# A unified approach to mapping and clustering of bibliometric networks

**Nees Jan van Eck**, Ludo Waltman, and Ed C. M. Noyons
Centre for Science and Technology Studies, Leiden University

STI 2010, Leiden, The Netherlands
September 9, 2010

Universiteit Leiden

CWTS

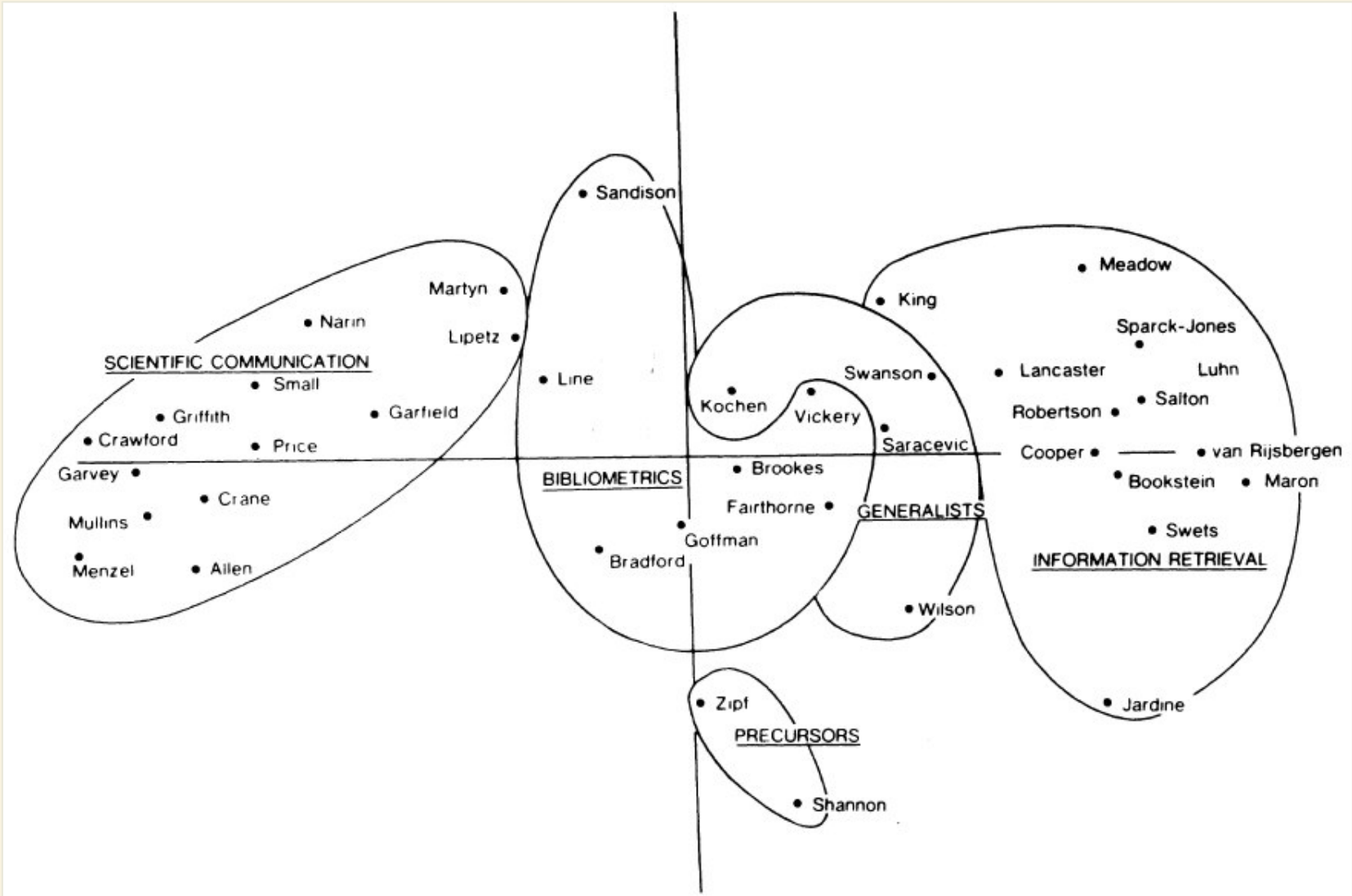
# Outline

- Introduction
- Unified approach to mapping and clustering
- Application
- Conclusion

