NEMO Working Paper #19

The dynamics of actor configurations in public R&D programmes

Barbara Heller-Schuh
Michael Barber

Austrian Institute of Technology GmbH

Supported by the EU FP6-NEST-Adventure Programme
Contract n° 028875 (NEST)
(Milestone 4.5)

This version June 2009

Contents

1 Introduction

Since 1984, the European Union has established seven Framework Programmes for research and technological development, in which funding activities are seen as a key concept to link Europe's research excellence in transnational R&D networks. The Framework Programmes aim at focussing and integrating the research activities within the European Union and have significantly advanced international research collaboration in Europe (Luukkonen 2001).

The main source of data for the NEMO project is the sysres EUPRO database, which comprises information on all projects funded in FP1-FP6 (current status). The database was constructed from the entire raw data publicly available on the CORDIS website of the European Commission (www.cordis.lu). It contains detailed information on funded projects and project participants of EU FPs. For the funded projects, CORDIS lists information on project objectives and achievements, project costs, project funding, temporal location and contract type. It also provides information on the project participants including the full name, the geographical location, the participating department, the contact person and the type of the organisation.

In the absence of more detailed information on how organisations collaborate in the R&D projects (i.e., the intra-project structures), we construct networks from the available data (Roediger-Schluga and Barber 2006). To this end, we start with the affiliation network of collaborative research projects and participating organisations. An affiliation network represents a special type of two-mode network with one set of actors, which belong to a set of events (Wassermann and Faust 1994). Typical examples of affiliation networks are directors serving on company boards, researchers co-authoring scientific publications, people participating in some joint event, etc.

In our empirical analyses, we focus on organisation and project networks separately, rather than analysing them as bipartite networks. The networks were constructed by projecting the bipartite graph onto the set of organisations and the set of projects, producing the o-graphs and p-graphs, respectively. For the o-graphs we assume that organisations are connected if they participate in the same project. Thus, we draw an edge between e.g. Alcatel and ABB if and only if these organisations participate in the same R&D project. Edges can be weighted, if there are multiple collaborations between the same organisations. Thus, if Shell and the University of Cambridge participate jointly in two projects, the corresponding edge has a weight of two and so on. In the same manner we proceed for the p-graphs: two projects are connected if the same organisation participates in both projects.

In constructing the networks, we thus assume for the o-graph each project to be a fully connected subgraph of organisations and for each p-graph to be a fully connected subgraph of projects. This is an idealised graph type that, although not fully representative, is a reasonable approximation to the actual intra-project structure of all but very large projects. Since the vast majority of projects in our data set have fewer than 15 participants, our construction rule is considerably more accurate than assuming the other idealised type of a star structure, in which each participant is only connected to the project coordinator as central vertex.

2 Networks related to thematic priorities and instruments

For a more detailed analysis of different instruments within different thematic priorities in FP4 to FP6, we identified 20 sub-networks in the fields of Aerospace, Energy & Environment, ICT and Life Sciences. These fields represent four of the thematic areas in FP6 (Aeronautics and space (AERO); Sustainable development, global change and ecosystems (ENV); Information Society Technologies (ICT); Life sciences, genomics and biotechnology for health (LIFESCI)) for which previous programmes in FP4 and FP5 could be easily identified.

Each of these sub-networks comprises for a given thematic priority only those projects (and their participants) of a specific contract type. We selected for FP4 and FP5 the contract type Costs Shared Contracts (CSC), which is in use since the first Framework Programme. CSC are considered as collaborative RTD projects with the aim of obtaining new knowledge, demonstration projects with the aim of showing the viability of new technologies, and support measures for access to research infrastructures. They require a minimum of two partners (EC 2003).

For FP6 we choose Specific Targeted Research Projects (STREP) which represent the former Shared-Cost Contracts and comprise objective driven research of limited scope, focused on a single issue. A minimum of three partners is required. SMEs usually state a clear preference for this instrument. (EC 2004). Additionally, we selected for FP6 two other contract types: Integrated Projects (IP) and Networks of Excellence (NoE). Integrated Projects (IP) are a new instrument in FP6 devoted to basic as well as applied objective-driven research with a “programme approach.” IPs are expected to assemble the necessary critical mass of activities, expertise and resources to achieve ambitious objectives. In practice, organisations with skills in management, dissemination and knowledge transfer, as well as potential users and other stakeholders, are recommmended, as well as a project size of 10-20 participants (EC 2004). Networks of Excellence (NoE) are also a new instrument in FP6 and are designed to strengthen scientific and technological excellence on a focused research topic. NoE are designed to integrate the critical mass of resources and expertise needed to provide European leadership and be a world force in that topic. NoEs are therefore an instrument aimed at tackling fragmentation of existing research capacities and aim at gathering research centres, universities, research and technology organisations, and to a lesser extent enterprises. 6-12 participants are recommended (EC 2004).

