IRIS (1978-2006) Historical Reflection through Visual Analysis

Judith Molka-Danielsen
Associate Professor, Molde University College, Norway
j.molka-danielsen [at] himolde.no

Matthias Trier
Dr-Ing., Technical University of Berlin, Germany
trier[at]sysedv.tu-berlin.de

Vadim Shlyk
Technical University of Berlin, Germany

Annette Bobrik
Technical University of Berlin, Germany

Markku I. Nurminen
Professor, University of Turku, Finland

Abstract. IRIS (Information Systems Research in Scandinavia) Conference is an annual meeting for researchers in Scandinavia. Through 29 annual conference meetings a social network of co-authorships and general research collaboration has emerged. This article contributes a historical reflection and analysis of the social network of researchers associated with the IRIS in the period from 1978 to 2006. We make use of the network visualization and analysis tool called Commetrix. We will examine the influence of certain researchers in this network by identifying their betweenness centrality and degree centrality. We will reflect on the nature and characteristic of the research topics by analyzing the network of keywords found in the conference paper titles, and visualizing their representation in the network.

Introduction

Information Systems Research in Scandinavia or IRIS is an annual meeting of researchers of information systems that began in 1978, where the first meeting took place in Tampere, Finland. Indeed the first 4 meetings were in Finland, championed by founders Pertti Järvinen and Pentti Kerola, and a majority of Finnish contributors. After that time the meetings have rotated among the Scandinavian countries of Denmark, Finland, Norway and Sweden. One participant that came in the early years and has attended nearly all the meetings is Markku I. Nurminen. He describes through his experience the birth of the conference and the founding elements of the research community, "The scientific community in Information Systems Research was, indeed, existing in the Scandinavian countries already before that date (of the first IRIS)."(Nurminen, 1997, p.48) But, how do we evaluate the significance of IRIS? It was founded as a forum where the ideas of Scandinavian researchers could be debated, expressed, and explored. This also prepared works for a greater impact on the international scientific community. Nurminen expresses this motive in the early meetings:

> "The discussion in all four meetings was focused on systemeering and systemeering method(ologie)s, which also was the explicit intention. The specific theme of IRIS 2 was ISAC[1] and of IRIS 4 the PSC[2] model. Another passion was the definition of the basic concepts and their underlying assumptions."(Nurminen, 1997, p.51)

In 2006, some of these terms will no longer be found in present day article titles. Terms like "infologically oriented systemeering" have transferred meaning to terms like "Information Systems Development" in more recent literature. (Lundeberg, et al., 1981), (Goldkuhl, 2006) While citations are often used to determine impact of research in research communities, we think this is only one measure. Social networks of themselves have great social value and lead to future collaborations. This paper will examine a social network of co-authors that presented works in association with the annual IRIS conference. We will use social network analysis (SNA) to examine the centrality of some of the IRIS community members in contributing to the growth of this community of researchers. In addition, we will reflect on the evolution of topics in this network over time using a novel technical method to visually animate the actual dynamics of the co-authorship network over time.

This paper is organized as follows: relevant research is presented in the Literature Review. The section Research Methodology informs of our source of data and approach. Visual Results through Commetrix are next presented. This

[1] "In 1973 ISAC was an acronym for Information Systems for Administrative Control. Later it was changed to Information Systems work and Analysis of Changes." (Goldkuhl, 2006, p.1)

[2] PSC is a systemeering model with three phases: semantic (S), pragmatic (P) and constructive (C). It was much discussed at the early IRIS meetings. (Kerola, 1979) (Kerola and Järvinen, 1979) (Kerola and Taggart, 1980) (Taggart, 1980)

section is followed importantly by Interpretation of Results in a Historical Context and summarized in Concluding Remarks.

Literature Review

"Social networks are information networks of humans interacting and creating relationships. A link between two persons in the network can be a close friendship, partnership, or working relationship. (Molka-Danielsen, Berge Søvik, 2005) Many studies examine the general structural and dynamic network characteristics of social networks, and in particular of emerging researcher networks." (Barabási, Jeong, et al., 2002) (De Castro, Grossman, 1999) (Grossman, 2004) (Newman, 2001)

In our study of the IRIS co-authorship network we use social network analysis (SNA) to identify key persons in the network and the spreading of keywords in the article titles over time.

Authors are the actor *nodes* of the social network. They are individuals. Co-authors on articles may be from the same or different countries.

Links or ties between the actor nodes (authors) are existence of at least two persons names (listed as authors) on one article that was presented at one of the IRIS conferences. It is not necessary for the author to have been present at the conference. The ideas are considered shared in retrospect, because the conference took place. (We would have liked to do analysis on the "working groups" at IRIS, but time and incomplete data did not permit it.)

The *relationship* between the nodes of this network can be strengthened with each new co-authored article between any two authors. The co-authorship relationship is not directional. We pursue two modes of analysis. In the accumulative analysis the strength of a tie does not depreciate with time. Alternatively, it is possible to look at time slices of the network, i.e. to examine relationships within a time window of 1 year, 5 years, any length up to the full 29 year period. This allows indicating the situation within defined period (cf. figure 4).

We use several metrics of SNA commonly used to examine the individuals in the network, these are: *direct contacts* (that is the number of first degree co-authors), the *betweenness centrality* and the *degree centrality*. These measures have been defined by Wasserman and Faust (1994) and by Freeman (1979). We adhere to these definitions and note applied differences in the appendix.

Direct contacts are a measure of the activity of an individual actor. This is the number of direct co-authorships with another person.

The *count of papers* are a count of how many papers a person contributed to, that is, was an author or a co-author. Individually authored papers are counted in the network as self-relations.

The *degree centrality* measures the importance of an actor by counting the number of direct contacts of an actor in ratio with the standardized maximum number of direct contacts. This measure depends on what connected component of the network the actor is a part of, and the time period in consideration.

