

RISIS – Exploitations of EUPRO in the context of ERA

Research infrastructures for the assessment of science, technology and innovation policy

Thomas Scherngell, Michael Barber, Barbara Heller-Schuh

European Commission, DG RTD, 01/14-01/17

Grant Agreement no: 313082

RISIS Annual Week, 26-29 January 2015, Rome

Group Meeting 7

Agenda

- I. Overview and basic characteristics of EUPRO
- II. Some exploitations in the context of ERA
 - I. Background
 - II. Structure of FP networks
 - III. Integration of FP networks
 - IV. Communities in FP networks
- III. Outlook

What is EUPRO?

- EUPRO is a database comprising information on R&D collaborations within the European Framework Programmes (FPs),
- developed and maintained by AIT by standardising information collected from the EU CORDIS database,
- including systematic information on more than 60,000 collaborative research projects of (FP1-FP7) and more than 60,000 participating organisations,
- such as the geographical location and the organisation type etc.

Characteristics of the dataset

Programmes	Period	Number of projects	Number of participations	Number of organisations
FP1	1984 - 1987	3.283	7.929	1.982
FP2	1987 - 1991	3.878	19.068	4.406
FP3	1990 - 1994	5.528	31.378	7.051
FP4	1994 - 1998	14.562	67.858	18.635
FP5	1998 - 2002	16.034	78.728	22.105
FP6	2002 - 2006	9.709	72.012	19.479
FP7 (July 2012)	2007 - 2013	10.866	60.135	15.355
Total (July 2012)		63.860	337.108	78.336

- Standardisation procedure
 - Harmonisation of organisation names (heterogenous spellings, different languages)
 - Harmonisation and assignment of organisation types
 - Regionalisation of geographical data to NUTS3 level

EUPRO's contribution to study Europeanization I

- Initiatives to foster collaborative R&D in precompetitive research have become a key instrument of STI-policy at the regional, national and supranational levels
- Main examples in Europe: European Framework Programmes (FPs) on Research and Technological Development
- (Co-)funding of thousands of R&D projects supporting transnational collaboration and coordination in research and transnational mobility for training purposes; seven FPs since 1984

EUPRO's contribution to study Europeanization II

- Key information source for the empirical analysis of network structures and network processes within the European Research Area (ERA)
 - Coverage of a diverse set of countries and regions enables analysis of progress towards ERA from a detailed geographical perspective
 - Investigation of structural properties of collaboration networks provides information for appraising the effectiveness of EU policy in this area
 - Observing the involvement of individual organisations (number of participations, profile of partners, positioning in the collaboration network) supports the identification of key actors in the FPs

