WP4

Results and experiments with socio-scientific indicators:

overview of D4.3 and outlook for D4.4

Selection process

1. Pool of indicators for case studies
2. Partner needs and problem secifications
3. Selection pressure on the initial pool of indicators
4. Experimenting with a specific set of indicators for each case study

The making of a socio-scientific indicator

1. Pool of proposed indicators

Network measures for grasping net/overall relatedness in diverse socio-scientific networks

$$d_1(a,b) =_{def} \alpha d_{path}^{co-author}(a,b) + \beta d_{path}^{cit.flow}(a,b) + \gamma d_{path}^{co-cit}(a,b)$$

$$d_2(a,b) =_{def} \alpha d_{structural}^{co-author}(a,b) + \beta d_{structural}^{cit.flow}(a,b) + \gamma d_{structural}^{co-cit}(a,b)$$

The making of a socio-scientific indicator

2. The indicators required by Frontiers

Goal: test hypotheses on network effects in peer review

Dimensions:

- (1) Position and (2) distance of actors affect review scores

Data provided:

- (a) „Review database" of Frontiers
- (b) Co-author network of Frontier contributors (authors/reviewers) basen on an extended search (Fr + Scopus)

The making of a socio-scientific indicator

3. What do we have in this setting?

Indicators selected by the problem setting and available data:

$$d_1(a,b) =_{def} \alpha d^{co-author}_{path}(a,b) + \beta d^{cit.flow}_{path}(a,b) + \gamma d^{co-cit}_{path}(a,b)$$

$$d_2(a,b) =_{def} \alpha d^{co-author}_{structural}(a,b) + \beta d^{cit.flow}_{structural}(a,b) + \gamma d^{co-cit}_{structural}(a,b)$$

The making of a socio-scientific indicator

3. Construction of the underlying network

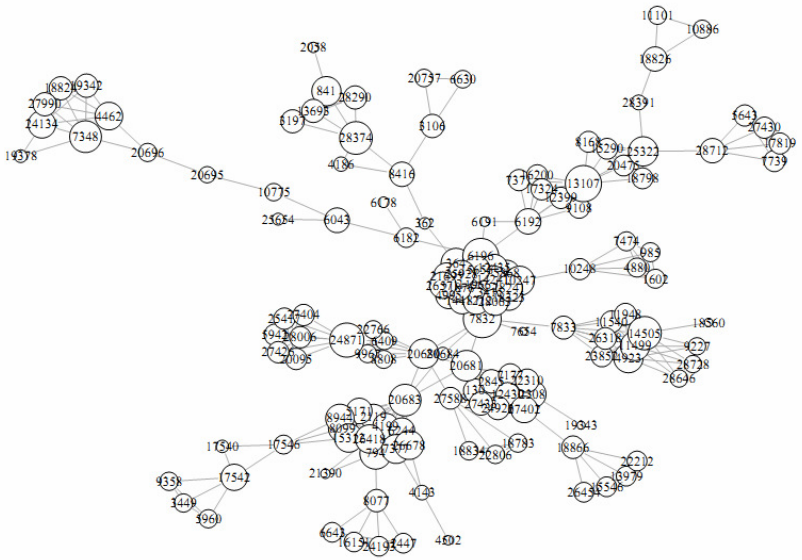
As the reference network representing actor relatedness, the co-author graph of anonimized actors (authors, reviewers) has been built. Co-author relations were obtained from two sources: (1) the author–paper table of Frontiers pubs and (2) the author–coauthor table of Frontiers contributiors (authors, reviewers) retrieved from Scopus.

As constrained by the two sources, this exercise resulted in an (a) unweighted (and undirected) graph with (b) **n = 18 958 authors**. The graph, as in the most standard case, consisted of a giant component along with several small „islands". For our study, we selected the **giant component containing n = 15 842 (~ 16 000)** actors.