The *betweenness centrality* is a measure of interpersonal influence that an actor has on others. It refers to how many pairs of actors the actor in measure stands between. If two other actors in a network can only contact each other through a third actor, this third actors holds a social capital that can manipulate the resources on the sides. (Burt, 1997) Our measure of betweenness centrality is based on Wasserman and Faust (1994) however, because of the size of the network; it is based on only one shortest path between actors. This alternative approach is supported in (Brandes, 2001).

Degree centrality and *betweenness centrality* are expressed as egocentric measures. These along with *average degree* or *average betweenness* are also part of the general or global network properties. Another global property includes the *diameter*. These measures for IRIS (reported later in Table 4) are given for the largest connected network component. The *diameter* is the average longest shortest path between any two actors of a network. This can be an expression of the spread of the network. The *average degree* can be used to measure the likelihood of a large connected component, where most members of the network are a member of the large component. In addition to these global characteristics we examine ego-centric characteristics of key members of the network.

Research Methodology

The data on the IRIS co-authorship community has been collected through a variety of sources. These include the IRIS-History web-pages that contained data from 1978 to 1998. (Accessed in 2006: http://www.cs.utu.fi/IRIS/hist.htm) From 1999 to 2006 we had PDF files of proceedings on CD and sourced the IRIS Conference web site: (http://www.cs.auc.dk/IRIS/conference/conference.html). Finally, we sought and found "missing pieces" of information from the personal (printed) archives of Markku Nurminen.

In order to prepare the data for network visualization with the employed analytical tool Commetrix, the data was first entered into a database. Figure 1 gives a simplified account of the adopted data model. It can be represented using the tables *Author and Message.*

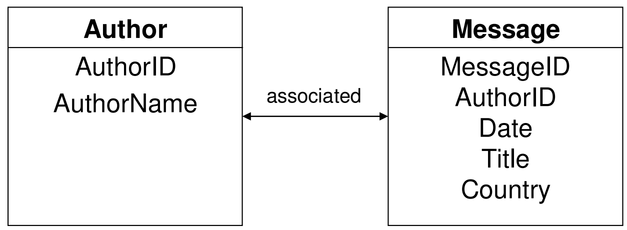

Figure 1. Simplified IRIS data model

The table *"Author"* has the field "AuthorID." For every author an AuthorID and AuthorName were stored.

The table *"Message"* represents all submitted papers of the conference with attributes: MessageID (every paper has a unique number), AuthorID (the first author of the paper), Date (it is an unique parameter for time (year) of the conference that the paper was submitted), Title refers to the title of the paper and Country (numerically refers to the origin land of the all authors of the work in the year submitted). As Commetrix initially has the intention to visualize message based structures, papers were stored as Messages together with the parameters described.

Difficulties in creating the database can be described in two categories: problems with incorrectly written names or given names of the participants and secondly the unequivocal identification of the origin of the participant. In the first case it concerns the orthography of the Scandinavian names and given name. In contrast to English there exist other letters (å, ø, ä, ö, æ, etc.) in all Scandinavian languages (as well as in German, Spanish, etc.) that were written in some cases in the different ways. If we look, for example, at an author with the name Anders Mørch, we saw that his name appeared several times under different spellings: Anders Mørch, Anders Morch, Anders Moerch or Anders Mörch. So we sometimes had to do research after respective persons on the Internet, so that we could be sure that these were really not different persons. Also the first names with the surnames were sometimes reversed or represented differently (abbreviated) in different years. (For example: Marie Luise Christ Neumann and next time Marie L. Christ-Neumann). Data cleansing took several iterations to be sure that the same person did not appear as new persons and did not receive a new AuthorID.

The second factor of difficulty was that for a time span of 29 years that many people have done research in different places. They may be associated with different countries or universities. The location is therefore associated with the authors of the paper in the respective year.

In the second phase of our methodology the data in the database was imported into Commetrix so that visual analysis could be performed on the network. Originally, Commetrix was developed as a software application to support visual representation of electronic communications (in the form of e-mail exchanges) of large geographically separated groups. (Trier, 2005) Later, its application has been extended and generalized towards the dynamic visualization and analysis of other sources of network data (e.g. business process networks). This made it possible to adapt the method for the purpose of our present analysis. With that, we are able for the first time to create novel insights about the dynamics and the evolution of co-authorship networks. Changes or transitions over time can now become a subject of study. With the incorporation of topic filtering we are further able to study the lifecycles of defined topics in the overall academic discourse of our dataset. Our methodology with the use of Commetrix, was as follows:

- Import of the author names, academic and location affiliations of the authors, IRIS conference article titles, conference locations, and conference years.
- Extraction of association of authors as represented through co-authorships on conference articles. This network is the visually represented knowledge community.
- Visualization and evaluation of community properties, structures and the dynamic behaviors of the network portrayed through a longitudinal representation of data over 29 years.

Visual Results through Commetrix

The key factors that we examine are:

- How has the IRIS network of co-authorship associations grown over time;
- Who are the central figures represented through betweenness and centrality statistics;
- What are central topics (themes) as represented in the titles of the IRIS articles and how do they change over time?

By relating our visual analysis to these questions we may be better able to understand the impact and benefits that the IRIS conference has brought to the information systems research community in Scandinavia and minimally, of the impact of contributing authors' individual efforts on the community as a whole.

General Network Characteristics

In Figure 2 we see the entire IRIS network of accumulated co-author associations from 1978 to 2006. The network is comprised of 1252 participants of whom Figure 2 shows the 1219 persons as nodes who are authors of papers. 80 percent of them (976 persons) participated as a co-author on a paper. This collaborative effort is represented by the links in the network of Figure 2.