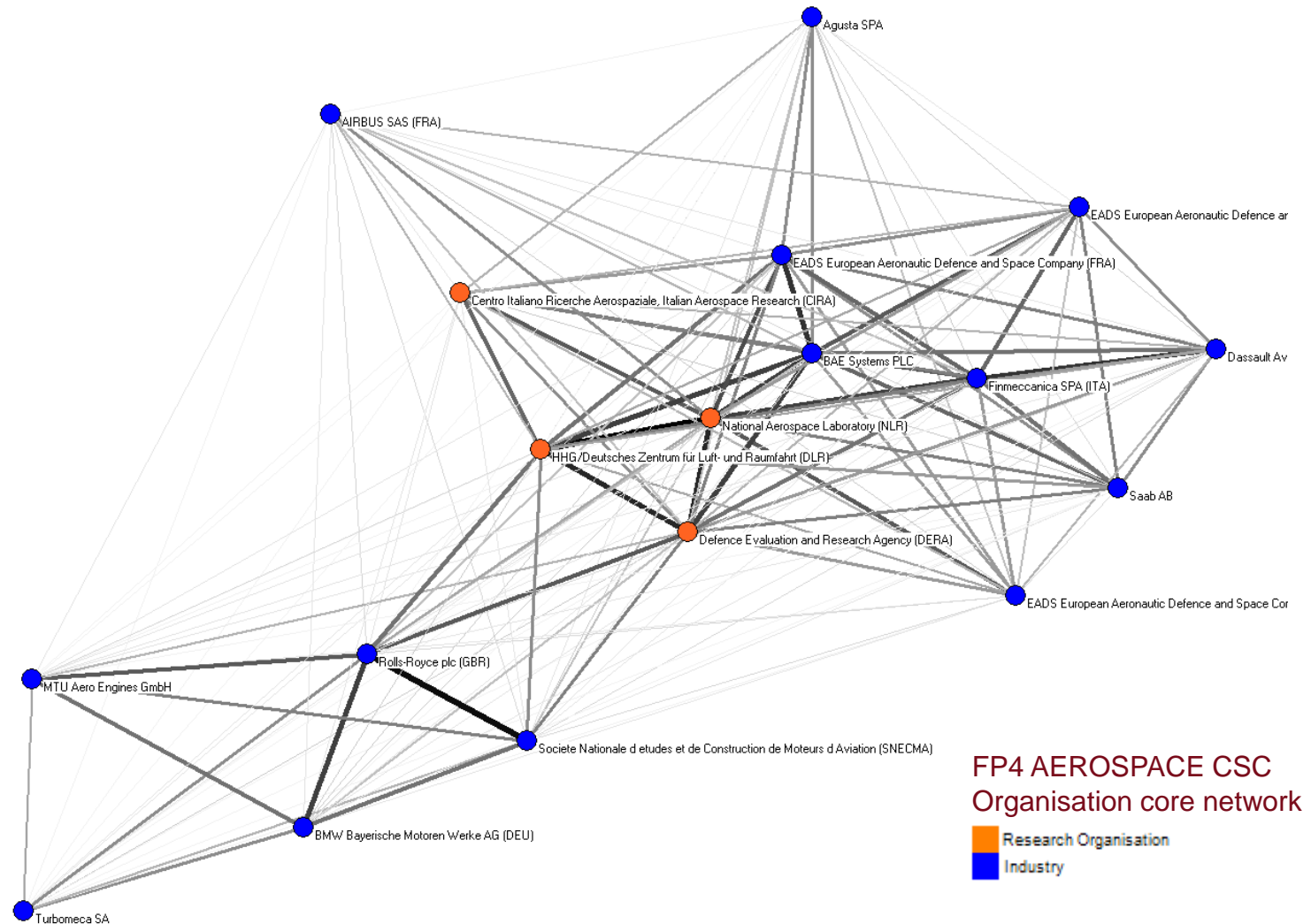
Example A: Structure of FP networks I

- Background
 - Analysis of collaborative networks promoted by the European Framework Programmes (1984-2006), joint study with JRC-IPTS
- Objectives
 - Exploit the richness of FP data through social network analysis to contribute to the progress of monitoring the move towards the European Research Area (ERA)
- Research questions
 - Does the density of collaborative organisational links increase over time?
 - Is it possible to identify a backbone of core research organisations in ERA?
 - Who are the key players in the FPs and where are they located within the FP networks?