The making of a socio-scientific indicator

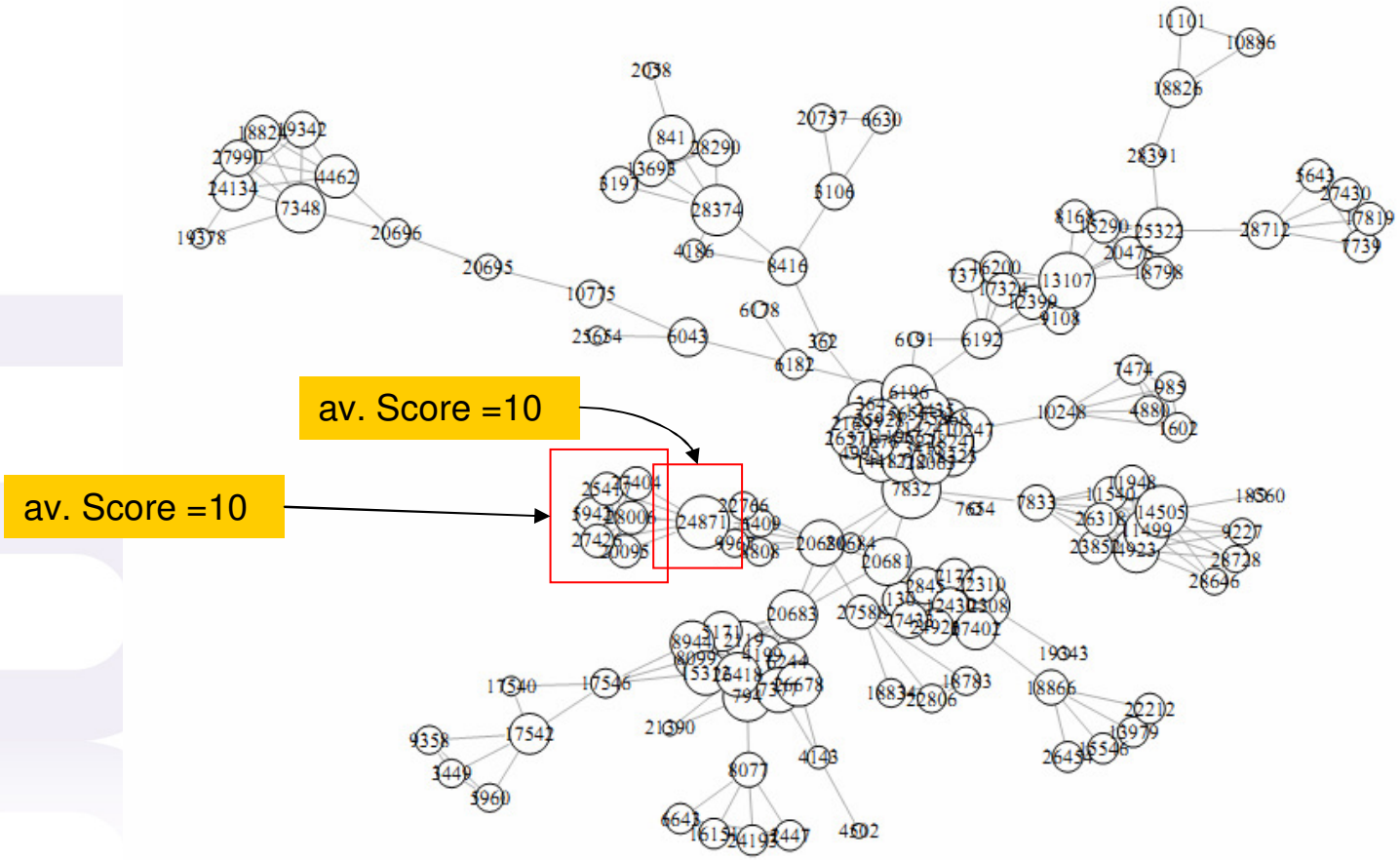
3. Hypothesis-family 1: Positional measures

For „structural similarity/distance”, calculating diverse centrality values of actors



Peer review

Problem with traditional perspective: „the more central the **author** is, the more awarded in peer review” (Scores to papers, not authors)



The making of a socio-scientific indicator

3. Hypothesis-family 1: Positional measures

- Solution to the difficulty above: turn it upside down!

Paper centrality (instead of author ~):

for each paper P with authors {A1, … , An} and author centralities AC = { C(A1), …, C(An)}, the maximum value of AC was obtained along each measures.

- New question, operationalized: whether reviewer scores for papers reflect the authors include high centrality ones.

- Links and scores made independent, empirically commensurable

The making of a socio-scientific indicator

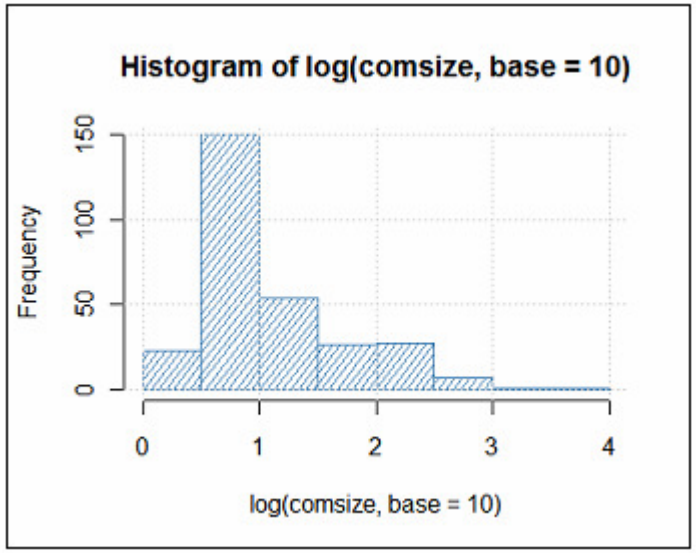
What is the big deal for WP4?

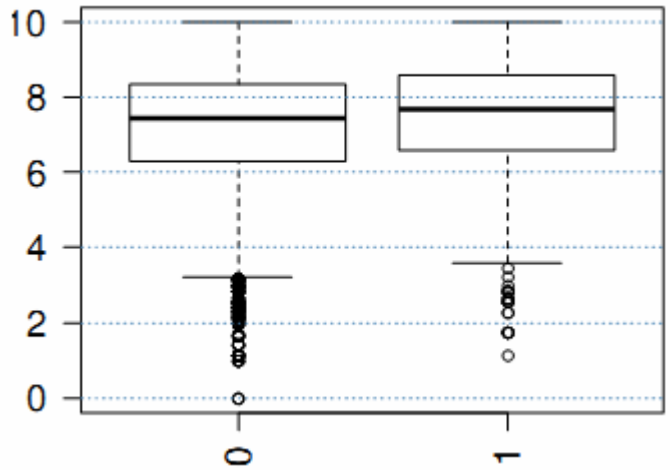
The making of a socio-scientific indicator