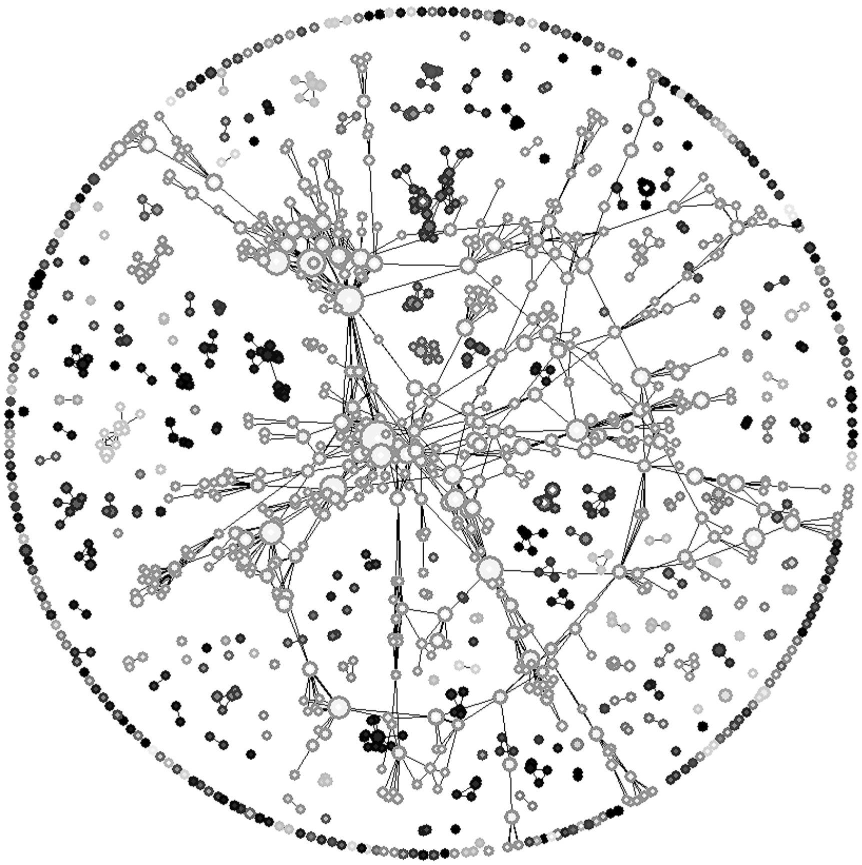

Figure 2: The complete author and co-author relationship network.

The node size in Figure 2 represents each person's activity measured by paper-counts in which the author was a participating co-author. Different node shades represent the various groups of linked authors (a larger colored version is available at http://www.commetrix.de/iris[3]). The network graph shows a co-author network comprised of one large component of 541 linked authors (44%) and several smaller clusters with the second largest cluster only containing 23 authors (2%). Over time, the 44% authors of the large component have produced a very interrelated and large researcher network. Within that cluster, there are several highly connected sections around core sets of linked participants. The graph implies that the network is dominated by those 500 actors and their collaboration. More peripheral co-authorships, which do not belong to this central cluster, appear as small network components, meaning that they remained much

[3] All the graph visualizations in this article are available via this paper's online supplement at http://www.commetrix.de/iris and in addition further detail of all the images of this article can be seen in this pdf file by adjusting the view to 200% or larger.

more isolated and less complex. About 10 actors appear very active (large nodes) in this network. Looking at their embeddedness it can be seen, that those active authors are also the more central persons in the IRIS network. This finding is also supported in Figure 3, which shows a reduced subset of 256 actors with a minimum of three contributed papers. The larger the node size the more contacts a node has in the researcher network and the lighter the nodes' color shade the larger is the contributed number of co-authored papers.

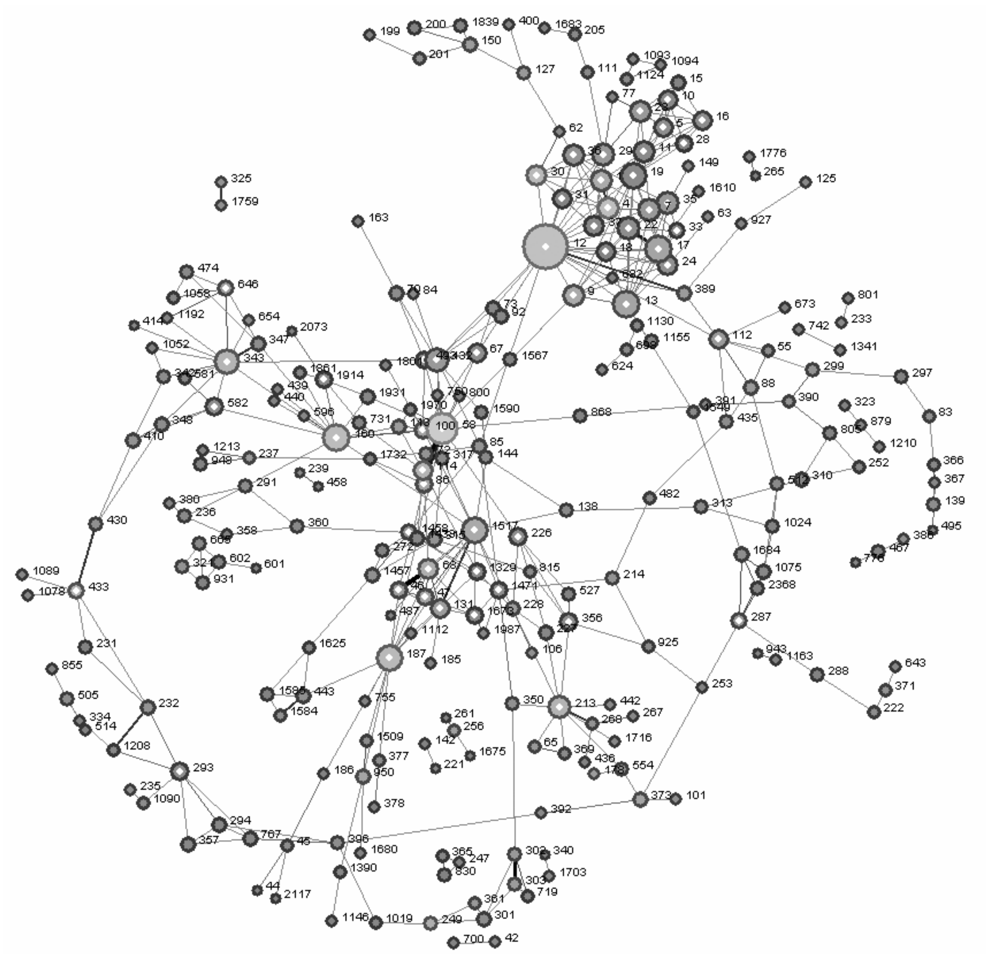

Figure 3. Subset of 256 actors with a minimum of three contributed papers (440 relationships). (density=1.35%; average degree = 4,182 links).