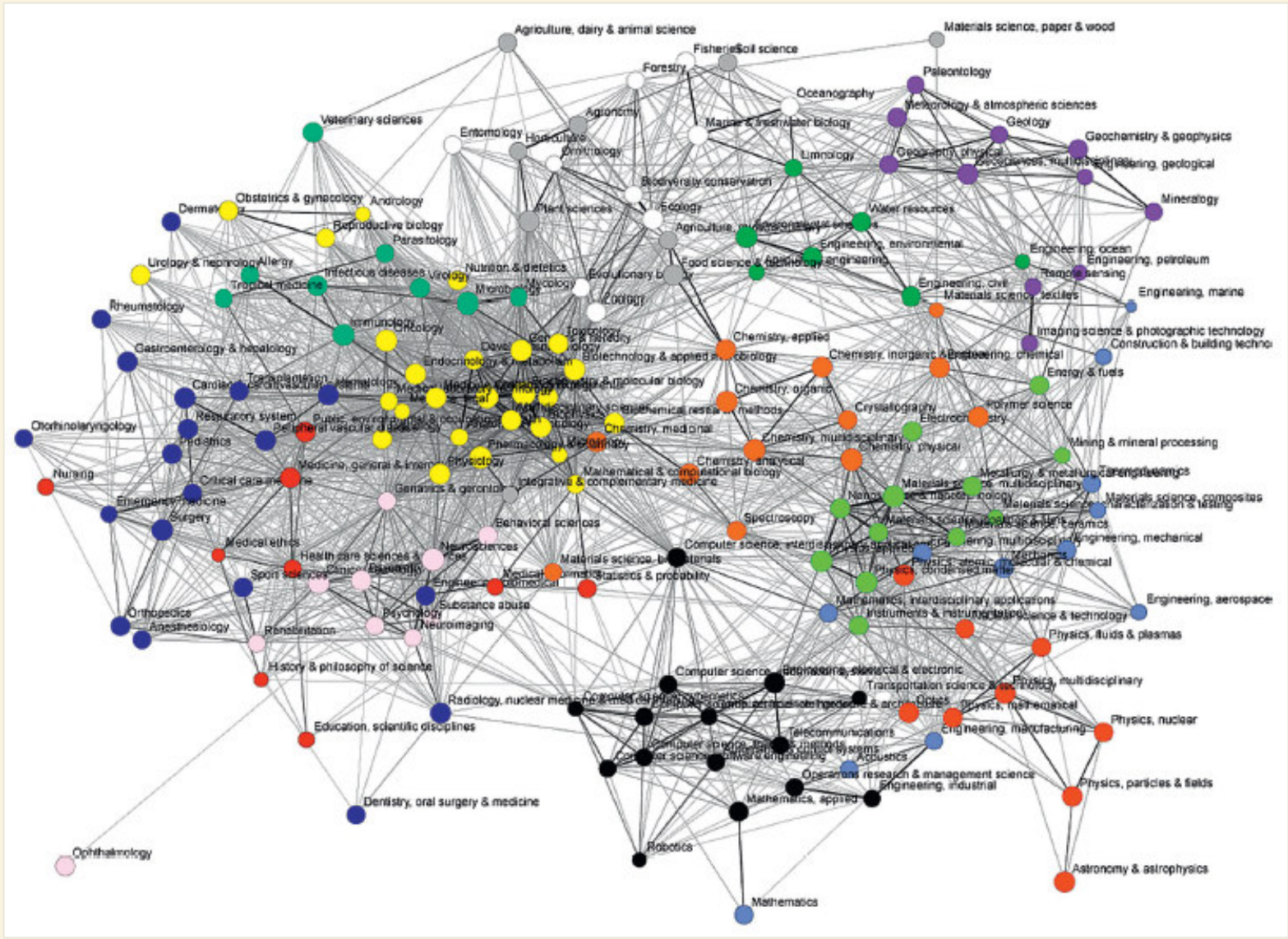
# Introduction

- Mapping and clustering are used to address questions such as:
  - What are the main topics within a scientific domain?
  - How do these topics relate to each other?
  - How has a scientific domain developed over time?
- Mapping and clustering are often used together