Table 1 gives an overview of the identified 20 sub-networks. For each thematic priority the names of the relevant sub-programmes across FPs are presented as well as the number of projects and participating organisations (on the subentity level) in each of these sub-programmes. Based on the information about projects and their participants we constructed affiliation networks and derived from that organisation and project networks as described before.

Table 1: Thematic sub-networks by instruments – numbers of projects and organisations

Sub-network	FP	Sub-programme	Instru-ment	Projects	Parti-cipants
AERO_FP4_CSC	FP4	BRITE/EURAM 3 (Aeronautics)	CSC	89	321
AERO_FP5_CSC	FP5	GROWTH (Aeronautics)	CSC	131	801
AERO_FP6_IP	FP6	FP6-AEROSPACE	IP	27	595
AERO_FP6_NoE	FP6	FP6-AEROSPACE	NoE	3	43
AERO_FP6_STREP	FP6	FP6-AEROSPACE	STREP	96	620
ENV_FP4_CSC	FP4	ENV 2C; NNE-JOULE C*	CSC	1,111	2,599
ENV_FP5_CSC	FP5	EESD	CSC	1,017	3,947
ENV_FP6_IP	FP6	FP6-SUSTDEV	IP	97	1,953
ENV_FP6_NoE	FP6	FP6-SUSTDEV	NoE	16	449
ENV_FP6_STREP	FP6	FP6-SUSTDEV	STREP	167	1,373
ICT_FP4_CSC	FP4	ESPRIT 4	CSC	940	2,622
ICT_FP5_CSC	FP5	IST	CSC	1554	5,462
ICT_FP6_IP	FP6	FP6-IST	IP	176	2,119
ICT_FP6_NoE	FP6	FP6-IST	NoE	49	914
ICT_FP6_STREP	FP6	FP6-IST	STREP	483	2,393
LIFESCI_FP4_CSC	FP4	BIOMED 2; BIOTECH 2	CSC	764	1,473
LIFESCI_FP5_CSC	FP5	LIFE QUALITY**	CSC	794	2, 335
LIFESCI_FP6_IP	FP6	FP6-LIFESCIHEALTH	IP	87	917
LIFESCI_FP6_NoE	FP6	FP6-LIFESCIHEALTH	NoE	29	568
LIFESCI_FP6_STREP	FP6	FP6-LIFESCIHEALTH	STREP	140	746

Note *: NNE-JOULE C comprises 460 collaborative projects, but since in one project the names of the participating organisations are not available, this project is not included in the calculation of sub-networks; ** in LIFE QUALITY projects in Key actions *Food, nutrition and health* and *Sustainable agriculture, fisheries and forestry* are not included

Furthermore, the investigation of the spatial structure of FP networks is a crucial part of various empirical analyses of NEMO, as requested in the project proposal (see WP4/Task 4.3). In order to be able to investigate the spatial structure of FP networks, we need an appropriate representation of geographical space. In NEMO, we use regions as units of observation for the spatial analysis. The European coverage is achieved by disaggregating our sysres EUPRO FP network data (see Section 2) into $i, j = 1, \ldots, n = 255$ NUTS-2 regions[1] (NUTS revision 2003) of the 25 pre-2007 EU member-states, as well as Norway and Switzerland. We use a concordance scheme between postal codes and NUTS regions provided by Eurostat to trace the specific NUTS-2 region of an organisation. By this extension, the sysres EUPRO database represents an extremely valuable resource for any kind of empirical analysis on the geography of knowledge creation and diffusion across Europe.

By assigning each organization to a spatial unit, we can, for instance, analyze the spatial distribution of participating organization, disaggregated by organization type. However, in

---

[1] NUTS is an acronym of the French for the “nomenclature of territorial units for statistics", which is a hierarchical system of regions used by the statistical office of the European Community for the production of regional statistics. At the top of the hierarchy are NUTS-0 regions (countries) below which are NUTS-1 regions and then NUTS-2 regions. Although varying considerably in size, NUTS-2 regions are widely viewed as the most appropriate unit for modelling and analysis purposes (see, for example, Fingleton 2001).

order to investigate the spatial configuration of EU FP collaboration from a network perspective, we need to focus on the spatial properties of the network links itself. Thus, we construct a region-by-region collaboration matrix that we label ***P*** by aggregating the number of individual collaborative activities to the regional level which leads to the observed number of R&D collaborations p_{ij} between two regions i and j. For instance, for a project with three participating organisations in three different regions, say regions a, b, and c, we count three links: from region a to region b, from b to c and from a to c[2]. When all three participants are located in one region we count three intraregional links. Note that we have excluded self loops to eliminate artificial self collaborations. The resulting regional collaboration matrix ***P*** then contains the collaboration intensities between all (i, j)-region pairs, given the $i = 1, \ldots, n = 255$ regions in the rows and the $j = 1, \ldots, n = 255$ regions in the columns. The n-by-n matrix is symmetric by construction ($p_{ij} = p_{ji}$).