Generally more connected (larger) nodes are also more actively publishing at IRIS (lighter color shades). In Figure 3's core structure of the IRIS co-authorship network, a central axis emerges with nodes 12, 58, 1517, 160, and 187 at the center. Several will be analyzed in more detail in our analysis of centrality and ego-networks later in this section. Next to this highly integrated central structure a second pattern can be found in the core network of IRIS co-authorships. It appears as long paths reaching out towards the periphery of the network and as very large (intertwined) rings with more than 13 steps. The authors positioned in these structures form large chains of peer relationships. Each participant of such a structure entertains only a very small set of two or three co-authoring peers. Those peers are embedded in the same structure and so on.

Further insights can be achieved by analyzing the evolution of the network. Figure 4 demonstrates four snapshots of the IRIS network's development across time. The visualization and a corresponding animation of the actual evolution of this network over time can be best examined using the online supplement at the referenced website.

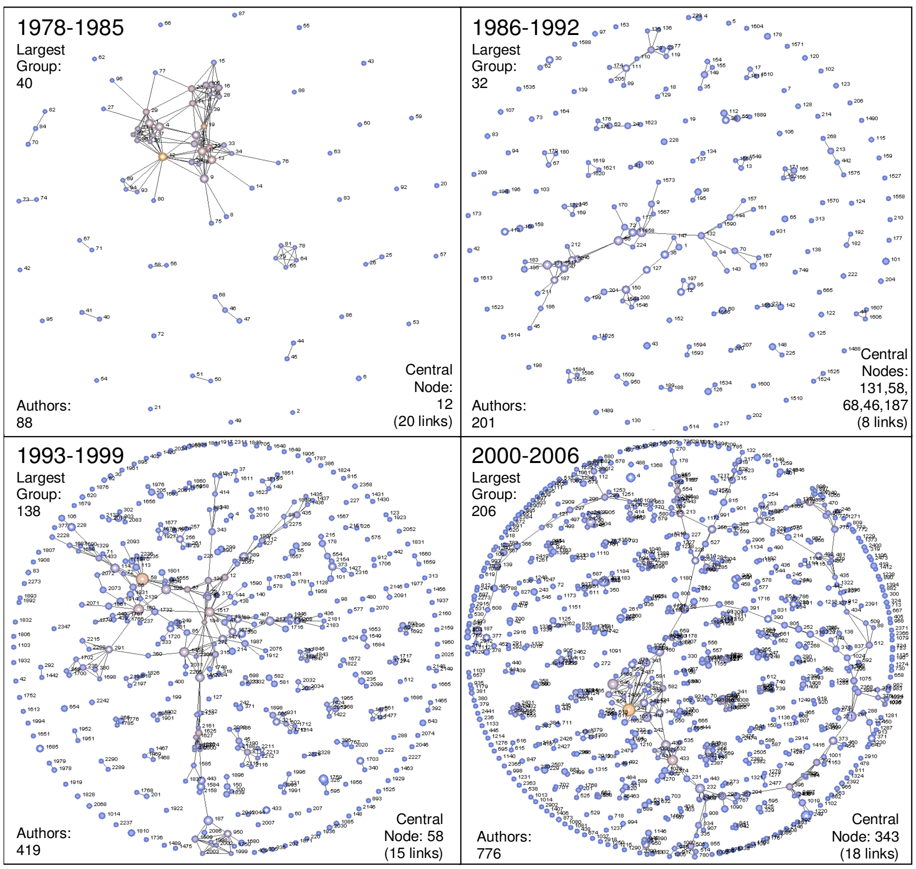

Figure 4: The different states of IRIS network in four periods (relationships outside the period are ignored).

This Figure 4 further shows the number of authors during the selected four periods, e.g. between 1978 and 1985 the co-authorship network contains 88 authors (those who alone contributed papers are included). Nodes and links before and after the filtered period are not displayed. This gives an impression of the actual differences in the network structures. The node size is related to the contributed papers of an author, the color shade represents the direct relationships. Also, the number of the largest group and the ID of the most central node are given. During the period 1986-1992, this position was shared by 5 authors (46, 58, 68, 131, and 187). Names are not shown in the graphs, but a low number generally indicates an early first occurrence of the author in our 29 years of data, indicating that these central nodes in the second period were already participating in the first period. The figure shows, that in the first quarter of the sample period (1978-1985) many individual papers have been published, but one dominating triangular cluster with 40 members and many co-authorships was present. There is one very central author (node 12). In the next seven years this cluster loses its central position and declines to a pair of small linear subnetworks which however reaches out to new peers. Next to this change, around the small triads of the first period a different dominant cluster with 32 participants appears. It also has a linear shape with a central axis of co-authors. The overall number of actors of IRIS has risen recognizably but still many authors are isolated contributors. The volume of the conference further grows in the third period (1993-1999). The largest cluster of the second period grows 138 people. From the central nodes of period two, node 58 emerges as the best connected IRIS participant. From the initial core participants, only author 12, who was the most central node in the first observed period reconnects to the large cluster, establishes a larger number of direct ties, and becomes an integral part of the active axis of period 3. Most of the remaining structures formed by the initial core group disappear. The network structure of this component is shaped like a star with a relatively dense connected center and linear chains of actors reaching towards the periphery. Finally in period four, the center of the network moves again. The main change is the strong development on the indirect path between the dominating node of period 2 and 3 and a new node, 343, which has been in a peripheral side arm in period 3. This node emerges to become the best connected scientist in period 4 and establishes a very dense center around him. Node 58 is still integral in period 4. The previous star structure somewhat dissolves and very long paths (which sometimes close to form ring structures) appear.