- We took a set of measures (centralities) describing social standing of actor in scientific communities
- We applied it to characterize publication data, namely, introduced a new parameter for papers (turn of perspective, novelty)
- The measure can be easily implemented in SISOB
- The SISOB system, with this functionality, can serve peer-review-like case studies, analyses
- We also simulated the analyses → D9.2

Four experimental designs

	bw.paper	d.paper	c.paper	e.paper	score.paper
bw.paper	1	0,830893	0,684468	0,551389	-0,007728
d.paper	0,830893	1	0,641275	0,530933	0,01020766
c.paper	0,684468	0,641275	1	0,868645	-0,0003843
e.paper	0,551389	0,530933	0,868645	1	-0,0142683
score.paper	-0,00773	0,010208	-0,00038	-0,01427	1





1. Pool of proposed indicators

Advancing *non-network* measures for capturing (1) types of mobility
(2) mobility-related performance (⇨ needs and data of FR)

Two examples:

(2) H-core dynamics: $$H_{dyn} =_{def} \frac{1}{N_P} \sum_{i=1}^{N_p} (Y_i - Y_{\min})$$

(1) Career entropy: $$\frac{1}{N} \sum_{i=1}^{m} p_i \ln p_i$$

2. Requried indicators for the case study (so far)

(1) mobility-related performance/types of mobility:

Rank shift: An auxiliary „indicator” for measuring **vertical** mobility. A set of indicators for HEI ranking + method

(2) types of mobility: measuring **thematic mobility** (specific task)

Rank shift indicator for the ranking of HEIs

Based on the data compiled by Thomson Evidence on UK Higher Edu Institutions (HEIs data), a (system of) time-variant ranking(s) is to be constructed with the following minimal set of features:

- A separate ranking is required (1) for each year and (2) for both disciplinary categories provided (natural sciences, engineering).
- The **ranking system** should provide **means for registering significant career steps** (through e.g. derived threshold values or scales for each ranking).

Rank shift indicator for the ranking of HEIs

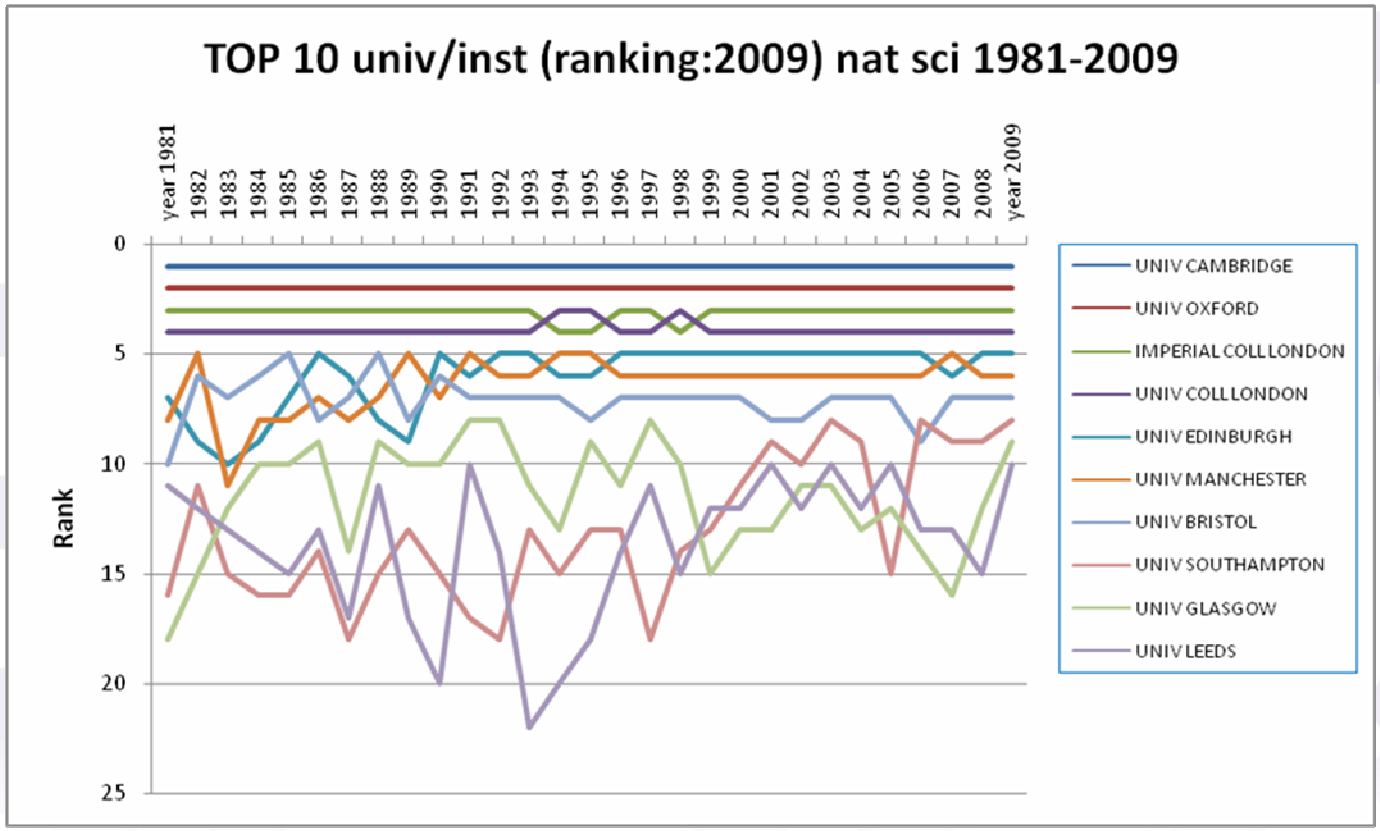
- Based on the evaluative indicator „cumulative impact weighted productivity”

