geneva

September 2020 - Creativity workshop on EEG & Epilepsy

- *Formalize current challenges for understanding epilepsy*
- Multidisciplinary attendance: clinicians, engineers, visualization experts, neurologists
- Follow-up paper based on our experience of running CVO workshops*

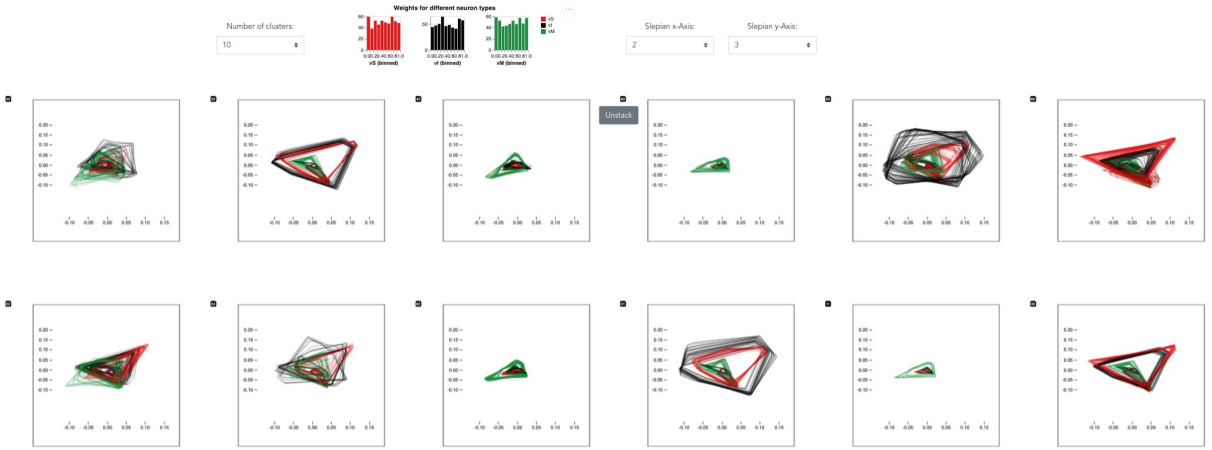


* *Extending Recommendations for Creative Visualization-Opportunities Workshops.*
C. Knoll, A. Çetin, T. Möller and M. Meyer
IN: 2020 IEEE Workshop on Evaluation and Beyond - Methodological Approaches to Visualization (BELIV), Salt Lake City, UT, USA, 2020, pp. 81-88
https://doi.org/10.1109/BELIV51497.2020.00017

chist-era

Uni Vienna - progress

- Guided graph spectral embedding: EPFL proposed novel algorithm to zoom on graphs*
- C. Elegans connectome dataset (279 interconnected neurons)
- Cooperation matrix for defining focus on specific neuron types
- Visual tool for exploring the sampled input space



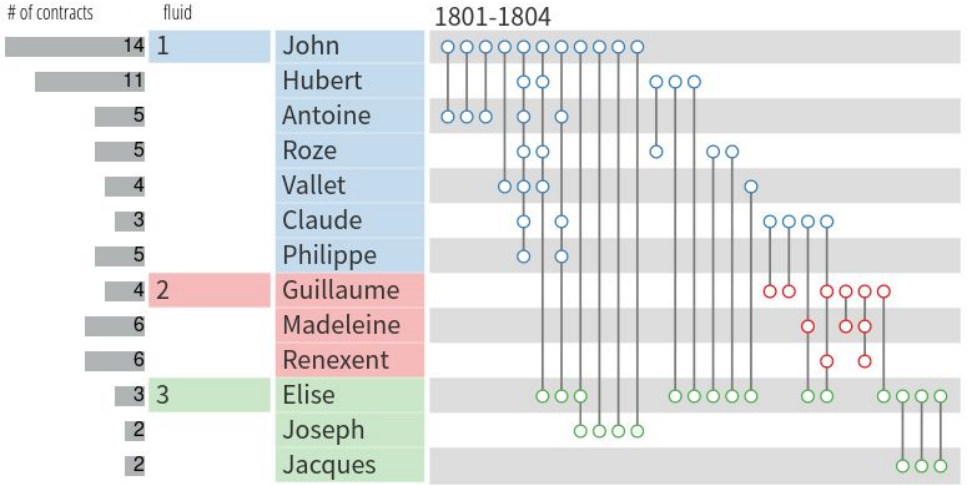
* *Guided graph spectral embedding: Application to the C. elegans connectome*
Miljan Petrovic, Thomas A. W. Bolton, Maria Giulia Preti, Raphaël Liégeois, and Dimitri Van De Ville
Network Neuroscience 2019 3:3, 807-826 (https://www.mitpressjournals.org/doi/full/10.1162/netn_a_00084)