3 Structural properties and temporal evolution of networks

Network links in the organisation graphs of the empirical networks represent participations of organisations in joint research projects, and can thus be interpreted as channels for the diffusion of information or the potential for joint knowledge production. The global structural properties of the networks can thus give some indication of network performance.

Based on the available data, we can characterize network properties and their change over time (FP1 to FP6) for FP networks in general as well as for the thematic sub-networks. For example, the clustering coefficient within the FP networks reflects the density of connections between networked organisations and projects from FP1 to FP6. A rise over time suggests that the intensity of collaboration between organisations has increased, indicating that Europe has already been moving towards a more closely integrated European Research Area in the earlier Framework Programmes.

In light of the focus on structural features of FP networks, also the analysis of the spatial structure of such networks, as well as the analysis of the determinants of the observed spatial structures is of crucial importance. Analyses in this direction enhance our understanding of knowledge diffusion processes on FP network from a spatial perspective. It provides important new insights not only in a scientific context, but also in a European policy context (see Scherngell and Barber 2009).

3.1 Global network characteristics

We identified the basic structural properties for FP1 to FP6 networks and for the selected 20 sub-networks based on different instruments in four thematic priorities, using indicators developed in statistical physics and social network analysis. For this purpose we analysed

---

[2] We refer to this counting method as full counting. Another counting method would be fractional counting by dividing each link in a project by the total number of links in a project. The full counting procedure used in the current study overestimates the impact of large projects, while the impact of large projects would be underestimated using the fractional counting method. We prefer the full counting method since full rather than fractional counting does justice to the true integer nature of R&D collaborations and is applicable in the context of a Poisson model specification.

both, the organisation projections (o-graphs) and the project projections (p-graphs) of the affiliation network. Several groups of parameters were calculated for each FP and for each sub-network. The descriptors introduced here are used in the following tables coherently.

Vertices and edges in the network

- Number of vertices *N*: in the organisation projection (o-graph) a vertex (in social network theory also referred to as a node) represents an organisation, whereas in the project projection (p-graph) a vertex represents a project
- Number of edges *M*: in the o-graph, an edge (in social network theory also referred to as a link) represents a participation in a joint project; in the p-graph, an edge represents an organisation participating in both projects

Measures of fragmentation of the network

- Number of components: Components are connected sub-networks. Thus, a higher number of components is associated with a higher fragmentation of the network
- *N* for largest component: number of vertices in the largest component
- Share of total *N* (%): fraction of the vertices in the largest component in the total number of vertices
- *M* for largest component: number of edges in the largest component
- Share of total *M* (%): share of the edges in the largest component in the total number of edges
- *N* for 2nd largest component: number of vertices in the second largest component
- *M* for 2nd largest component: number of edges in the second largest component

Other structural measures for the network

- Clustering coefficient: For a given vertex the clustering coefficient measures the local density of a network by indicating the extent to which its direct neighbours are also connected. The clustering coefficient of a network is the mean clustering coefficient of all vertices (Watts and Strogatz 1998).
- Diameter of largest component: The distance between two vertices is the shortest path between them. The diameter of a network is the longest distance between any two of its vertices. It can be interpreted in the context of information flow through the whole network.
- ℓ largest component: ℓ denotes the characteristic path length, i.e. the average distance between pairs of vertices; it can be interpreted in the context of information flow.
- Mean degree: The degree of a vertex denotes the number of its direct neighbours; for the o-graph this means the overall number of partners of an organisation, for the p-graph the overall number of linked projects
- Fraction of *N* above the mean (%): the share of vertices with degree higher than the mean degree; indicative of the skewness of the degree distribution
- Mean vertex size *P*: In the o-graph *P* denotes the mean number of projects of an organisation; in the p-graph *P* denotes the mean number of participating organisations in a project
- Standard deviation of *P*: a measure of the width of the distribution of *P*; indicative of the skewness of the distribution of vertex sizes

3.2 Characteristics of FP organisation networks

In Table 2, we give some basic properties of the organisation projections for FP1 to FP6. Note that both FP1 and FP6 are exceptional - FP1 has the poorest source data, while FP6 is not complete at the time of writing.

The increase in the number of vertices *N* in the organisation network shows that a growing number of organisations have participated in subsequent FPs (FP6 goes against this trend). Most of these are linked to each other. A giant component is present in all FP R&D collaboration networks. In each case, the great majority of vertices and essentially all edges are in the giant component. The existence of a giant component indicates that two arbitrary vertices are connected, either directly or indirectly through a path of connected vertices, with high probability, ensuring that information or objects can spread or diffuse in a network. This means that even in the absence of other, unobserved communication channels, information can spread in the observed networks.