Resulting from this dynamic analysis, a constant shift of the epicenter of activity can be observed. The initial cluster almost dissolves and gives way to a second core which builds for a long time and slowly transitions in its center. By separating the analysis into four shorter temporal periods it uncovers that the final network clustering is reduced to predominantly linear networking formations. As most papers are written by two or three authors, this implies that there are small working groups of two or three peers which are only sequentially connected.

Statistics of the Yearly Paper Count

A summary of the yearly paper count and number of unique authors are represented in Table 1.

Year	Host Country	Host Location	Author Count	Paper Count
1978	Finland	Tampere	15	16
1979	Finland	Dragsfjärd	10	14
1980	Finland	Saarijärvi	19	18
1981	Finland	Oulu	26	23
1982	Sweden	Stockholm	17	12
1983	Norway	Øystese	25	17
1984	Finland	Helsinki	37	26
1985	Denmark	Århus	30	23
1986	Sweden	Lund	36	29
1987	Finland	Tampere	52	42
1988	Norway	Røros	39	35
1989	Denmark	Skagen	63	37
1990	Finland	Turku	53	38
1991	Sweden	Umeå-Lövånger	45	36
1992	Norway	Larkollen	84	56
1993	Denmark	Copenhagen	113	81
1994	Finland	Syöte	97	75
1995	Denmark	Gjern	77	53
1996	Sweden	Lökeberg	79	54
1997	Norway	Hankø	88	56
1998	Denmark	Sæby	108	68
1999	Finland	Keuruu	146	92
2000	Sweden	Uddevalla	193	119
2001	Norway	Bergen	259	163
2002	Denmark	Bautahöj	132	92
2003	Finland	Porvoo	223	137
2004	Sweden	Falkenberg	167	115
2005	Norway	Kristiansand	147	92
2006	Denmark	Helsingør	133	77

Table 1. IRIS Conference paper counts by year and location.

There are 1696 papers for all years 1978-2006: these are 1598 single nation (written by one or multiple authors all from one nation designated National in Table 2), 44 Scandinavian collaborations, 54 external international collaborations. Of the 1598 national papers, 1424 are from Scandinavian countries and 174 are from non-Scandinavian countries). Not shown in Table 2, are that 44 papers have joint co-authors of two to four Scandinavian countries. And 54 papers are by co-authors from non-Scandinavian countries, these are external international. The columns for paper counts with Scandinavian authors and international authors cannot be added down, and they will not sum to 44 and 54 respectively, as papers with authors from two or more nations will be listed in multiple rows.

Country	Paper count with all author(s) National	Paper count with Scandinavian author(s)	Paper count with International author(s)
Sweden	574	28	14
Finland	424	16	17
Denmark	232	21	17
Norway	194	22	11
UK	41	13	
Germany	38	5	
Australia	25	5	
Italy	14	5	
USA	13	14	
Others	43		
TOTAL	1598		

Table 2. Paper Count by Country

General network properties are made on the largest connected network component. These are represented in Table 3. Besides the properties defined earlier, we describe several more here. The amount of links, 1066, is the number of co-authorship relationships in the largest connected component. The amount of authors, 541, is the number of unique authors in this connected component. The number 254 of core group members is the number of persons of the connected component that produced 80% of the papers overall. That means of the 1696 papers produced for all years that the ratio of 17.3% of contributing authors produced 80% of the articles. The density of the largest component is fairly high, that is 72% is the ratio of the number of existing links compared to the number of possible links between unique authors. Finally, the average degree is the average amount of contacts per author in the core.

Largest Network Component Properties	
First Article	1978
Last Article	2006
Amount of authors	541
Amount of links	1066
Diameter	15
Core group members	254
Average link-strength	2.41
Density	0.72%
Average degree (contacts)	4.3
Average betweenness	0.99%
Average centrality	18.56%
Average path length	6.38

Table 3: General network properties of the largest network component

Table 4 shows the top ten authors that have contributed the most articles to IRIS. We will display the egocentric network of some of these top authors to give an impression of the interconnection of researchers that they have supported.

Active Period	Author		Author_ID	Paper Count
1983-2004	Lars	Mathiassen	ID_58	29
1978-2002	Markku	Nurminen	ID_12	23
1978-2006	Pertti	Järvinen	ID_4	22
1989-2003	Carsten	Sørensen	ID_160	21
1981-2006	Per	Flensburg	ID_30	19
1990-2005	Karl Heinz	Kautz	ID_187	19
1986-2005	Peter Axel	Nielsen	ID_114	17
1997-2006	Lars	Svensson	ID_343	17
1994-2006	Urban	Nuldén	ID_433	15
1991-2006	Ole	Hanseth	ID_213	15

Table 4: Top 10 authors and the number of papers contributed.

Tables 5 and 6 show other top listed individual actor measures.

No	Author		Author_ID	Betweenness Centrality
1	Markku	Nurminen	12	31.29%
2	Kristin	Braa	1517	23.32%
3	Ole	Hanseth	213	18.46%
4	Karl Heinz	Kautz	187	17.54%
5	Carsten	Sørensen	160	16.69%
6	Lars	Mathiassen	58	15.63%
7	Joan	Greenbaum	132	15.09%
8	Lars	Svensson	343	14.15%
9	Jukka	Heikkilä	112	10.16%
10	Jonny	Holmström	554	8.46%

Table 5. Actors with the top betweenness centrality measures.