chist-era

Inria - progress

Integrating Prior Knowledge in Mixed Initiative Social Network Clustering

- Address the *black box* problem of network clustering
- Expert have *prior knowledge* (PK) on their networks
- How can we include it in the clustering process?
- Start with a network, add PK and run an ensemble of algorithms on it
- Keep only the solutions that match the PK
- Then, finish the clustering



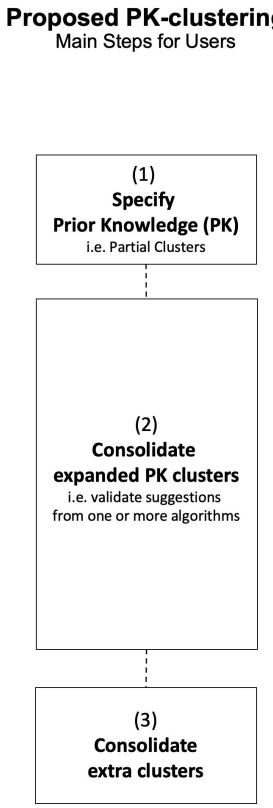
chist-era

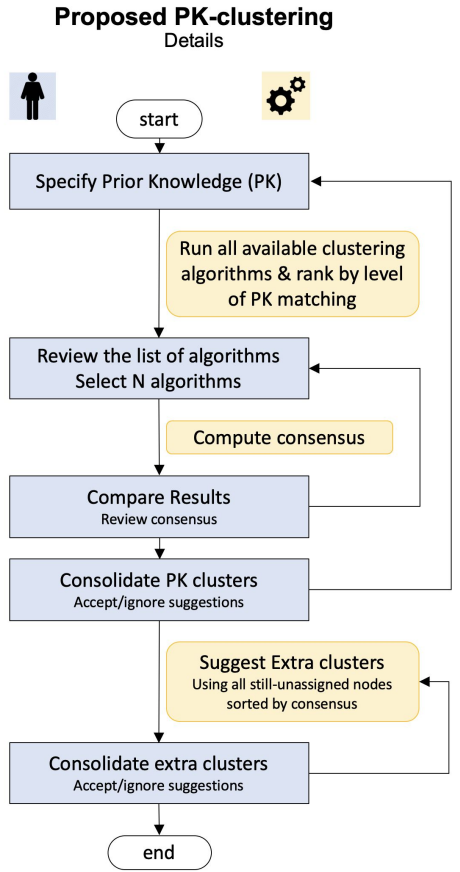
PK-Clustering

Classical clustering approaches:

- Automatic
- Ensemble
- Semi-supervised

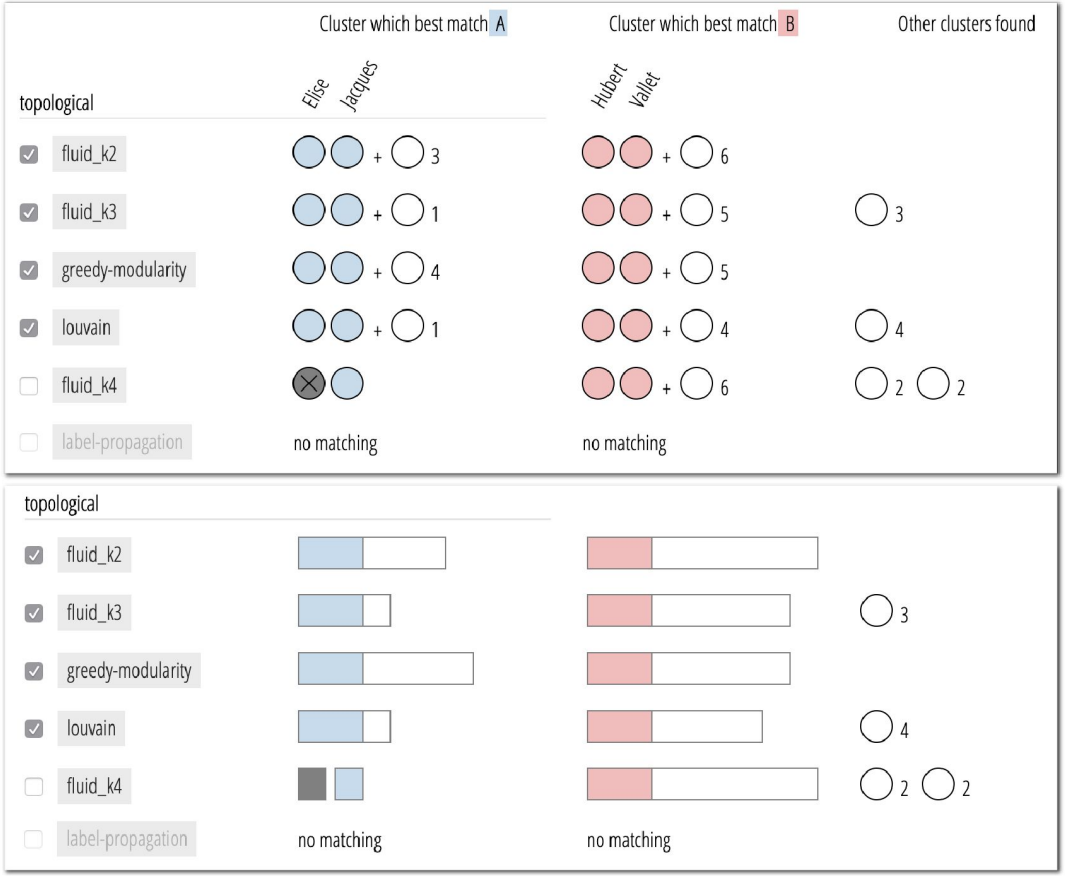
We propose a new approach: **PK-Clustering**, which combines aspects of the three options in order to **give more control to the user in the analysis loop.**





Algorithms ranking view

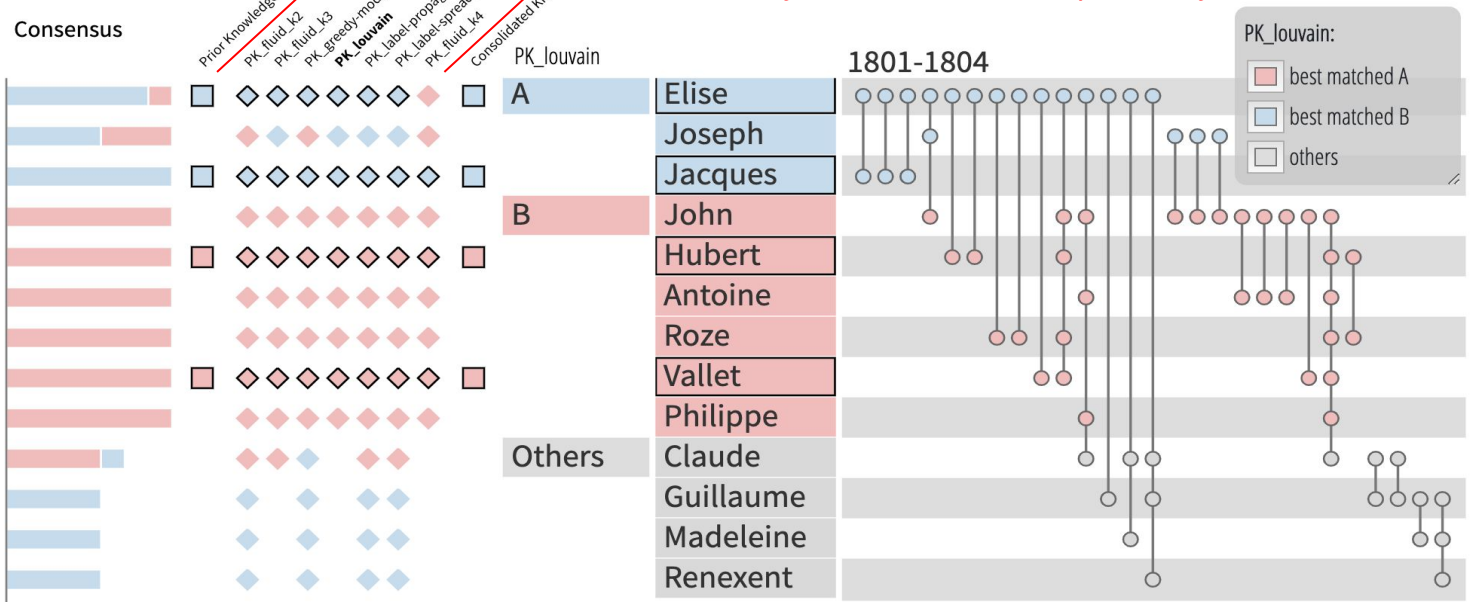
- First, **all** the clustering algorithms are run
- The results are ranked by the **edit distance** between the **PK** and the clustering
  - Ties are ranked
- The user then selects a set of algorithms she thinks are appropriate regarding the PK and the contents of the clusters