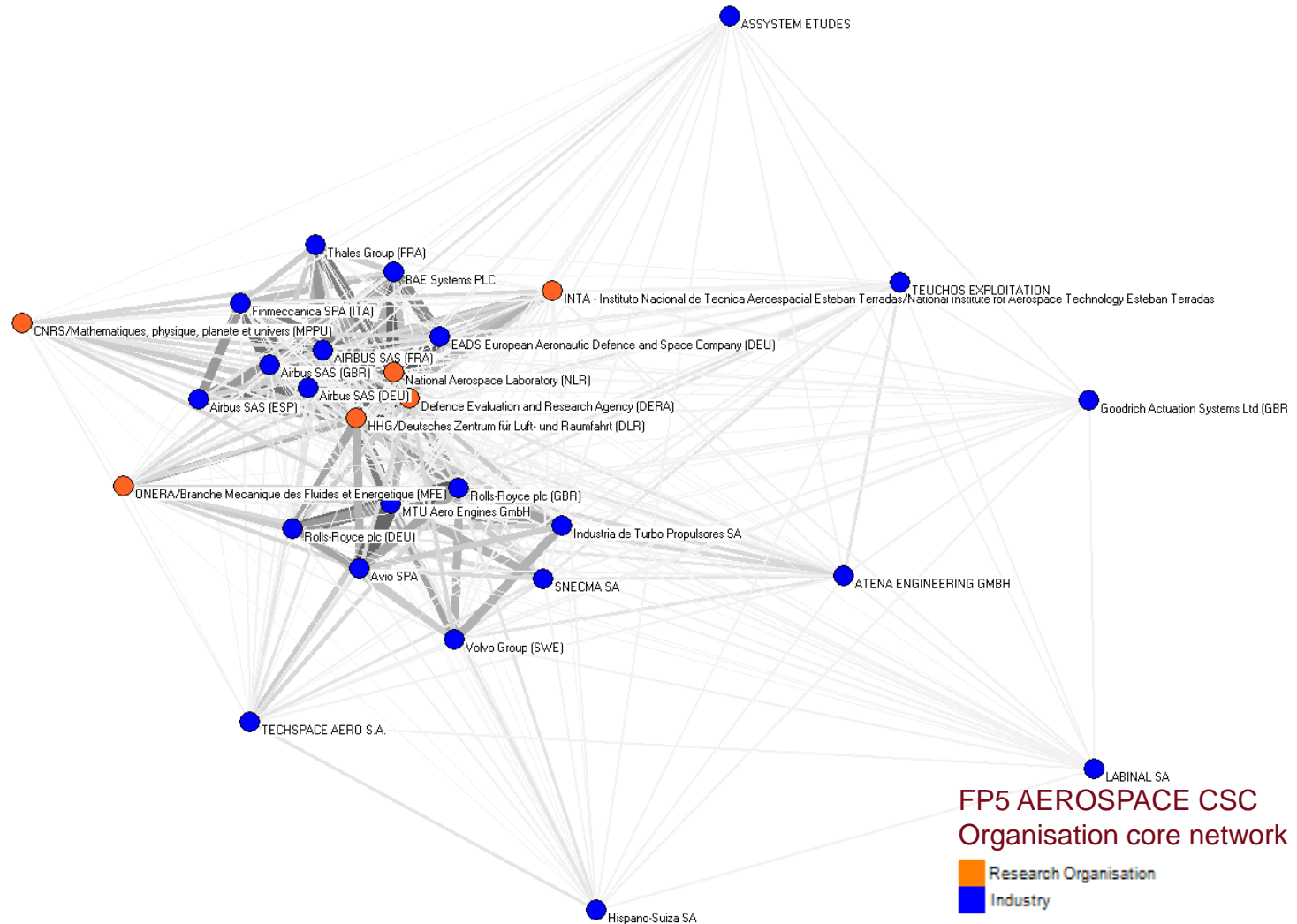
Example A: Structure of FP networks II

Graph Characteristic	FP1	FP2	FP3	FP4	FP5	FP6
No. of vertices N	2,116	5,758	9,035	21,599	25,840	17,632
No. of edges M	9,489	62,194	108,868	238,585	385,740	392,879
No. of components	53	45	123	364	630	26
N for largest component	1,969	5,631	8,669	20,753	24,364	17,542
Share of total (%)	93.05	97.79	95.95	96.08	94.29	99.49
M for largest component	9,327	62,044	108,388	237,632	384,316	392,705
Share of total (%)	98.29	99.76	99.56	99.60	99.63	99.96
N for 2nd largest component	8	6	9	10	12	9
M for 2nd largest component	44	30	72	90	132	72
Clustering coefficient	0.65	0.74	0.74	0.78	0.76	0.80
Diameter of largest component	9	7	8	11	10	7
l largest component	3.62	3.21	3.27	3.45	3.30	3.03
Mean degree	9.0	21.6	24.1	22.1	29.9	44.6
Fraction of N above the mean (%)	29.4	28.0	23.6	22.4	23.5	26.1
Mean vertex size	3.0	3.1	3.3	3.0	2.8	2.7
Standard deviation	5.0	6.1	7.7	7.9	6.8	5.4