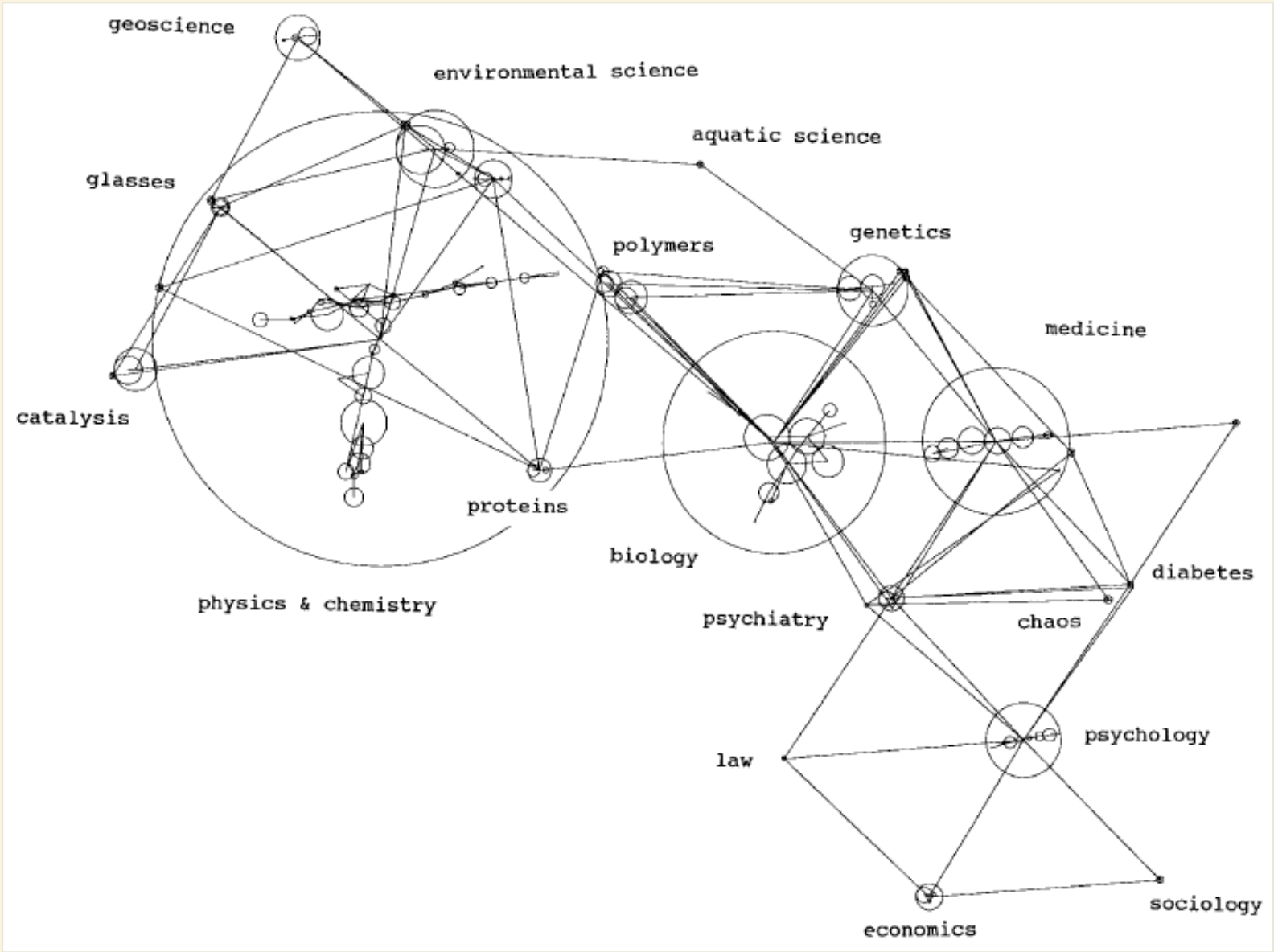
# Example: White & Griffith (1981)