No	Author		Author_ID	Degree Centrality	Co-author Count
1	Markku	Nurminen	12	6.64%	36
2	Lars	Mathiassen	58	4.61%	25
3	Lars	Svensson	343	3.69%	20
4	Carsten	Sørensen	160	3.69%	20
5	Ole	Hanseth	213	3.51%	19
6	Karl Heinz	Kautz	187	3.32%	18
7	Vesa	Savolainen	29	3.14%	17
8	Göran	Goldkuhl	13	2.77%	15
9	Matti	Jakobsson	19	2.96%	16
10	Per	Flensburg	30	2.96%	16

Table 6. Actors with the top degree centrality measures.

Tables 5 and 6 show different ways participants have been core or central to the network. Using the Author_ID in Tables 5 and 6 the participants can be identified also in the broad network visualizations.

Figure 5 shows the relationship network among the top 5 central actors of the above listed in Table 6.

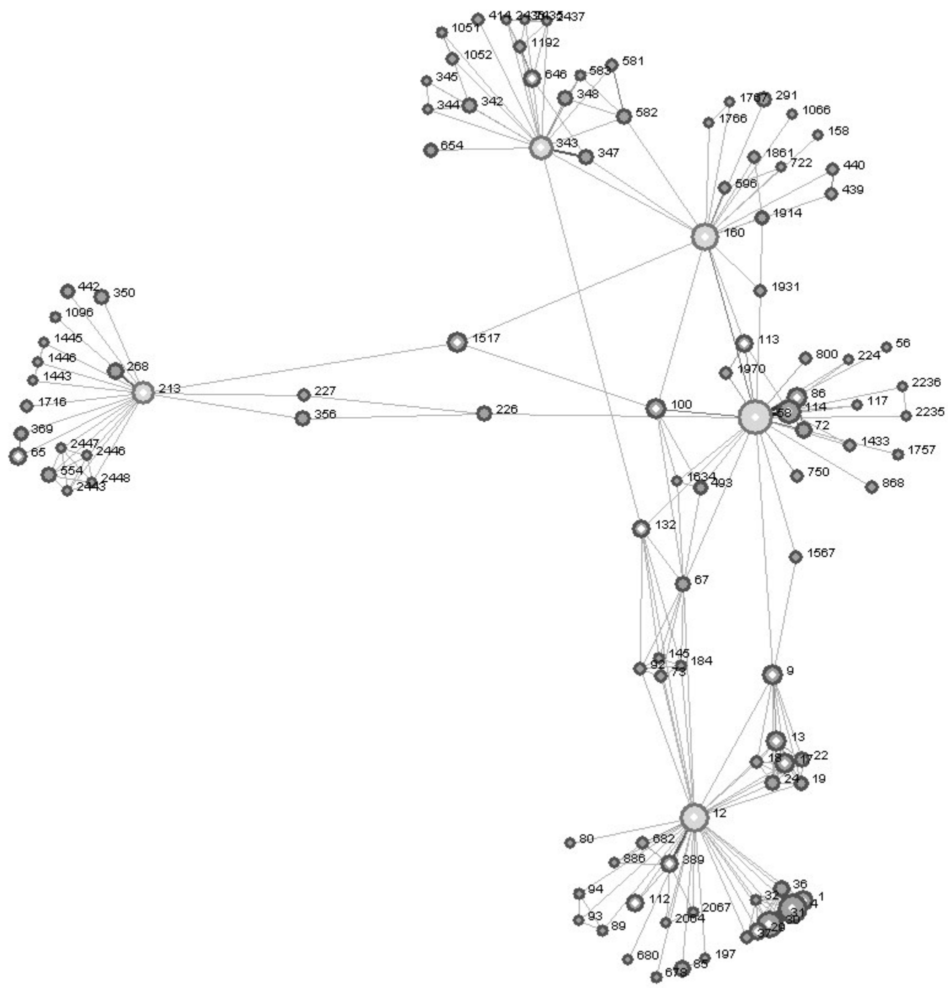

Figure 5. Combined Relationship ego-network among the top 5 central authors (with 107 authors and 248 relationships)

The lighter color nodes in Figure 5 represent the top five central actors selected as ego. Interestingly, four actors of the top6-top10 actors in Table 4 are also found in direct relationship to these top5.

It can be observed, how the combined ego-network forms around initial personal networks of the top 5 actors.

This combined ego-network can also be analyzed for its behavior over time. Figure 6 observes two time periods, the first half until 1991 and the second half of the sampling period. Compared to Figure 4, it can also be seen, that the development of the top 5 central nodes corresponds to the general development of the IRIS network and its shifting epicenters throughout the analyzed four time periods. Node 12 was very central during the first period. Later node 58 moved to the central position in period two. In this period, node 160 was embedded in a very small peripheral component, but it moved towards a very central position in period three (starting in 1993). The same applies for node 213, except that this author started its centralization efforts later in 1998. Node 343 finally is the most connected node in the fourth period although the first co-authorship link was established only in 1997. For the observed IRIS network, this implies that the active centers of the overall structure are formed by very active and prominent authors during their lifecycle of activity. Although three out of the top five have even collaborated directly, their respective groups of co-authors remain much separated (yielding in a high betweenness of the central actors) and create almost no links among each other, i.e. there is almost no triadic closure effect.