As expected from prior works (Barber et al. 2006; Breschi and Cusmano 2004), all networks are of small-world type (Watts and Strogatz 1998): They exhibit a high clustering coefficient (a measure of local connectedness), a small characteristic path length and a diameter that scales at most logarithmically with the number of vertices. Networks with high clustering coefficients are called cliquish. In terms of what we presently know about knowledge creation and knowledge diffusion in social networks (Cowan 2006), this is a positive result. When path lengths are short, new knowledge can spread rapidly and widely through the population and thus fuel local knowledge creation. Dense local connections facilitate learning. Agents can only learn from each other if they know different things but are sufficiently similar to communicate. As in a barter economy, there must be a double coincidence of wants. This constraint is relaxed if agents can communicate through joint neighbours, of which there are many in cliquish networks. There is a slight increase in the clustering coefficient from FP1 to FP5. This suggests that integration between collaborating organisations has increased over time, indicating that Europe has already been moving towards a more closely integrated ERA in the earlier Framework Programs (on this, see Breschi and Cusmano, 2004). The increasing trend continues in FP6, but some caution is warranted: the marked change in FP6 project sizes would be expected to increase the clustering coefficient as well.

The mean degree in the R&D collaboration networks is roughly constant through FP5, with a sharp jump for FP6. We interpret this as evidence that organisations have a roughly constant capability to maintain connections to one another. However, the mean degree is not terribly informative as it does not divide the population into two roughly equal halves. Rather, only around a quarter of organisations and a third of projects have a degree higher than the mean, indicating a skewed degree distribution.

Table 2: Characteristics of the organisation projection of FP networks (o-graph)

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 *N* (%)	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 *M* (%)	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 *P*	3.0	3.1	3.3	3.0	2.8	2.7
Standard deviation of *P*	5.0	6.1	7.7	7.9	6.8	5.4

Note: For an explanation of the descriptors see page 4.

3.3 Characteristics of organisation sub-networks by instruments

The organisation sub-networks based on different instruments (CSC/STREP, IP and NoE) in four thematic priorities show quite similar properties (Table 3 to Table 5) compared to the FP organisation networks in general. In each sub-network a giant component consists of the majority of nodes (94-100%) and each of the sub-networks shows small world network characteristics: high clustering coefficient (between 0.65 and 0.80) and a small characteristic path length (~3). Again, the clustering coefficient indicates increased clustering from FP4 to FP6 in most of the thematic priorities and and decreasing mean vertex size *P* shows that organisations tend to participate in FP6 in a smaller number of projects than in previous projects.

Only the sub-networks in Aerospace differ in some respects and show, especially for FP4 and FP5, even highly intensified clustering: The number of participating organisations is compared to other thematic priorities smaller (number of vertices), but the mean number of partners in FP4 (27) and FP5 (59) as well as the clustering coefficient in each FP is significantly higher. Additionally, the network consists of one single component comprising all participants in Aerospace projects.

The differences between instruments correspond to different project sizes. Integrated Projects and Networks of Excellence are large projects with many participants; therefore the mean number of partners per organisation (mean degree) as well as the clustering coefficient in these networks is higher than in networks based on projects of contract type Cost Shared Contracts.

Based on these results it might be assumed that conditions and instruments of different FPs show more influence on collaboration structures than different technological regimes.

Table 3: Characteristics of the organisation projection of FP sub-networks by instruments (CSC and STREP)

Graph Characteristics	**AERO_4_CSC**	**AERO_5_CSC**	**AERO_6_STREP**	**ENV_4_CSC**	**ENV_5_CSC**	**ENV_6_STREP**	**ICT_4_CSC**	**ICT_5_CSC**	**ICT_6_STREP**	**LIFE SCI_4_CSC**	**LIFE SCI_5_CSC**	**LIFE SCI_6_STREP**
No. of vertices *N*	321	801	620	2.599	3.947	1.373	2.622	5.462	2.393	1.473	2.335	746
No. of edges *M*	4.354	23.463	5.993	20.353	42.555	10.651	12.035	40.299	15.952	13.407	23.243	4.685
No. of components	1	1	3	11	19	4	34	29	3	7	3	2
***N* for largest component**	321	801	607	2.561	3.855	1.349	2.489	5.304	2.376	1.458	2.311	743
Share of total (%)	100,00	100,00	97,90	98,54	97,67	98,25	94,93	97,11	99,29	98,98	98,97	99,60
***M* for largest component**	4.354	23.463	5.955	20.284	42.337	10.560	11.772	39.903	15.888	13.395	23.150	4.682
Share of total (%)	100,00	100,00	99,37	99,66	99,49	99,15	97,81	99,02	99,60	99,91	99,60	99,94
***N* for 2nd largest**	0	0	8	8	9	11	12	16	9	3	17	3
***M* for 2nd largest**	0	0	56	56	72	110	132	72	72	6	144	6
Clustering coefficient	0,85	0,89	0,87	0,77	0,81	0,90	0,82	0,82	0,84	0,73	0,76	0,86
Diameter of largest component	3	4	6	7	6	7	11	8	9	7	7	6
ℓ largest component	2,10	2,16	2,63	3,19	3,14	3,66	3,77	3,44	3,39	2,89	2,92	3,05
Mean degree	27,13	58,58	19,33	15,66	21,56	15,51	9,18	14,76	13,33	18,20	19,91	12,56
Fraction of *N* above the mean (%)	27,73	35,33	23,87	28,05	28,10	34,60	25,97	25,94	27,12	25,93	27,62	31,23
Mean vertex size	3,06	2,34	1,86	2,74	2,26	1,33	1,91	2,07	1,67	3,14	2,59	1,54
Standard deviation	5,57	4,71	2,84	5,14	3,74	0,90	3,14	3,68	2,09	5,77	4,39	1,67

Note: For an explanation of the descriptors see page 4; for an explanation of abbreviations see Annex.