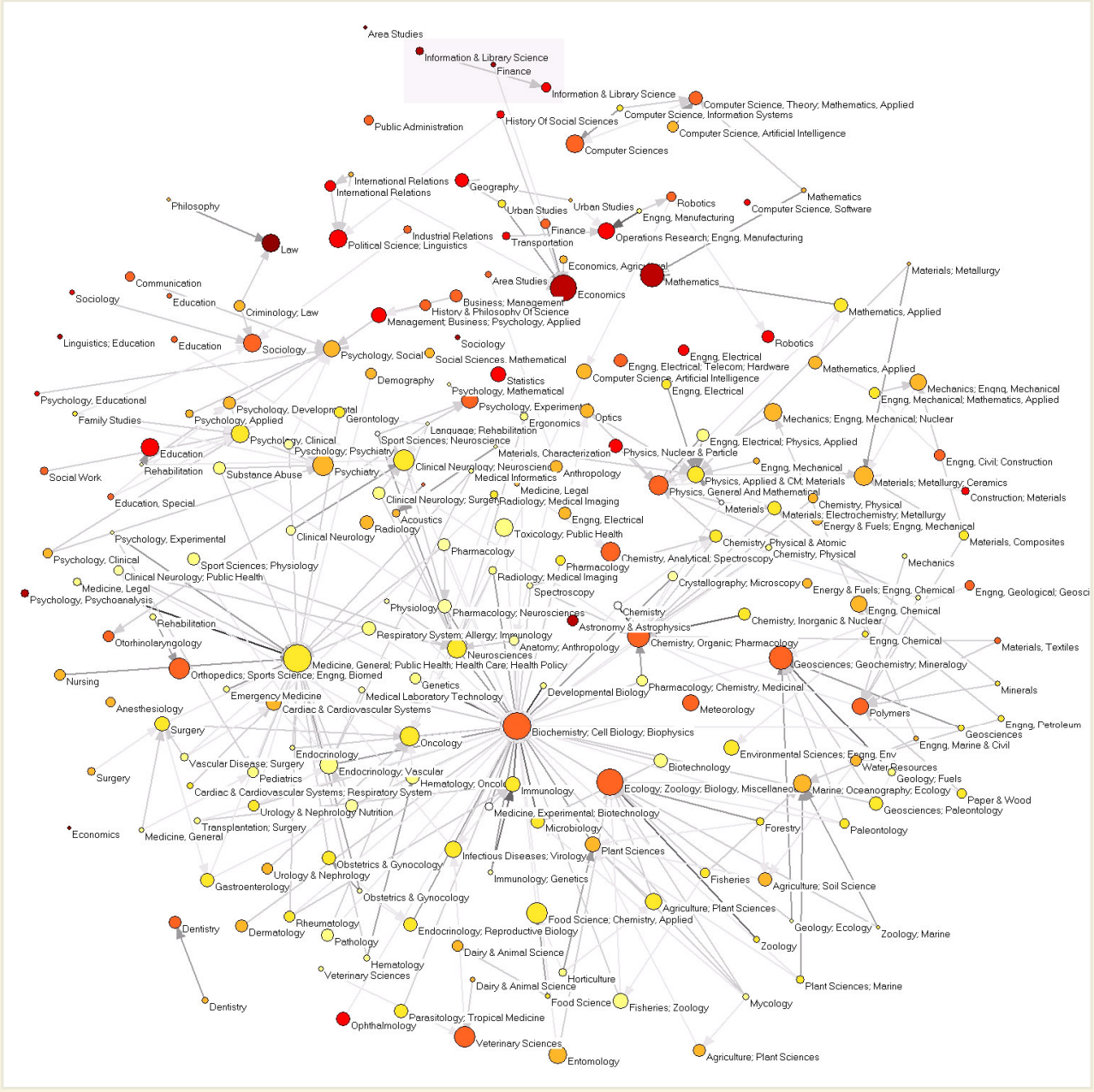
# Example: Leydesdorff & Rafols (2009)