$$CIWP(HEI, year) =_{def} \sum_{i \le year} RI(HEI, i) \times P(HEI, i)$$

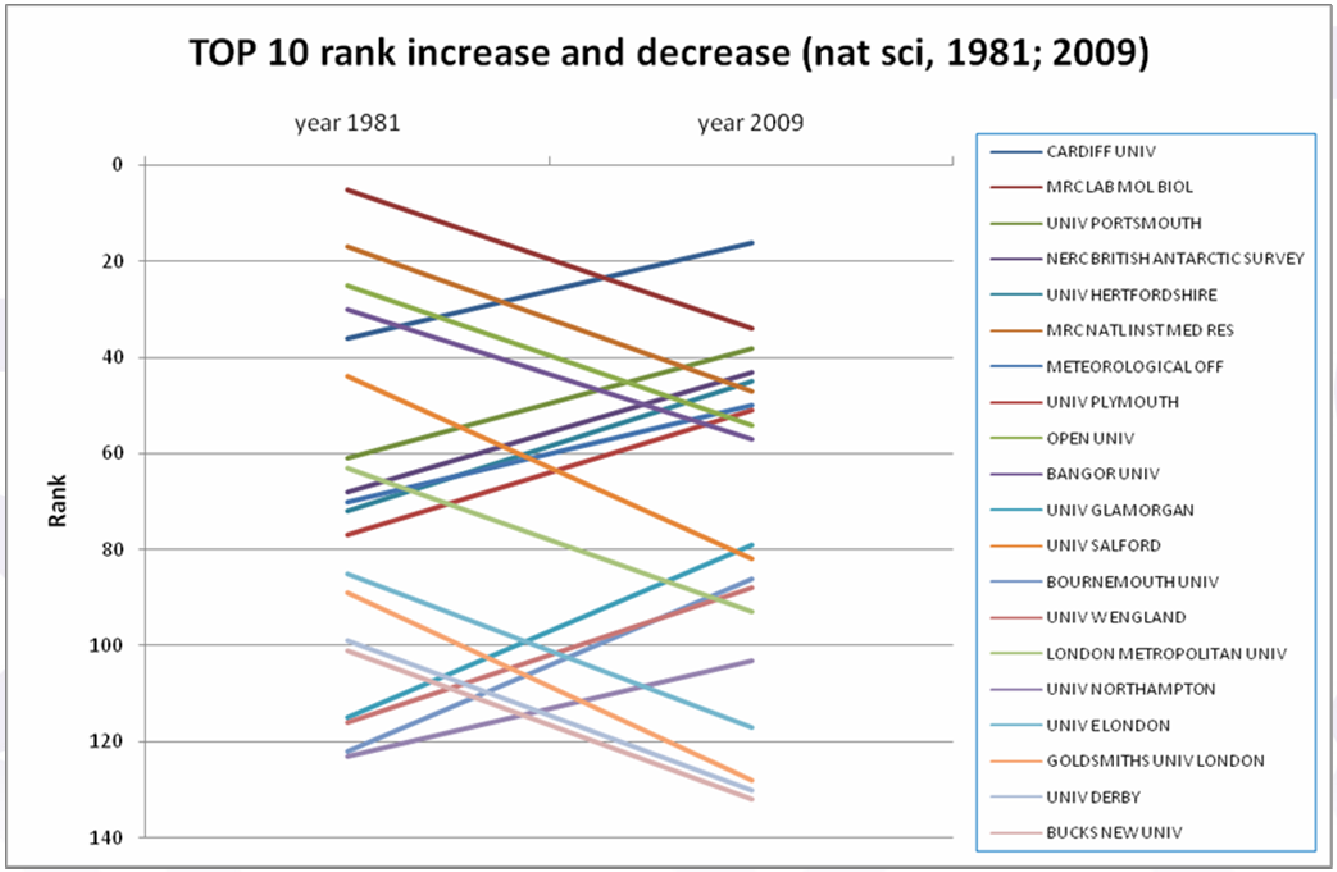
- Problem: skewed distributions, ordinal ranks suppress distances. Solution: find a ranking method reflecting quality shifts along the list. Comparison of distances.

- (1) Percentile-based ranking
- (2) Rank shift indicator
- (3) **Ranking based on the „characteristic scores and scales”**- forming „internally comparable” groups