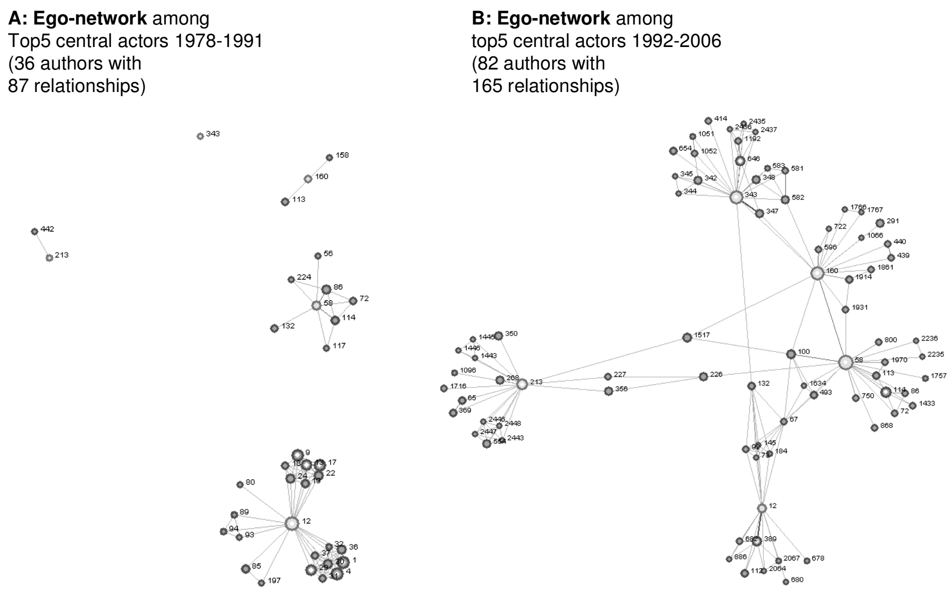

Figure 6. Top 5 key actors and their ego-centric relationships during in two sampling periods.

Keyword Filtering of Words in the Article Titles

Keyword Filtering will enable us to see the influence of themes and topics in IRIS as represented through the frequency of the keyword in the network.

For the list of most prominent terms, similar terms like genitive forms or plurals have been grouped. The following Table 7 lists the most frequently used terms and thus indicates the dominant keywords. We measured the number of participating actors (acts*), which used a certain keyword in a paper.

	Occurrences			Occurrences	
Term	acts*	Rank	Term	acts*	Rank
system	416	1	methods	35	20
information	284	2	practice	35	20
development	221	3	electronic	33	23
design	146	4	theory	32	24
software	109	5	organizational	31	25
knowledge	88	6	quality	31	25
technology	76	7	human	30	27
process	75	8	object	30	27
management	69	9	communication	29	29
computer	64	10	study	29	29
approach	60	11	project	28	31
analysis	56	12	social	28	31
research	54	13	framework	27	33
mobile	53	14	group	27	33
learning	51	15	methodology	27	33
implementation	45	16	cooperative	26	36
support	45	16	improvement	25	37
systemeering	45	16	evaluation	24	38
model	42	19	method	24	38
business	35	20	action	23	40

Table 7. Top 40 Keywords[4]

Measuring the term co-occurrences yields the result that the terms information and system, system and development, software and process, knowledge and management, as well as computer and support, often have been used together in a paper title. This suggests that those terms were used as combined keywords, which gives a first impression of which aspects of Information Systems Research were published most frequently. Table 8 shows keyword co-occurrences.

[4] Acts*= Contributing Acts is the number of Participants in the co-authored paper with the respective keyword. (e.g. 416 mean that 416 actors participated in a paper using the term “system”).

Co-Occurrence in % Matrix of Top 20 Terms		
Term A	Term B	Terms A and B
information	**system**	106
system	**development**	66
software	**process**	19
knowledge	**management**	14
computer	**support**	13
information	development	30
system	design	29
computer	design	11
software	development	15
information	technology	15
process	development	11
development	management	10
system	knowledge	18
system	method	16
system	implementation	15
system	model	15
information	research	11
information	management	10
system	management	13
software	system	14
system	analysis	12
system	research	10

Table 8. Keyword Co-Occurrences

To create a dynamic view of keyword occurrences we split the dataset into time slices with a length of four years. Then the keyword occurrence in each time slice has been counted to check for longitudinal changes in the keywords used. This is shown in Table 9.

Keyword	1978-1981	1982-1985	1986-1989	1990-1993	1994-1997	1998-2001	2002-2005
system	21	46	36	27	10	12	8
information	25	21	14	10	10	11	10
development	11	9	18	15	2	8	9

design	18	9	3	9	9	8	5
software	0	3	1	3	5	9	4
management	0	4	0	3	1	8	6
technology	0	0	6	1	5	3	7
process	1	0	3	1	5	6	4
knowledge	3	4	3	3	1	7	4
method	14	12	2	5	5	2	2
support	0	4	7	2	6	3	3
model	27	4	3	4	0	2	2
computer	0	13	6	4	3	3	1
organization	0	0	6	<1	4	5	1
mobile	0	0	0	0	0	5	5
implementation	0	5	3	0	3	4	3
practice	0	0	1	1	2	4	4
analysis	4	15	6	3	3	<1	1
learning	0	0	1	2	4	2	3
research	13	5	1	1	0	2	2
systemeering	30	1	0	0	0	0	0
PSC	14	3	0	0	0	0	0

Table 9. Development of Keyword Occurrence defined as number of papers with keyword across time in percent of overall papers in that period. Further graphs available via the online supplement at http://www.commetrix.de/iris.

Table 9 highlights two temporal effects. First, the share in percent of occurrence of those keywords (i.e. percent of papers with selected keyword versus overall papers) decreases noticeably over time. This highlights reduced coverage of the selected top twenty and indicates increased pluralism and heterogeneity in terms of keywords: over time the main topics differentiate to a larger set of different terms. Quantitatively, the initial 307 detected keywords per four-year-period grow to 855 keywords (with a peak of 998 around the year 2000 after doubling within the preceding four years).

The second observation is that different topics have different lifecycles: The keyword analysis seems to indicate a trend from basic terms like “systemeering”, “methods”, “model”, “analysis” to terms associated with more application of IT like “mobile”, “technology”, or “management”. A third group seems to remain stable, including the core terms “information” or “development”. Both trends suggest, that a general lifecycle of the whole discipline from general issues to more diverse and application oriented issues can be observed here.