Table 4: Characteristics of the organisation projection of FP sub-networks by instruments (IP)

Graph Characteristics	AERO_6_IP	ENV_6_IP	ICT_6_IP	LIFESCI_6_IP
No. of vertices *N*	595	1953	2119	917
No. of edges *M*	16.630	47.658	41.885	15.370
No. of components	1	2	1	2
N for largest component	595	1.936	2.119	909
Share of total (%)	100,00	99,13	100,00	99,13
M for largest component	16.630	47.522	41.885	15.342
Share of total (%)	100,00	99,71	100,00	99,82
N for 2nd largest component	0	17	0	8
M for 2nd largest component	0	272	0	56
Clustering coefficient	0,89	0,88	0,87	0,83
Diameter of largest component	4	5	4	4
ℓ largest component	2,05	2,61	2,56	2,33
Mean degree	55,90	48,80	39,53	33,52
Fraction of *N* above the mean (%)	38,32	28,67	29,87	30,75
Mean vertex size	1,56	1,47	1,69	1,75
Standard deviation	1,59	1,26	2,17	1,98

Note: For an explanation of the descriptors see page 4; for an explanation of abbreviations see Annex.

Table 5: Characteristics of the organisation projection of FP sub-networks by instruments (NoE)

Graph Characteristics	AERO_6_NoE	ENV_6_NoE	ICT_6_NoE	LIFESCI_6_NoE
No. of vertices *N*	43	449	914	568
No. of edges *M*	394	10.905	24.231	17.158
No. of components	1	1	1	1
N for largest component	43	449	914	568
Share of total (%)	100	100	100	100
M for largest component	394	10.905	24.231	17.158
Share of total (%)	100	100	100	100
N for 2nd largest component	0	0	0	0
M for 2nd largest component	0	0	0	0
Clustering coefficient	0,95	0,94	0,86	0,88
Diameter of largest component	2	5	4	3
ℓ largest component	1,53	2,41	2,24	2,03
Mean degree	18,33	48,57	53,02	60,42
Fraction of *N* above the mean (%)	53,49	43,88	37,64	29,93
Mean vertex size	1,14	1,18	1,58	1,50
Standard deviation	0,46	0,51	1,33	1,13

Note: For an explanation of the descriptors see page 4; for an explanation of abbreviations see Annex.

3.4 Characteristics of project networks

Table 6 contains the characteristics for the networks of projects and is complementary to the tables of organisation projection characteristics. From FP1 to FP4 a growing number of single projects were funded. With the establishment of new instruments in FP5 and FP6, like Thematic Networks, Integrated Projects and Networks of Excellence, intended to integrate a larger number of organisations within the same project, the number of projects decreases. As well as for the organisation networks, the majority of the vertices of the project networks is comprised in a giant component. In FP5 and 6, organisations participate in larger projects (indicated by the increased mean vertex size in the p-graph, i.e. number of participants in projects), but fewer projects (indicated by the decreasing mean vertex size in the o-graph, i.e. number of projects per organisation). This leads to opposing trends in the development of the clustering coefficient in both graphs: while organisations are embedded in increasingly clustered networks (rising cluster coefficient in o-graph), the connectivity between projects decreases (sinking cluster coefficient in p-graph).

Table 6: Characteristics of the project projections (p-graph)