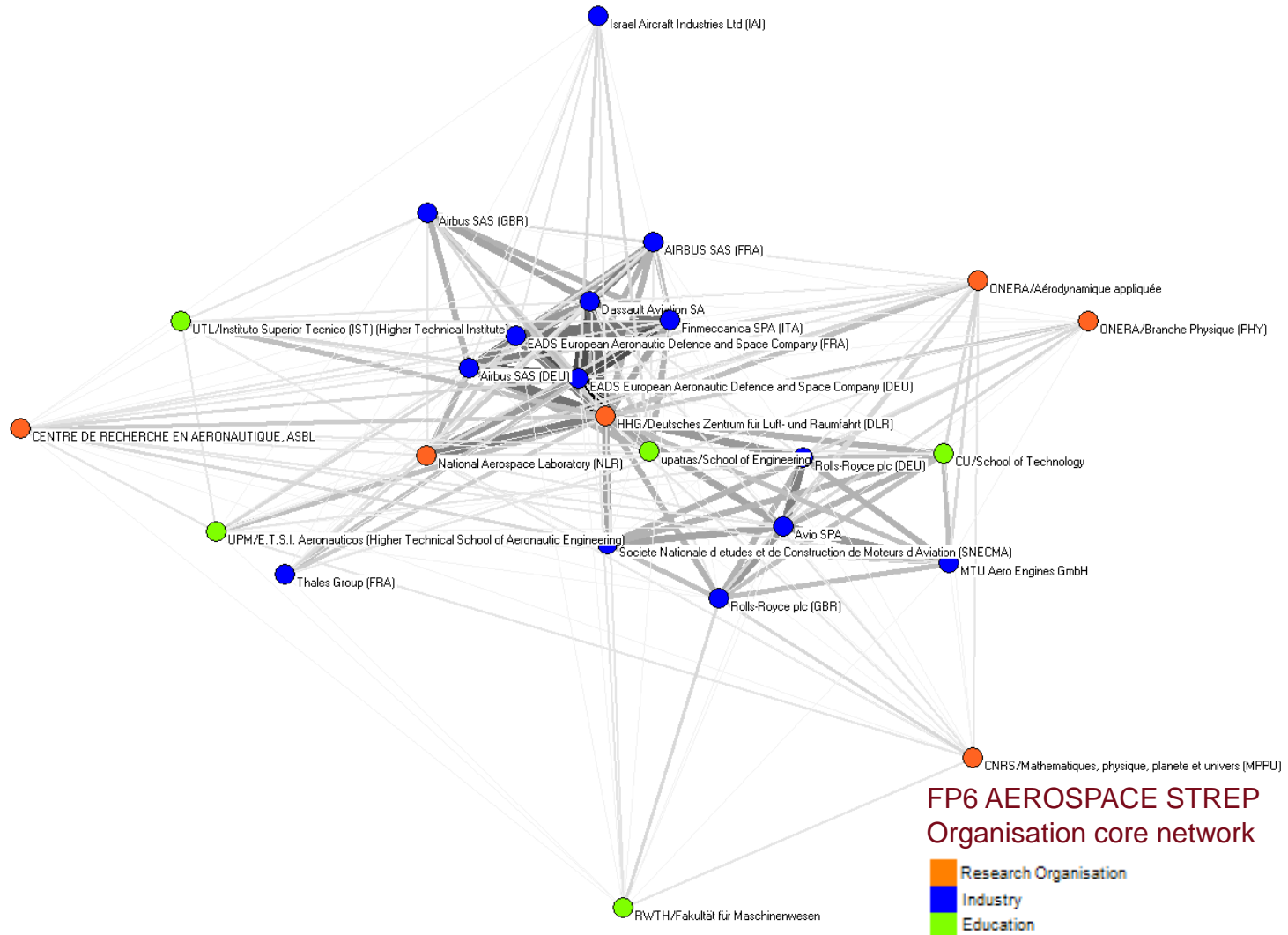
Example A: Structure of FP networks III



Example A: Structure of FP networks III



Example A: Structure of FP networks III

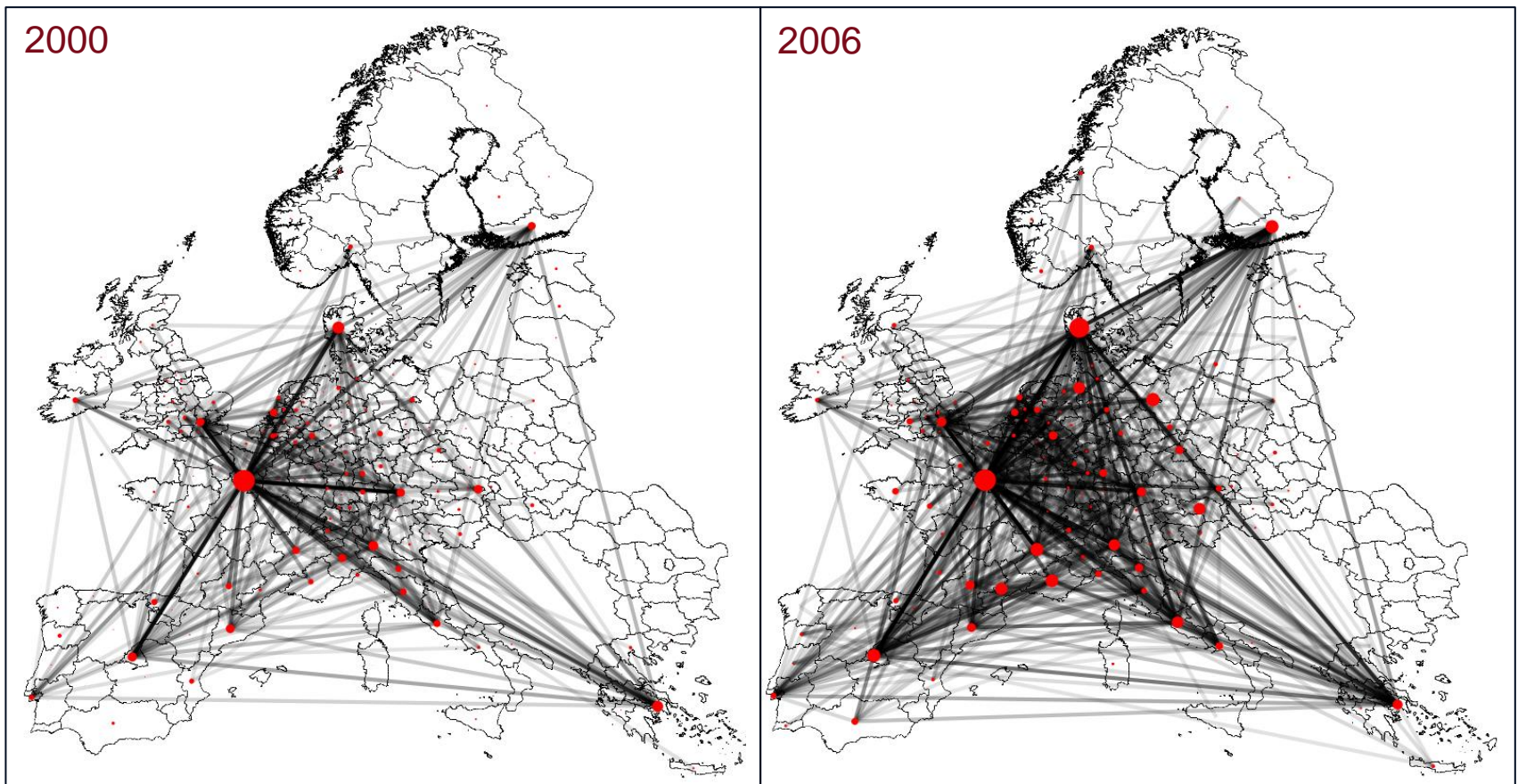


Example B: Integration of FP networks I

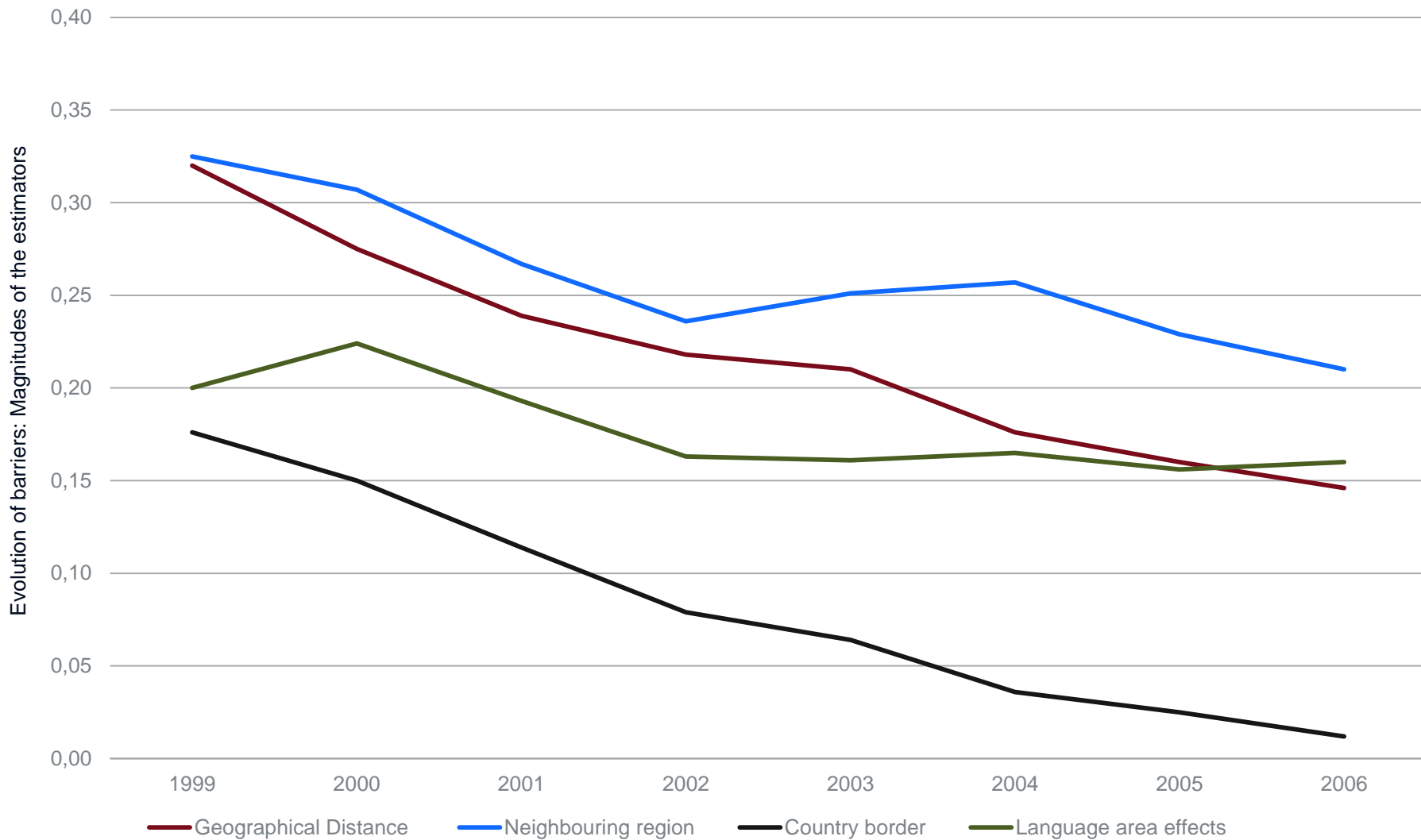
- Focus on integration of FP networks in terms of their spatial dimension
 - Identifying evolution of spatial patterns of FP networks, and
 - the evolution of separation effects affecting the probability of cross-region R&D collaborations in the European network of R&D cooperation,
 - Separation effects involve geographical, technological, cultural and institutional barriers
- Methodological and technical approach
 - Geographical Information (GI) tools to visualise the development of the networks in space in time
 - Spatial interaction models to identify the evolution of negative effects from spatial and non-spatial barriers to integration