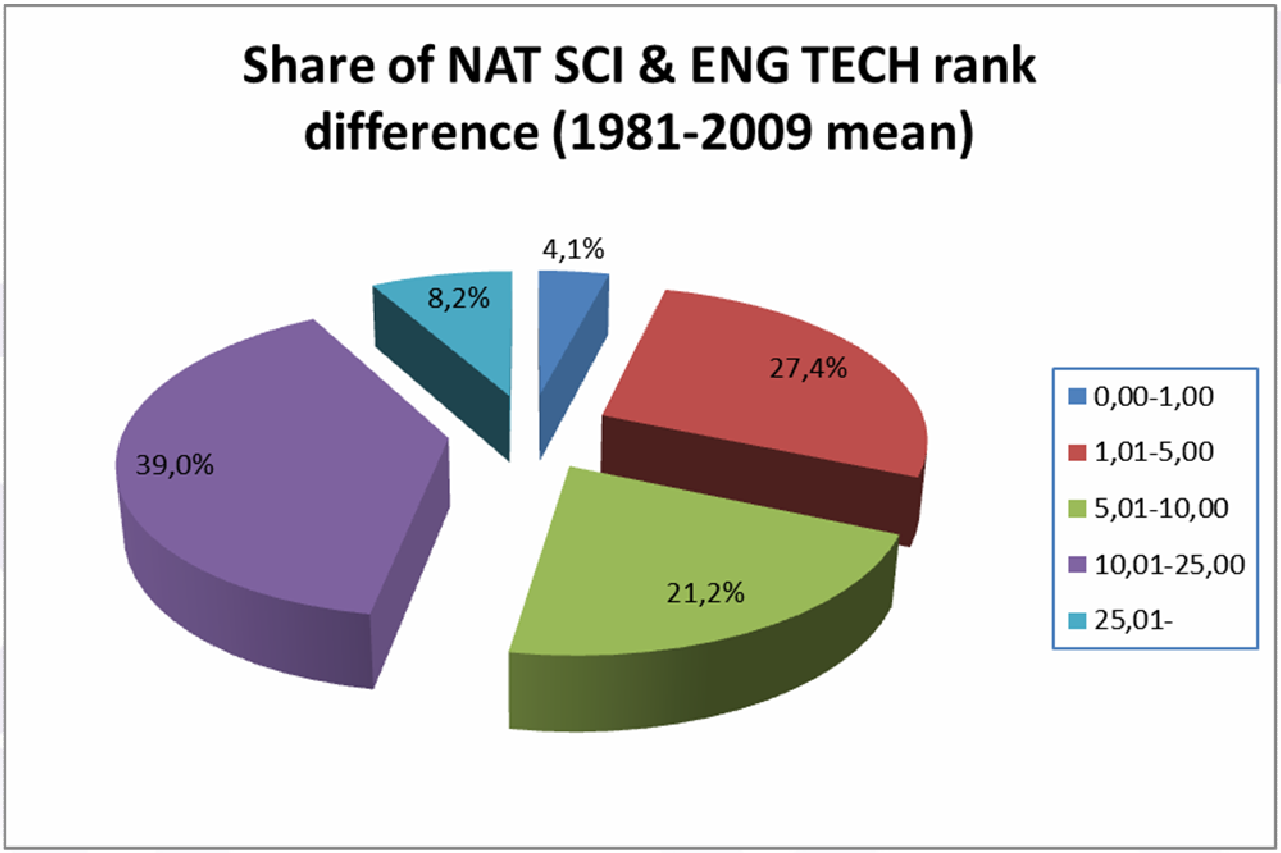
Actual ranking experiments



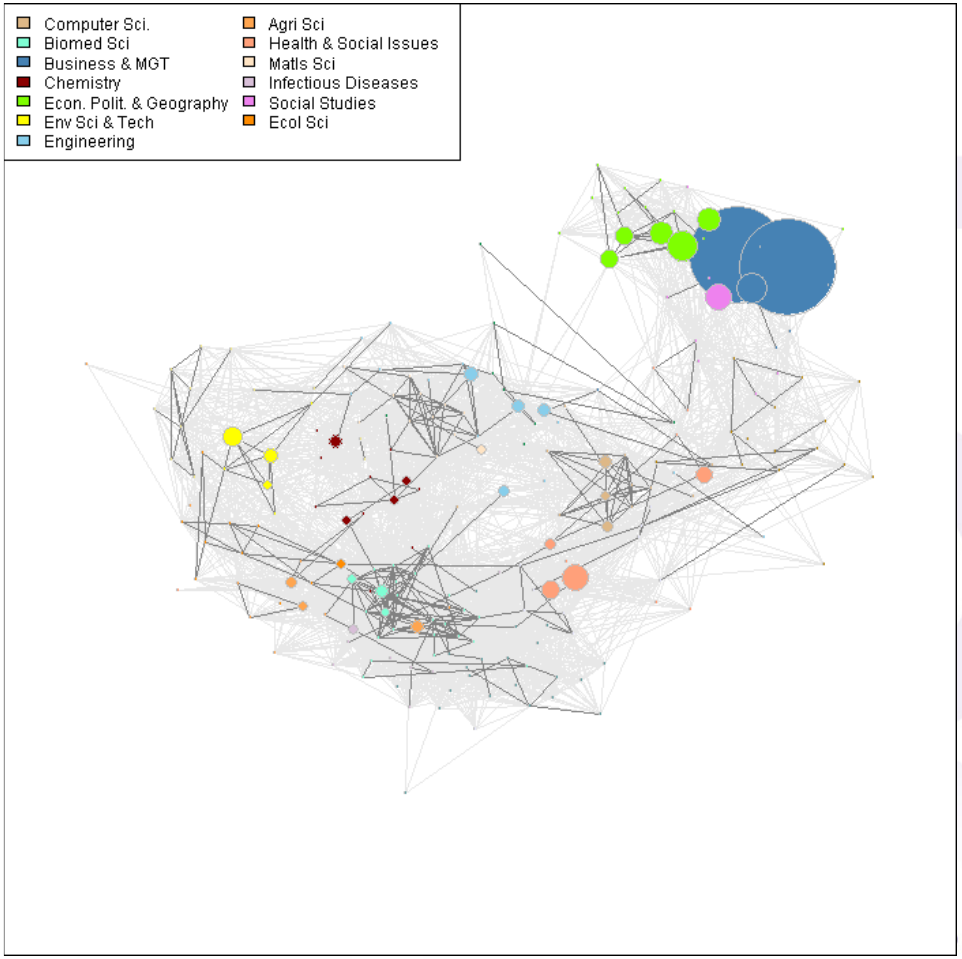
Actual ranking experiments

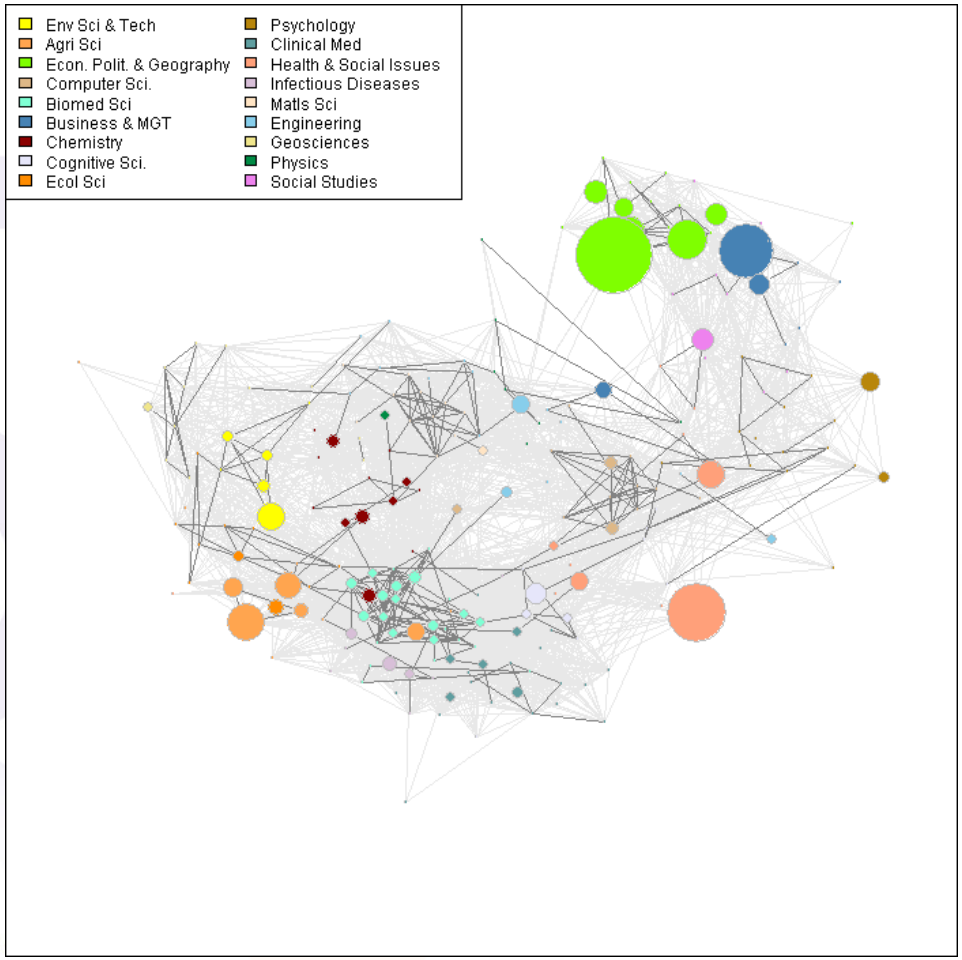


Actual ranking experiments



(1) Types of mobility: thematic mobility





(1)Types of mobility: thematic mobility

Apply dynamic IDR measures for a time series of individual scimaps

Table 1 *Typology of the Stirling index in measuring research diversity*

	Formula (versions of the generalized Stirling index)	d_{ij}	Underlying science map (level of aggregation)	Measuring diversity of…
1	$\sum_{ij(i\neq j)} d_{ij} p_i p_j$	$1 - s_{ij}$, where sij=cos(i,j)	Similarity network of (1) journals (2) ISI Subject Categories (based on the cited and citing dimension) Rafols, Meyer, Porter, Leydesdorff	(1) journals, (2) work of researchers, (3) output of organizations
2	$\sum_{ij(i\neq j)} d_{ij}$	g_{ij} shortest path from i to j (# edges)	Similarity network of papers (based on bibliographic coupling) Rafols, Meyer	particular research area

1. Pool of proposed indicators

Monitoring and comparing the differential career of concepts in different communities (scientific vs. non-scientific)

Primary example: $\Delta P(t) =_{df} \frac{P_S(t)}{P_P(t)}$

$\Delta C(t) =_{df} \frac{C_S(t)}{C_P(t)}$ $\Delta B(t) =_{df} \frac{B_S(t)}{B_P(t)}$

⇛

The experiments of UDE on contrasting „public” and „scientific” conceptual networks

1. Knowledge sharing „between science and society”

2. Knowledge sharing „within the scientific society”

SISOB-publication

András Schubert: *Measuring the similarity between the reference and citation distributions of journals*

- An indicator study for knowledge sharing
- Comparative in assessing the capabilities of an existing set
- Aim: to contrast the „community” of „incoming knowledge” with that of „disseminated knowledge”