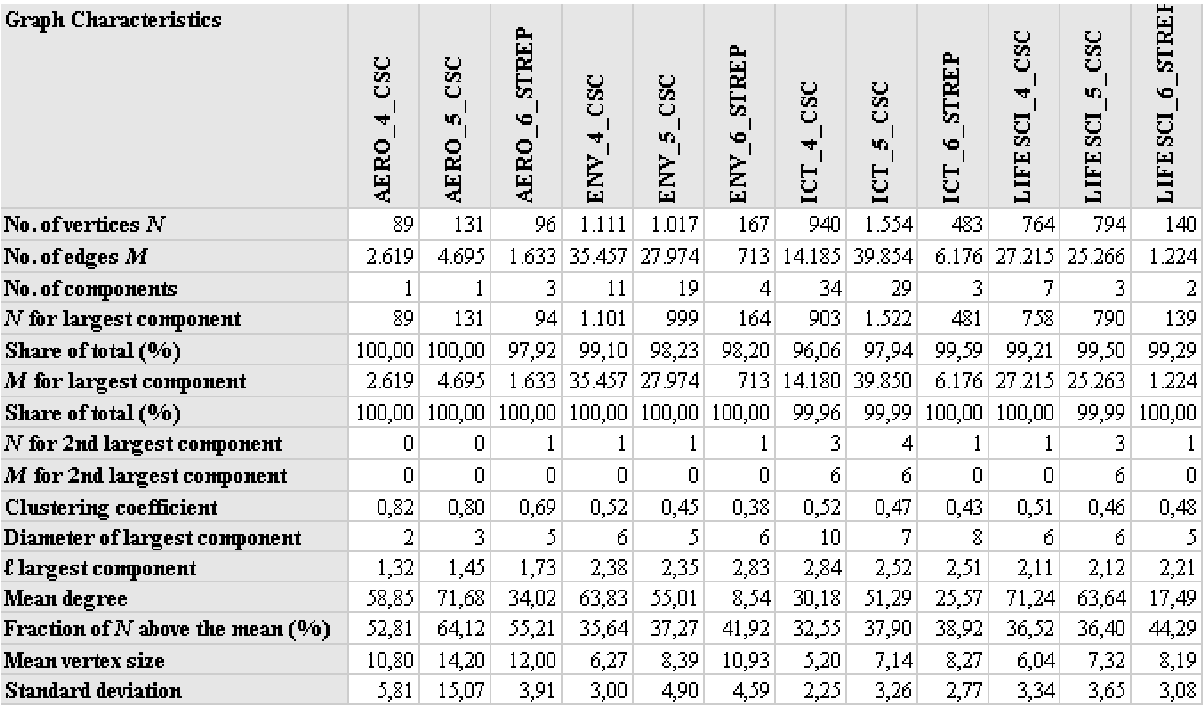
Graph Characteristic	FP1	FP2	FP3	FP4	FP5	FP6
No. of vertices N	1,696	3,013	4,611	11,374	9,491	3,981
No. of edges M	30,080	111,963	277,175	688,207	609,101	251,049
No. of components	53	45	123	367	631	26
N for largest component	1,637	2,969	4,485	10,970	8,839	3,956
Share of total N (%)	96,5	98,5	97,3	96,4	93,1	99,4
M for largest component	30,073	111,963	277,171	688,156	609,079	251,049
Share of total M (%)	99,9	100,0	99,9	99,9	99,9	100,0
N for 2nd largest component	2	1	2	5	2	1
M for 2nd largest component	2	0	2	20	2	0
Clustering coefficient	0,60	0,52	0,49	0,57	0,44	0,46
Diameter of largest component	8	6	7	10	9	6
ℓ largest component	2,90	2,54	2,48	2,73	2,59	2,34
Mean degree	35,5	74,3	120,2	121,0	128,4	126,1
Fraction of N above the mean (%)	39,2	37,4	35,6	33,2	34,0	36,6
Mean vertex size P	3,7	6,0	6,5	5,7	7,6	11,8
Standard deviation of P	2,1	4,5	4,4	4,6	6,4	9,6

Note: For an explanation of the descriptors see page 4.

3.5 Characteristics of project sub-networks by instruments

Table 7 to Table 9 present the properties of the project sub-networks based on different instruments in different thematic priorities. As for the organisation networks, the results of the thematic priorities are similar to the results of the FP project networks in general, but some differences across instruments and topics can be stated. Aerospace, e.g., is characterised by a smaller number of participating organisations, a smaller number of projects, but a higher clustering coefficient in the o-graph as well as in the p-graph. This indicates for Aerospace a more densely connected community than for other thematic priorities. In the field of Energy and Environment the comparatively small clustering coefficient in the p-graph of FP6 can be interpreted as evidence for a stronger differentiation between projects.

Table 7: Characteristics of the project projection of FP sub-networks by instruments (CSC and STREP)