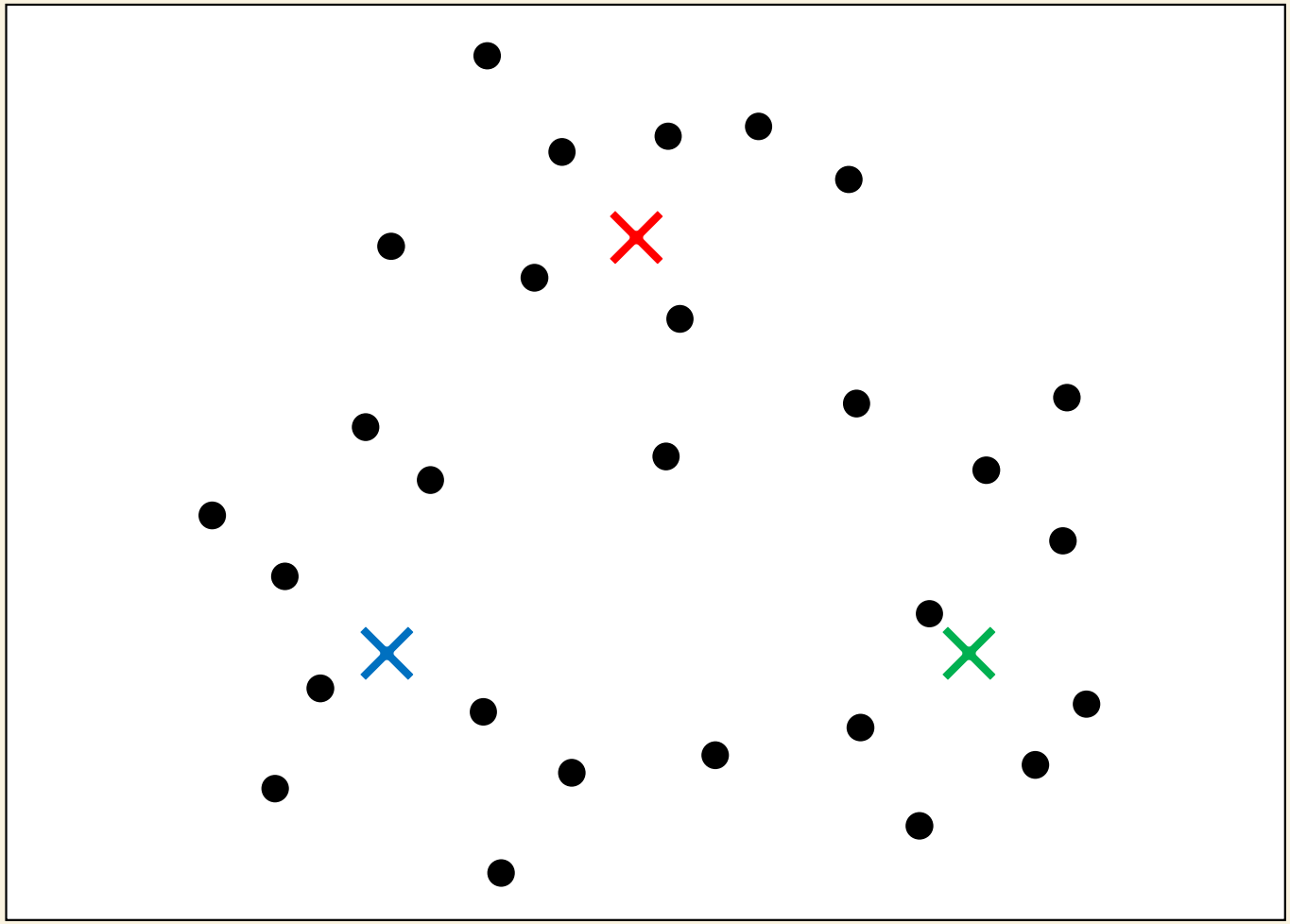
# Example: Small (1999)