Example B: Integration of FP networks II

The space-time evolution of FP networks across European NUTS-2 regions; node size corresponds to regional degree centrality, line transparency with the number of joint projects between two regions



Example B: Integration of FP networks III



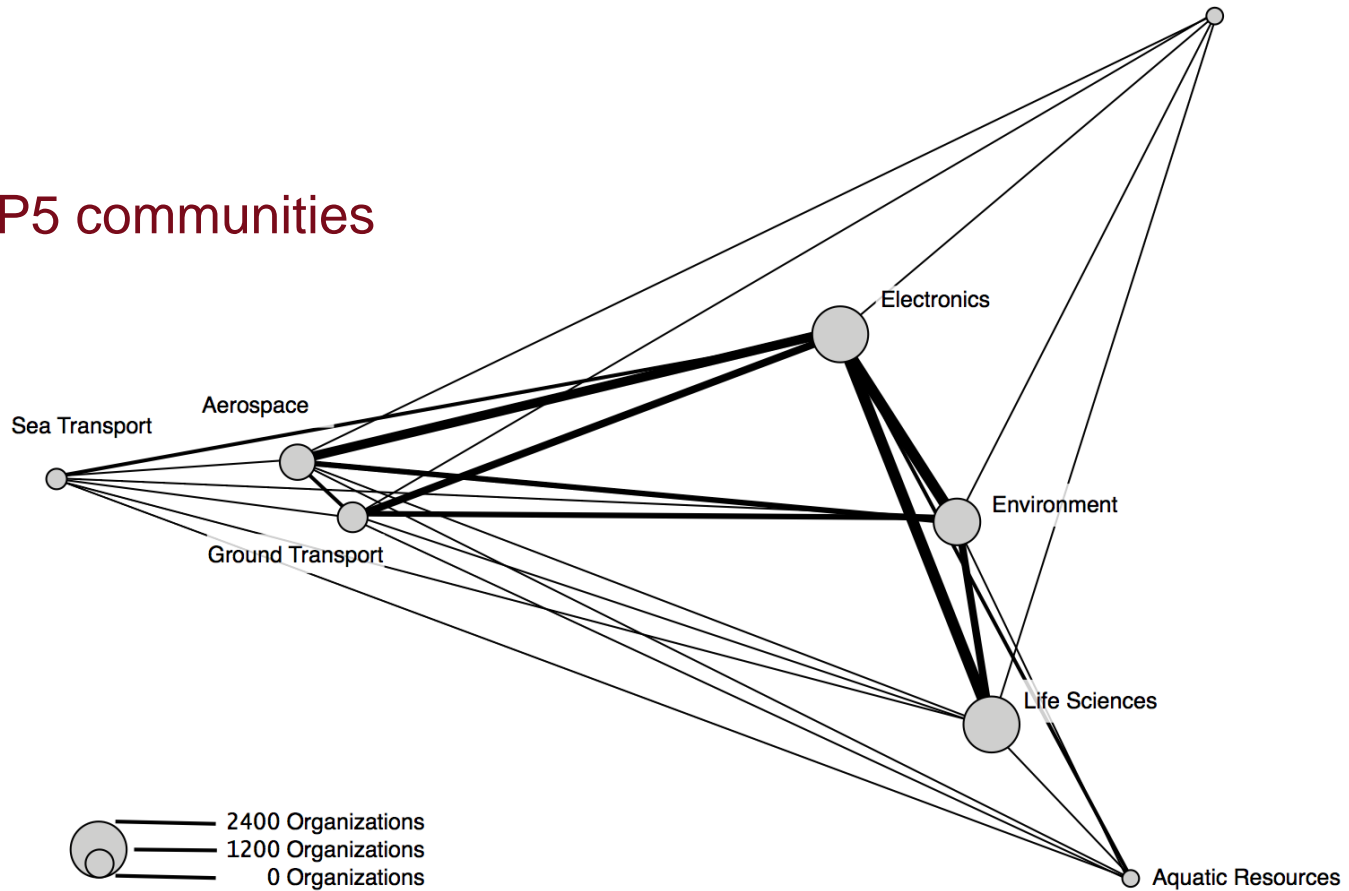
Note: Estimates received from a Negative Binomial spatial interaction model with Eigenvector spatial filtering

Example C: Communities in FP networks I

- Identify important communities in the FP networks
 - Communities are subnetworks with strong internal connections, but relatively weak connections to the rest of the network
 - Network communities often have distinct character, e.g., thematic or spatial properties
- Methodological and technical approach
 - Many ways to formalize idea of community.
 - Popular approaches based on optimizing a measure of community quality (e.g., modularity) with a heuristic search (e.g., greedy agglomerative clustering, label propagation algorithm, etc.).
 - Results provide a useful starting point for more detailed characterisations of the organisations and FP projects making up the communities.