Validation of the prior knowledge groups

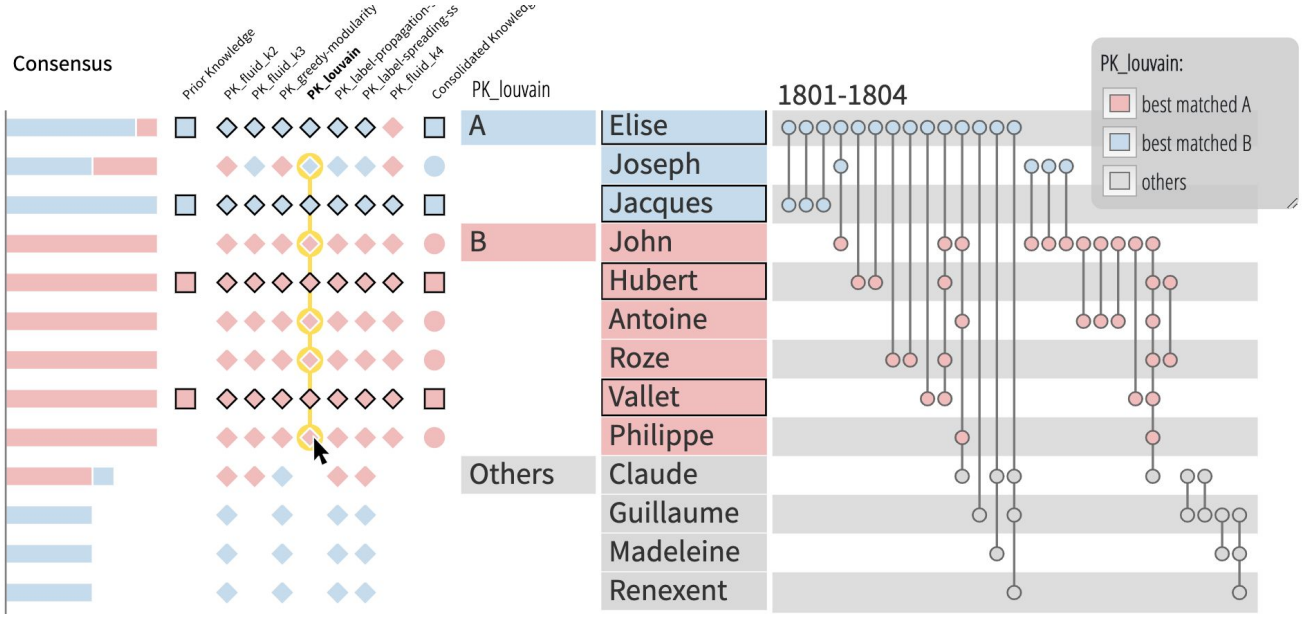
Selected algorithms

The consolidated knowledge is the final partition validated by the user with the help of the system



chist-era

Validation of the prior knowledge groups



The system can give **suggestions (circles)** based on the **consensus** of the algorithms