# Example: Boyack, Klavans, & Börner (2005)



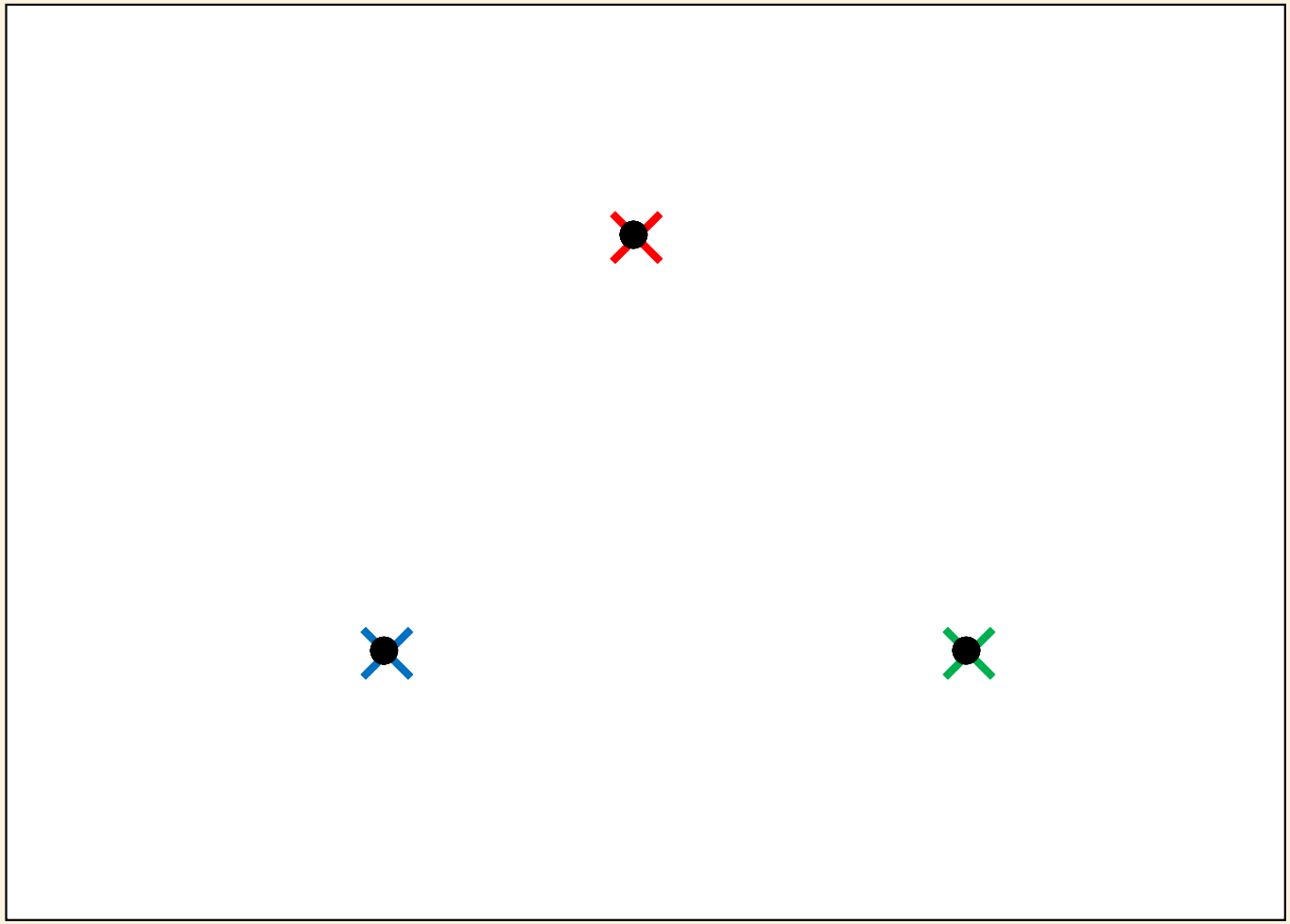
# Unified approach: Clustering seen as mapping in a restricted space