Graph Characteristics	AERO_4_CSC	AERO_5_CSC	AERO_6_STREP	ENV_4_CSC	ENV_5_CSC	ENV_6_STREP	ICT_4_CSC	ICT_5_CSC	ICT_6_STREP	LIFE SCI_4_CSC	LIFE SCI_5_CSC	LIFE SCI_6_STREP
No. of vertices *N*	89	131	96	1.111	1.017	167	940	1.554	483	764	794	140
No. of edges *M*	2.619	4.695	1.633	35.457	27.974	713	14.185	39.854	6.176	27.215	25.266	1.224
No. of components	1	1	3	11	19	4	34	29	3	7	3	2
***N* for largest component**	89	131	94	1.101	999	164	903	1.522	481	758	790	139
Share of total (%)	100,00	100,00	97,92	99,10	98,23	98,20	96,06	97,94	99,59	99,21	99,50	99,29
***M* for largest component**	2.619	4.695	1.633	35.457	27.974	713	14.180	39.850	6.176	27.215	25.263	1.224
Share of total (%)	100,00	100,00	100,00	100,00	100,00	100,00	99,96	99,99	100,00	100,00	99,99	100,00
***N* for 2nd largest component**	0	0	1	1	1	1	3	4	1	1	3	1
***M* for 2nd largest component**	0	0	0	0	0	0	6	6	0	0	6	0
Clustering coefficient	0,82	0,80	0,69	0,52	0,45	0,38	0,52	0,47	0,43	0,51	0,46	0,48
Diameter of largest component	2	3	5	6	5	6	10	7	8	6	6	5
ℓ largest component	1,32	1,45	1,73	2,38	2,35	2,83	2,84	2,52	2,51	2,11	2,12	2,21
Mean degree	58,85	71,68	34,02	63,83	55,01	8,54	30,18	51,29	25,57	71,24	63,64	17,49
Fraction of *N* above the mean (%)	52,81	64,12	55,21	35,64	37,27	41,92	32,55	37,90	38,92	36,52	36,40	44,29
Mean vertex size	10,80	14,20	12,00	6,27	8,39	10,93	5,20	7,14	8,27	6,04	7,32	8,19
Standard deviation	5,81	15,07	3,91	3,00	4,90	4,59	2,25	3,26	2,77	3,34	3,65	3,08

Note: For an explanation of the descriptors see page 4; for an explanation of abbreviations see Annex.

Table 8: Characteristics of the project projection of FP sub-networks by instruments (IP)

Graph Characteristics	AERO_6_IP	ENV_6_IP	ICT_6_IP	LIFESCI_6_IP
No. of vertices *N*	27	97	176	87
No. of edges *M*	255	1.228	4.347	1.646
No. of components	1	2	1	2
***N* for largest component**	27	96	176	86
Share of total (%)	100,00	98,97	100,00	98,85
***M* for largest component**	255	1.228	4.347	1.646
Share of total (%)	100,00	100,00	100,00	100,00
***N* for 2nd largest component**	0	1	0	1
***M* for 2nd largest component**	0	0	0	0
Clustering coefficient	0,85	0,49	0,53	0,60
Diameter of largest component	3	4	3	3
ℓ largest component	1,24	1,81	1,74	1,54
Mean degree	18,89	25,32	49,40	37,84
Fraction of *N* above the mean (%)	74,07	50,52	47,16	56,32
Mean vertex size	34,30	29,67	20,36	18,49
Standard deviation	14,69	13,19	10,61	6,79

Note: For an explanation of the descriptors see page 4; for an explanation of abbreviations see Annex.

Table 9: Characteristics of the project projection of FP sub-networks by instruments (NoE)

Graph Characteristics	AERO_6_NoE	ENV_6_NoE	ICT_6_NoE	LIFESCI_6_NoE
No. of vertices *N*	3	16	49	29
No. of edges *M*	3	43	600	261
No. of components	1	1	1	1
***N* for largest component**	3	16	49	29
Share of total (%)	100,00	100,00	100,00	100,00
***M* for largest component**	3	43	600	261
Share of total (%)	100,00	100,00	100,00	100,00
***N* for 2nd largest component**	0	0	0	0
***M* for 2nd largest component**	0	0	0	0
Clustering coefficient	1,00	0,55	0,71	0,75
Diameter of largest component	1	4	3	2
ℓ largest component	0,67	1,73	1,47	1,31
Mean degree	2,00	5,38	24,49	18,00
Fraction of *N* above the mean (%)	0,00	56,25	55,10	48,28
Mean vertex size	16,33	33,00	29,39	29,45
Standard deviation	4,03	18,41	14,19	19,68

Note: For an explanation of the descriptors see page 4; for an explanation of abbreviations see Annex.

4 Summary and Conclusions

Analysing empirical R&D collaboration networks is one of the key fields of research in Economics of Innovation. The focus of this Deliverable is on the empirical analysis of R&D collaboration constituted under the heading of the European Framework Programmes (FPs). Novel empirical insights into the European FP networks are not only of great current interest in a scientific context, they are a sine-qua non condition to reach the overall goal of NEMO that is the investigation of the interplay between political governance, structure and function of politically induced R&D collaboration networks. The empirical analyses described in this Deliverable produce promising results in the light of relevant theoretical and empirical literature, but also in a European policy context. It provides novel and unique information on the structure and evolution of R&D collaboration networks that have emerge within the European FPs, as well as the impact of different governance regimes on observable network structures and properties.

We identified and described global characteristics of the empirical networks for FP1 to FP6. In general, most empirical R&D collaboration networks analysed exhibit high clustering coefficients, small characteristic path lengths and diameter that scales at most logarithmically with the number of vertices. Networks with high clustering coefficients are called cliquish. In terms of what we presently know about knowledge creation and knowledge diffusion in social networks (Cowan 2006), this is a positive result. There is a slight increase in the clustering coefficient from FP1 to FP5. This suggests that integration between collaborating organisations has increased over time, indicating that Europe has already been moving towards a more closely integrated ERA in the earlier Framework Programs (on this, see Breschi and Cusmano, 2004). The increasing trend continues in FP6, but some caution is warranted: the marked change in FP6 project sizes would be expected to increase the clustering coefficient as well.