⇛

- Shows the superiority of the „Jaccardized Czekanowsky Index” to other similarity measure

$$Cz_{A,B} = 1-\Sigma_i|q_i^A-q_i^B|/\Sigma_i(q_i^A+q_i^B) = 1-(1/2)\Sigma_i|q_i^A-q_i^B|,$$

2. Knowledge sharing „within the scientific society”

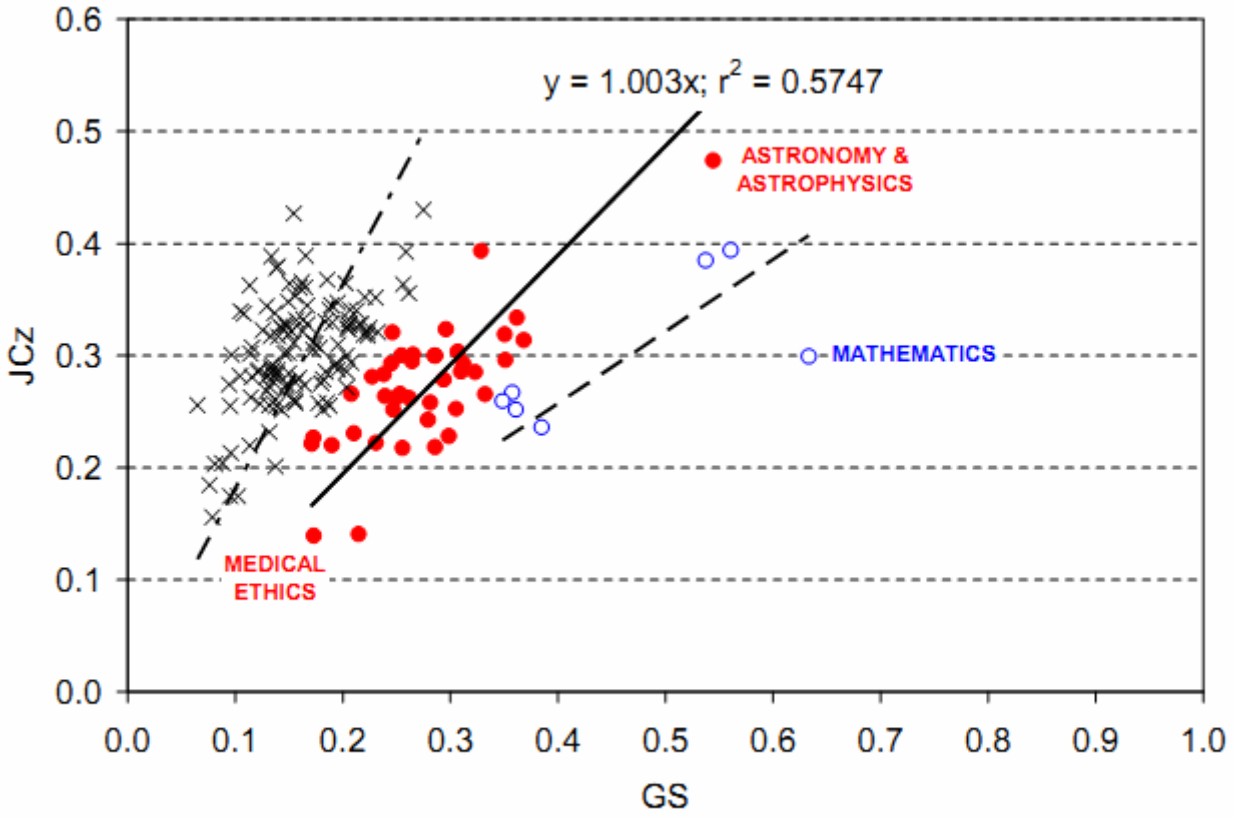


Figure 3 Regression plot of indices JCz and GS in three groups of subject categories