Example C: Communities in FP networks II

Main FP5 communities



...plus 3474 smaller communities, many corresponding to single projects

Main references

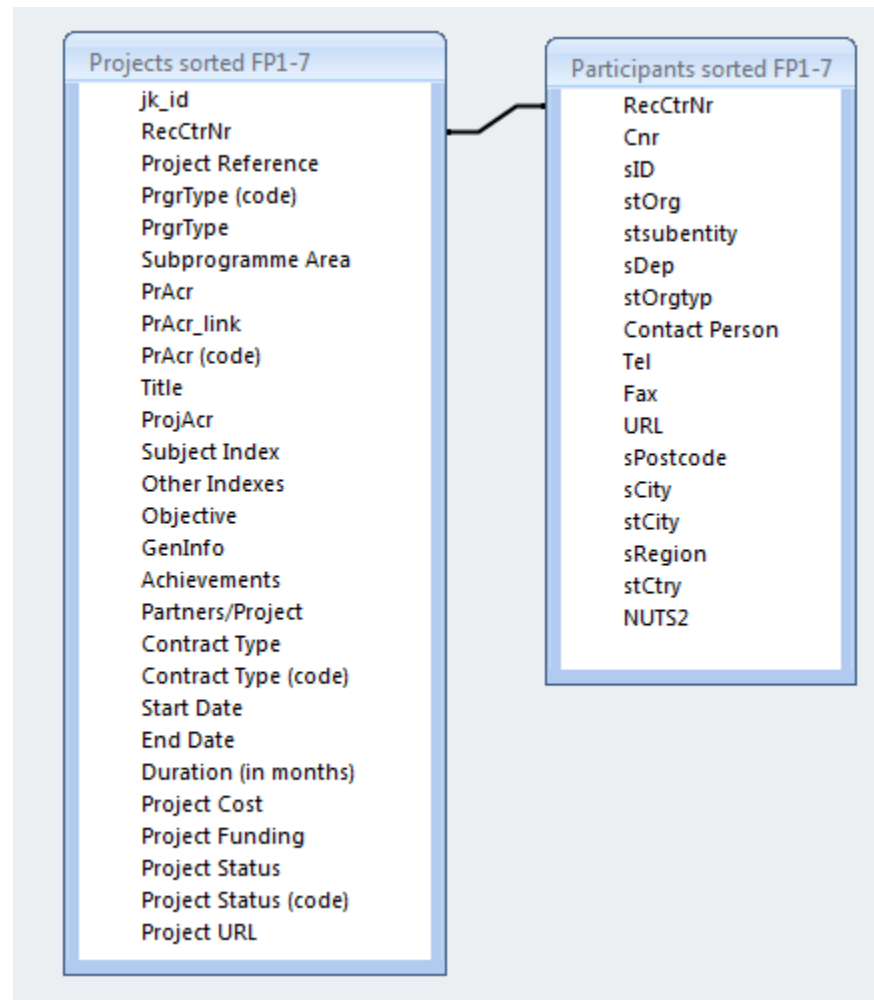
- Barber, M.J. and Scherngell, T. (2013): Is the European R&D network homogenous? Distinguishing relevant network communities using graph theoretic and spatial interaction modeling approaches, *Regional Studies* 47, 1283-1298
- Heller-Schuh, B., Barber M.J., Henriques, L., Paier, M., Pontikakis, D., Scherngell, T., Veltri, G.A. and Weber, M. (2011): *Analysis of Networks in European Framework Programmes (1984-2006)*, Sevilla, European Commission [ISBN: 9789279197147]
- Scherngell, T. and Lata, R. (2013): Towards an integrated European Research Area? Findings from Eigenvector spatially filtered spatial interaction models using European Framework Programme data, *Papers in Regional Science* 92, 555-577

Outlook

- Constant update of EUPRO in order to ensure ongoing monitoring of ERA dynamics in publicly funded R&D by the EU
- Potential extensions of EUPRO for evaluation purposes (e.g. integration of NABS categories)
- Extend richness of analyses on ERA dynamics by combining EUPRO with other RISIS datasets
 - Participation determinants of specific organisations and/or group of organisations
 - Impacts of FP networks (e.g. publications, patents, etc.), e.g. combinations of EUPRO with ETER, VICO, CIB, etc.

Appendix

EUPRO database structure



EUPRO project variables

RecCtrNr (data type: number)

- unique identifier (record control number) for each project in the database, identical with unique identifier of all projects in the CORDIS projects database

Project Reference (data type: number)

- project index, probably for internal use in the European Commission, structure and usage not clear

Prgr Type (code) (data type: number)

- code (1-7) for the names of the specific framework programme types FP1 to FP7

Subprogramme Area (data type: memo)

- full name of the subprogramme areas in each of the framework programmes

PrAcr (data type: text)

- subprogramme acronyms of the programmes (e.g., FP6-FOOD or FP6-INCO)

Title (data type: text)

- title of the project (full length)

ProjAcr (data type: text)

- (not-unique) project acronym or abbreviation of the project title

EUPRO project variables

Subject Index (data type: text)

- one or more of 45 standardized keywords;

Other Indexes (data type: memo)

- additional keywords to characterize project contents, freely chosen by the project team

Objective, GenInfo (general information), Achievements (data type: memo)

- distinction of these fields not clearly defined, in practice all of these fields (may) contain objectives and/or an abstract of the project.