The empirical results serve as a sound basis for further policy and governance related considerations in the final workpackage of NEMO. Further, some ideas for a future research agenda come to mind. *First*, the dynamic analysis of the empirical networks remains far from satisfactory since at the current stage we are only able to use the different FPs as reference points in time. Thus, data structures and new methods to gain a higher granularity of time will be need to enable future research in this direction. This could be reached by a combination of community identification methods and the implementation of participation windows of organization in time.

Further research could be directed towards the spatial evolution of the networks, combining spatial interaction modeling techniques – as a useful instrument to describe collaboration taxonomies – and community identification methods. Hereby, the evolution of geographical patterns and FP community groups could be analysed.

Annex

Table 10: Abbreviations for sub-networks

Sub-network	Thematic Priority	FP	Sub-Programme	Instrument
AERO_FP4_CSC	Aerospace	FP4	BRITE/EURAM 3-Aeronautics	CSC
AERO_FP5_CSC	Aerospace	FP5	GROWTH-Key Action Aeronautics	CSC
AERO_FP6_IP	Aerospace	FP6	FP6-AEROSPACE	IP
AERO_FP6_NoE	Aerospace	FP6	FP6-AEROSPACE	NoE
AERO_FP6_STREP	Aerospace	FP6	FP6-AEROSPACE	STREP
ENV_FP4_CSC	Energy/Environment	FP4	ENV 2C; NNE-JOULE C	CSC
ENV_FP5_CSC	Energy/Environment	FP5	EESD	CSC
ENV_FP6_IP	Energy/Environment	FP6	FP6-SUSTDEV	IP
ENV_FP6_NoE	Energy/Environment	FP6	FP6-SUSTDEV	NoE
ENV_FP6_STREP	Energy/Environment	FP6	FP6-SUSTDEV	STREP
ICT_FP4_CSC	ICT	FP4	ESPRIT 4	CSC
ICT_FP5_CSC	ICT	FP5	IST	CSC
ICT_FP6_IP	ICT	FP6	FP6-IST	IP
ICT_FP6_NoE	ICT	FP6	FP6-IST	NoE
ICT_FP6_STREP	ICT	FP6	FP6-IST	STREP
LIFESCI_FP4_CSC	Life Science	FP4	BIOMED 2; BIOTECH 2	CSC
LIFESCI_FP5_CSC	Life Science	FP5	LIFE QUALITY (without KA Food and Agric.)	CSC
LIFESCI_FP6_IP	Life Science	FP6	FP6-LIFESCIHEALTH	IP
LIFESCI_FP6_NoE	Life Science	FP6	FP6-LIFESCIHEALTH	NoE
LIFESCI_FP6_STREP	Life Science	FP6	FP6-LIFESCIHEALTH	STREP

References

Barber, M. J. (2007): Modularity and community detection in bipartite networks. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 76(6):066102.

Barber M., Heller-Schuh B., Roediger-Schluga Th., Scherngell Th. (2008): NEMO - Network Models, Governance and R&D Collaboration Networks. Deliverable D4.1: The sysres EUPRO database manual. Austrian Research Centers GmbH – ARC

Barber, M. J., Krueger, A., Krueger, T. and Roediger-Schluga, T. (2006): The network of EU-funded collaborative R&D projects. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 73(3), 036132 (arXiv:physics/0509119).

Breschi, S. and Cusmano, L. (2004): Unveiling the texture of a European Research Area: Emergence of oligarchic networks under EU Framework Programmes. *International Journal of Technology Management*, 27(8), 747-772.

Cowan, R. (2006): Network models of innovation and knowledge diffusion. In Breschi, S. and Malerba, F. (Eds.), *Clusters, Networks and Innovation*. Vol. 29-53. Oxford: Oxford University Press.

EC (2003): Participation Rules Vademecum, European Commission. http://cordis.europa.eu/fp5/management/particip/v_participation.htm, accessed 30 May 2009.

EC (2004): Classification of the FP6 Instruments. Detailed description. ftp://ftp.cordis.europa.eu/pub/fp6/docs/annex_instruments.pdf, accessed 30 May 2009.

Luukkonen, T. (2001): Old and new strategic roles for the European Union Framework Programme. Science and Public Policy. Vol. 28 No 3. 205-218.

Roediger-Schluga, T. and M. J. Barber (2006), 'The structure of R&D collaboration networks in the European Framework Programmes', UNU-MERIT working paper, 2006-036, available from http://www.merit.unu.edu/publications/wppdf/2006/wp2006-036.pdf

Wassermann, S. and Faust, K. (1994): Social network analysis. Methods and applications. Cambridge: Cambridge University Press.

Watts, D. J. and Strogatz, S. H. (1998): Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440-442.