Universiteit Leiden

CWTS

# Unified approach: Clustering seen as mapping in a restricted space



# Unified approach to mapping and clustering

Minimize $$V(x_1,\ldots,x_n)=\sum_{i<j}\frac{2mc_{ij}}{c_i c_j}d_{ij}^2-\sum_{i<j}d_{ij}$$

where

$n$: number of nodes in the network

$m$: number of links in the network

$c_{ij}$: number of links between nodes $i$ and $j$

$c_i$: number of links of node $i$

## Mapping

$x_i$: vector denoting the location of node $i$ in a $p$-dimensional map

$$d_{ij}=\left\|x_i-x_j\right\|=\sqrt{\sum_{k=1}^{p}(x_{ik}-x_{jk})^2}$$

## Clustering

$x_i$: integer denoting the cluster to which node $i$ belongs

$$d_{ij}=\begin{cases}0 & \text{if } x_i=x_j\\ 1/\gamma & \text{if } x_i\neq x_j\end{cases}$$

$\gamma$: resolution parameter

# Unified approach: Mapping

- Equivalent to the VOS mapping technique
- Closely related to multidimensional scaling (Van Eck et al., forthcoming in *JASIST*)

# Unified approach: Clustering

- Equivalent to a weighted and parameterized variant of modularity-based clustering

Maximize $$\hat{V}(x_1,\ldots,x_n) = \frac{1}{2m}\sum_{i<j}\delta(x_i,x_j)w_{ij}\left(c_{ij} - \gamma\frac{c_i c_j}{2m}\right)$$

where

$\delta(x_i, x_j)$ equals 1 if $x_i = x_j$ and 0 otherwise

$$w_{ij} = \frac{2m}{c_i c_j}$$

- Parameter $\gamma$ helps to deal with the resolution limit problem of modularity-based clustering

# Application

- Library and information science
- 1242 most frequently cited publications in the period 1999–2008