Partners/Project (data type: number)

- number of participants in each project; additionally generated information, calculated with the help of table “participants sorted”

Contract type (data type: text)

- different types of contracts which regulate size, financing and funding of the research projects

Start Date, End Date (data type: date)

- day, month and year of project start and ending

Duration (in months) (data type: number)

- duration of the project in months

EUPRO project variables

Project Cost (data type: number)

- official project costs

Project Funding (data type: funding)

- financing contribution of the European Union; since not all projects are financed completely in the Additional Costs Model, figures in “Project Funding” are equal or smaller than figures in “Cost”.

Project Status (data type: text)

- status of the projects (accepted, completed, execution) found in the CORDIS projects database in January, 2012

Project URL (data type: text)

- official website of the project

EUPRO participants variables

RecCtrNr (data type: number)

- unique identifier (record control number) for each project in the database, identical with unique identifier of all projects in the CORDIS projects database; corresponds with the entries in the field RecCtrNr in the table “projects sorted”

Cnr (data type: number)

- unique identifier (control number) assigned internally by systems research, as project participants are not uniquely indexed in the CORDIS projects database; all project-relevant information is indicated with “1”, prime contractor with “2”, and remaining participants with “3”, “4”, etc.

sID (data type: number)

- internal number of standardised project participants (ambiguous number, assigned only for standardisation procedures!)

stOrg (data type: text)

- standardised organisation name

Contact Person (data type: text)

- name of contact person, without name affixes like titles, etc.

sEmail (data type: text)

- electronic mail address for the contact person

sAddress (data type: text)

EUPRO participants variables

sPostcode (data type: text)

sCity (data type: text)

sReg (data type: text)

- region according to the CORDIS projects database

stCtry (data type: text)

- standardised country; country abbreviations are given as ISO Alpha-3 country codes

stOrgtype (data type: text)

- standardised organisation type; available types are: IND (industry), EDU (universities and other educational institutions), ROR (publicly funded research organisations), GOV (governmental institutions), CON (consultants), NCL (non-commercial and non-profit organisations), and OTH (rest group)

NUTS1-NUTS3 (data type: text)

- region code according to the NUTS classification scheme

Project References I

- Urban Research in den EU-Rahmenprogrammen (5.-7. RP),
JPI Urban Europe, 10/14-03/14
- F&E Transport - Leistungsfähigkeit und Entwicklungspfade von Forschung und Entwicklung im österreichischen Transportsektor im europäischen Kontext (Horizon 2020),
Austrian Federal Ministry for Transport, Innovation and Technology, 01/14-12/14
- Data preparation for the identification of R&D networks related to European water research,
Rathenau Institute, 02/13-03/13
- Data preparation and scientific advisory for the geography of mobile phone R&D networks,
University of Toulouse, 06/11-08/11
- Analyse der F&E-Netzwerke österreichischer Universitäten im 5., 6. und 7. Rahmenprogramm,
Austrian Federal Ministry for Science and Research, 11/10-06/11
- Analyse der F&E-Netzwerke österreichischer Universitätsinstitute im 7. Rahmenprogramm,
Austrian Federal Ministry for Science and Research, 11/10-06/11
- Vergleichende Untersuchung der Publikations- und F&E-Netzwerke österreichischer
Universitätsinstitute,
Austrian Federal Ministry for Science and Research, 11/10-06/11
- Beteiligung, Positionierung und Vernetzung österr. Akteure im 7. RP,
Austrian Research Promotion Agency (FFG), 07/10-11/10
- The spatial and temporal evolution of R&D Networks,
Austrian Science Fund (FWF), 08/09-01/12

Project References II

- Network Analysis Study on participations in Framework Programmes, European Commission, JRC, 09/08-03/09
- NEMO - Network Models, Governance and R&D Collaboration Networks, European Commission, DG Research, 09/06-11/09
- CEERails - Integration zentral- und osteuropäischer Bahntechnologiestrategien, Austrian Federal Ministry for Transport, Innovation and Technology/Austrian Research Promotion Agency (FFG), 11/07-01/09
- RailNet - Analyse und Konzeption von kollaborativen Netzwerken in der Bahnforschung, Austrian Federal Ministry for Transport, Innovation and Technology/Austrian Research Promotion Agency (FFG), 01/06-07/06
- Analyse der Forschungsfelder in den Rahmenprogrammen der Europäischen Union als Grundlage für die Identifizierung von Emerging Technologies, DaimlerChrysler AG, 02/06-06/06
- Systemforschung im Urbanen Raum - Teilprojekt Wissensnetzwerke, ARGE Systemforschung im urbanen Raum, 01/05-12/06