The user can:

- **validate these suggestions**, or
- copy the results of any algorithm, or
- manually select the cluster of one person

chist-era

Possible biases

PK-Clustering

- Users have a lot of influence and can overwrite labels
- If the PK is wrong, the system will suggest clusterings that match it

Traditional clustering

- **Priming bias**
- PK is often used to evaluate the results
- Do not propose alternate views

chist-era

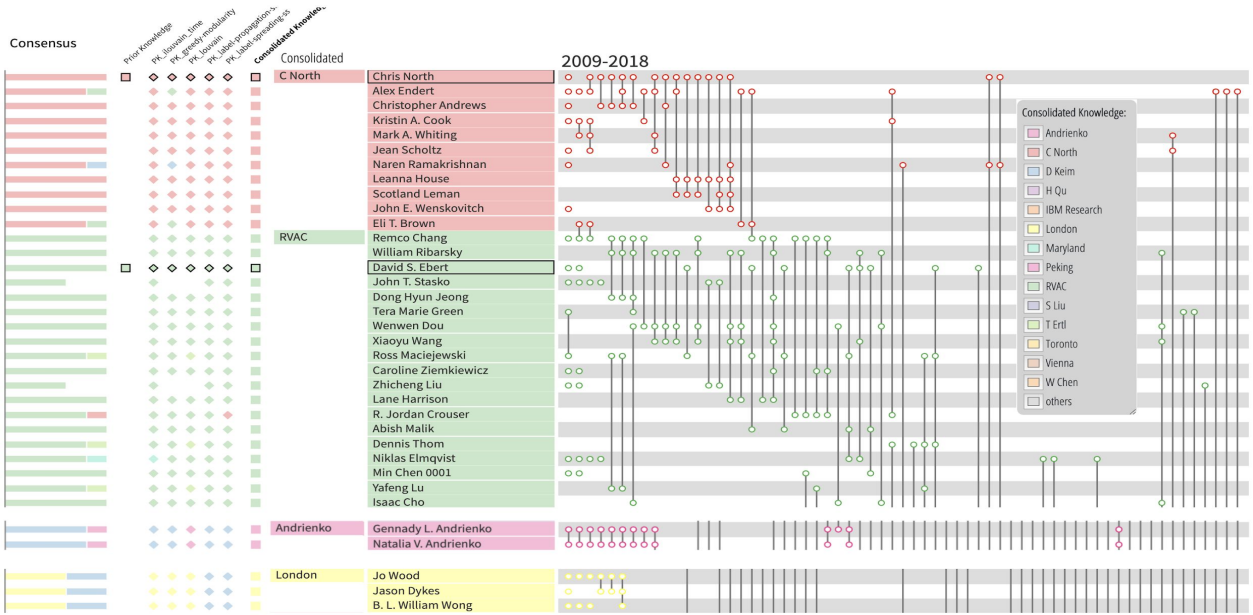
Increasing transparency

1) Summary Table
2) Summary report (e.g. of Marie Boucher use case)

Prior Knowledge: 3 groups, 3 persons
Final: 4 groups, 37 persons (out of 38) consolidated as follows:

- **12 (32.4%) using the consensus**:
  - 10 (27.0%) with an agreement of 8 algorithms
  - 2 (5.4%) with an agreement of 5 algorithms
- 18 (48.6%) using the result of 1 algorithm:
  - 18 (48.6%) with PK_ilouvain_time
- **4 (10.8%) manually** consolidated

13 algorithms considered
9 algorithms reviewed and compared



Project webpage with paper, video and demo
www.aviz.fr/Research/pkclustering

chist-era

Thank you for your attention!

Further information can be found on our website: https://ivan.ai

universität wien

Torsten Möller: torsten.moeller@univie.ac.at
Christian Knoll: christian.knoll@univie.ac.at

EPFL

Dimitri Van De Ville: dimitri.vandeville@epfl.ch
Raphaël Liégeois: raphael.liegeois@epfl.ch

Inria

Jean-Daniel Fekete: Jean-Daniel.Fekete@inria.fr
Paola Valdivia: paola.valdivia@inria.fr