2. Knowledge sharing „within the scientific society”

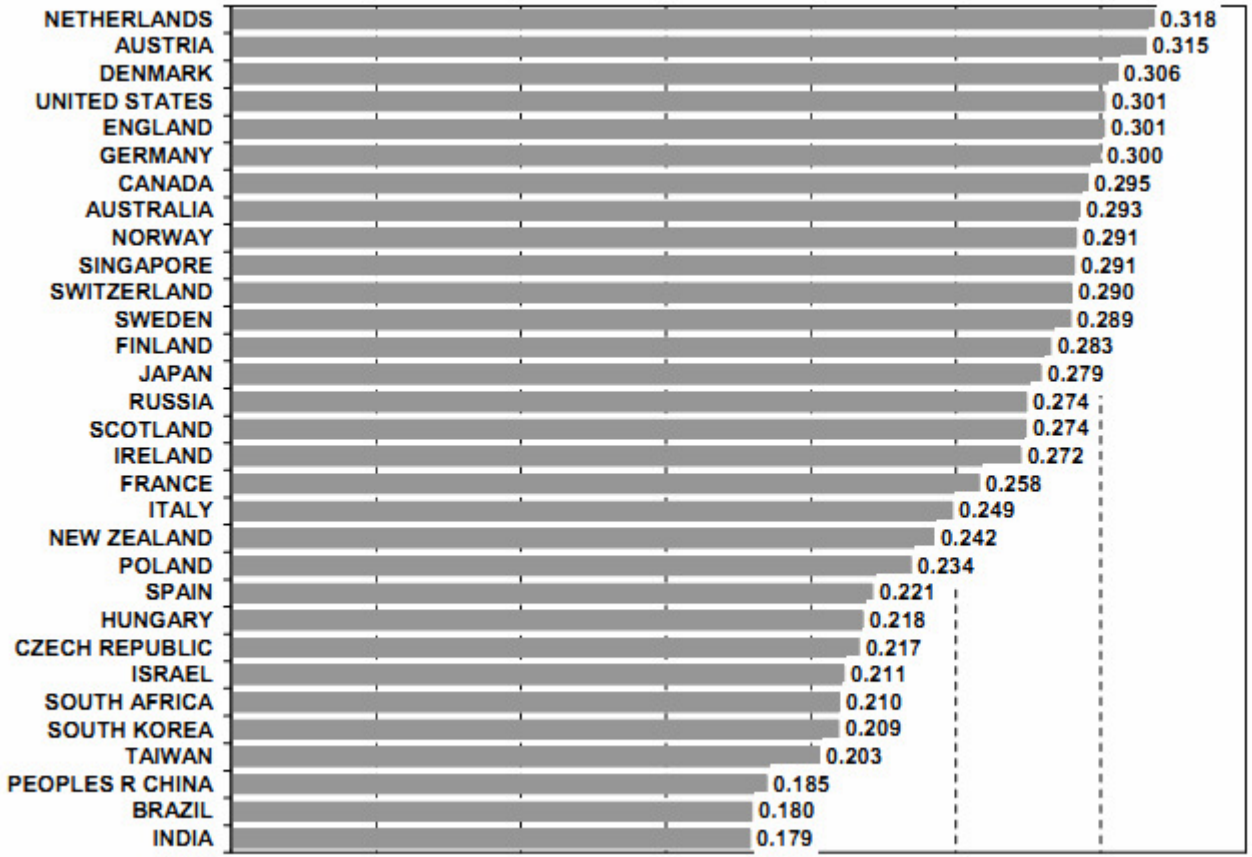


Figure 4 Country averages of the JCz index

2. Knowledge sharing „within the scientific society”

- Its a case of a „pure test” of a knowledge sharing indicator

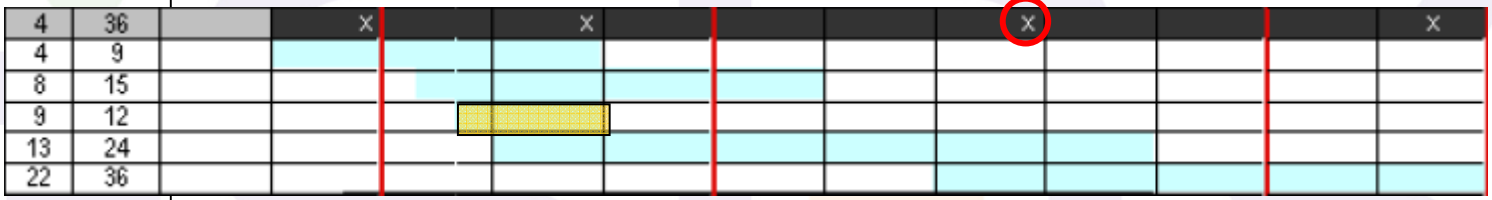
„It was shown that the indicator **characterizes the network properties** of individual journals and, in aggregated form, also that of subject categories or countries”.

„Evaluative aspects

The author feels the need to devote a specific paragraph to stress that the similarity index proposed in this paper **has no evaluative aspect**, whatsoever. Any attempt to find correlation between JCz and some impact factor-like indicator remained unsuccessful whether in the total sample or in selected subsamples (by subject category, country, journal type, etc.).”

T4.3 Contrast indicators with the information requirements

- Evaluate, using the SISOB database the quality of each indicator
- Refine the definition of the indicators

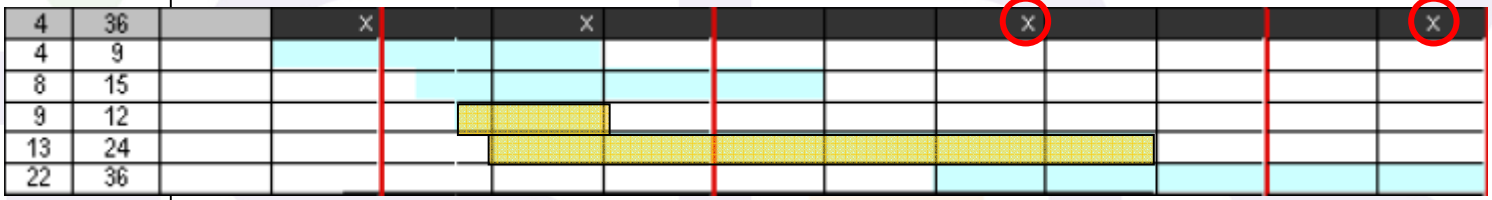


D4.3 Review of of the indicator quality test and refinements: Study about the quality of the indicators. (M24)

Review

T4.4 Implement the indicators within the generated modules in SISOB system

- Collect real data from the SISOB database, using control cases



D4.3 Review of of the indicator quality test and refinements: Study about the quality of the indicators. (M24)

D4.4 Report on the indicators operating within the SISOB modules. Comparative approach.
Report on the results obtained with the first prototype of SISOB with the basis to calcucate indicators.