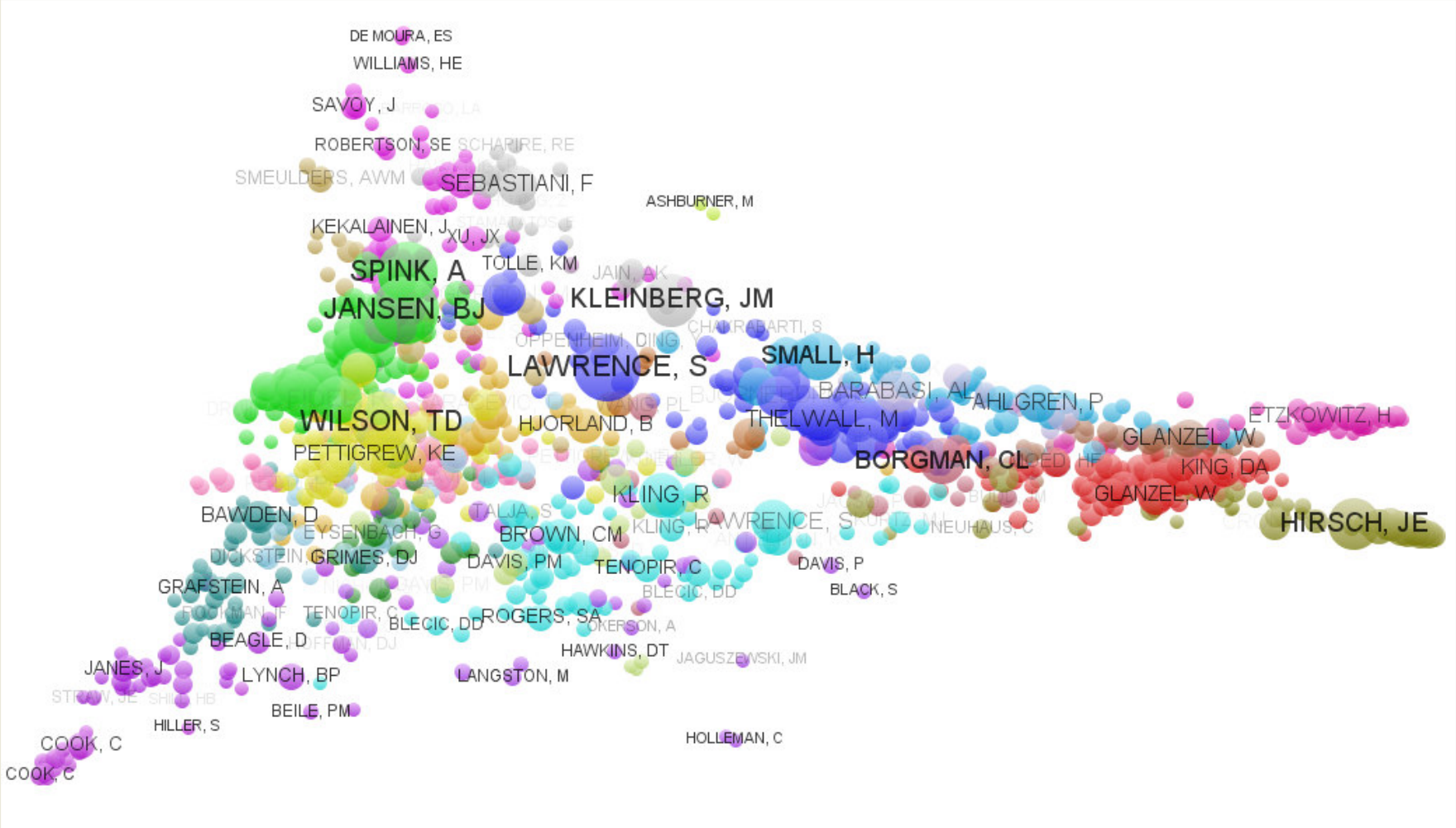
Universiteit Leiden

# Application: Clustering results

- 25 clusters of frequently cited publications
- 8 clusters on bibliometrics/scientometrics:

| No of pub. | Important authors | Main topics |
|---|---|---|
| 123 | Rousseau, R.; Glänzel, W.; Moed, H.F.; Van Raan, A.F.J. | Citation analysis; research evaluation; general scientometric topics |
| 101 | Thelwall, M.; Vaughan, L.; Bar-Ilan, J.; Wilkinson, D. | Webometrics |
| 73 | Leydesdorff, L.; Chen, C.M.; White, H.D.; Small, H. | Mapping and visualization of science |
| 53 | Egghe, L.; Burrell, Q.L.; Daniel, H.D.; Glänzel, W. | *h*-index; citation distributions; Google Scholar |
| 48 | Glänzel, W.; Cronin, B.; Bozeman, B.; Shaw, D. | Scientific collaboration; co-authorship |
| 46 | Meyer, M.; Leydesdorff, L.; Tijssen, R.J.W.; Zimmermann, E. | Science and technology studies; patent analysis |
| 26 | Nisonger, T.E.; Cronin, B.; Shaw, D.; Wilson, C.S. | Studies of the library and information science field |
| 14 | Newman, M.E.J.; Barabasi, A.L.; Albert, R.; Jeong, H. | Complex networks; scientific collaboration networks |

# Application: Combined mapping and clustering results




Universiteit Leiden

# Conclusion

- Advantages of a unified approach to mapping and clustering:
  - Consistent mapping and clustering results
  - Consistent analysis at different aggregation levels
  - Methodological simplicity and transparency
- Ongoing work:
  - Extension of unified approach to clustering with overlap
  - Application of unified approach for constructing a journal classification system for Scopus database

# Software and paper

- Software for unified approach:
  - VOSviewer
  - VOS mapping and clustering programs
- Software freely available at www.vosviewer.com
- Paper forthcoming in *Journal of Informetrics*

# Thank you for your attention!

# References


Boyack, K.W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics*, *64*(3), 351–374.

Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *JASIST*, *60*(2), 348–362.

Newman, M.E.J. (2004). Analysis of weighted networks. *Physical Review E*, *70*(5), 056131.

Small, H. (1999), Visualizing science by citation mapping. *JASIST*, *50*(9): 799–813.

Van Eck, N.J., Waltman, L., Dekker, R., & Van den Berg, J. (in press). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *JASIST*.

White, H.D., & Griffith, B.C. (1981). Author co-citation: A literature measure of intellectual structure. *JASIST*, *32*(3), 163–171.


Universiteit Leiden

CWTS