To give a visual account of how the keywords spread over time, we observed the two terms 'information' and 'knowledge' in more detail. The filtered networks of Figure 7 show the co-authorship networks which covered those topics in their papers. It can be seen, how the use of the terms in the papers expanded. These 2 terms are used in the first period. However, the co-authorship network around the term information appeared in 1980, whereas the network around the term knowledge appeared in 1985. The analysis of the usage of the term knowledge further showed that it can be assigned to the context of knowledge engineering. The context of knowledge management became dominant towards the end of the sample period.

A second insight is, that co-authorships using the keyword information grow fairly constant, where as knowledge started to increase its occurrence after the year 1997. The clusters of co-authors using the term knowledge are all rather small, whereas there is a dominant co-author network that used the term information. However, it appears to be a result of very early group formation processes in the first 10 years.

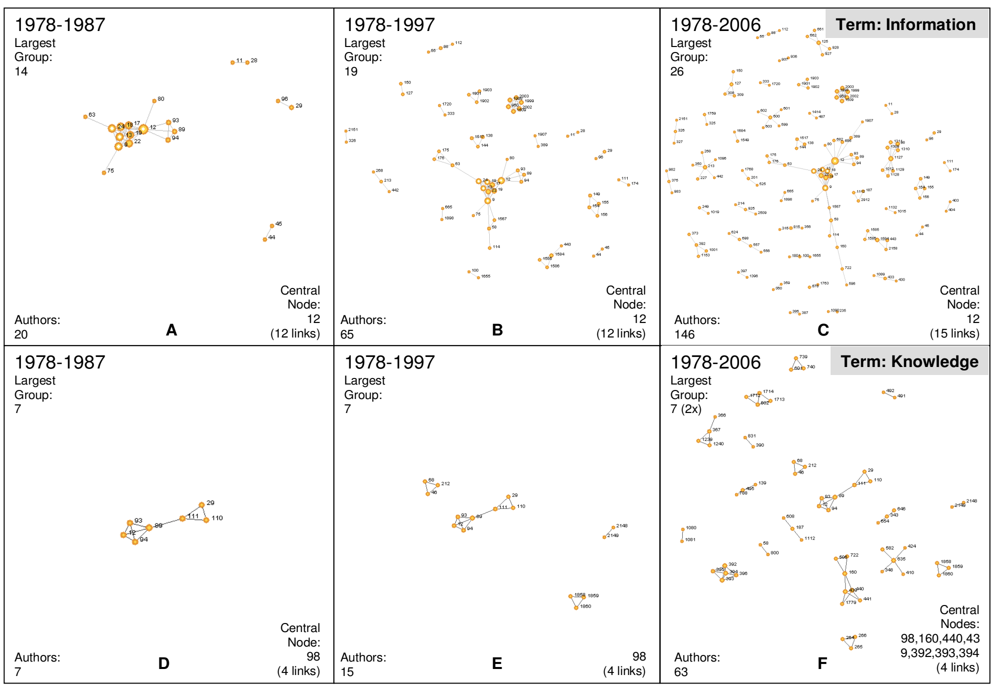

Figure 7. The co-authorship network of authors, who employed the terms information (A-C) and knowledge (D-F) in their papers, indicating the spread of these terms over time (video clip is available at www.commetrix.de/iris)

Interpretation of Results in a Historical Context

Through our approach of visual analysis, we conclude that using information visualization in social network analysis can create transparency on the evolving structures in co-authorship networks. This revealing visual approach allows us to efficiently convey an intuitive impression of the otherwise invisible network – even to users which are unfamiliar to the topic of SNA. We can better identify, explore and present interesting patterns, zones and growth processes within large complex networks. Central authors are not only a list of names, but their role for the network as a focal point of activity, attraction and group formation can be grasped. Dynamic analysis offers opportunities to quickly understand how and when parts of a large network disconnected and rejoined, quickly grew or stalled.

The previous sections have given a visual representation of the co-authorship network in IRIS. If we reflect on this specific network in an historical context, several qualitative interpretations can be made.

The core persons in this study are those who have been active in the submission of articles to the conference. This study however does not tell us about who is central within the Scandinavian school, or research tradition. A valuable extension to this study to place these persons in the larger context would be to identify the citation history of the participants in the conference. So, we may observe in Table 4 that Finnish contributors were very active in the early years, and Swedish actors became more prominent later. But, IRIS activities between countries and faculties do not tell use about activities that took place outside of IRIS.

The 29 years imply significant change for every participant. Juniors become seniors and retire, while some other students leave the academic carrier. The proportion between seniors and juniors has been changing. While again we see the active period of key persons in Table 4, some questions are difficult to evaluate given our present data. For example, are there generations and how did the new generation emerge and at what rate do the old ones leave? These issues would require further data collection and analysis. For example, the date of PhD degree and/or first positions (professor, associate) could be added to the data. We already see indications of such a life cycle in the data of overall network relationships of Figure 4 and in ego-centric relationships of Figure 6. Both of these figures indicate that different authors emerge as dominant over time, that epicenters of activity can be observed and that some indirect paths exist seeding the devolution and emergence of cores of associations.

The social networks do tell us, interestingly, who was writing with whom. It was not surprising to find out that people prefer to write with their colleagues, most prominently with those closest to home. More interestingly, the visualizations reveal the exceptions, collaboration patterns that emerge because of joint activities even over national boundaries. The activities can be independent and unrelated as visualized in Figures 5 and 6. In the whole network graph of

Figure 2, we even see islands, collaborations that are not part of the largest component. What may be the reason for these? We may speculate a local school or department having no contact with others.

The most exciting question that this study begins to answer is what are the key topics of research? We see in Table 9 that keyword terms such as “systemeering” faded out of use after the early years. However, other terms like “information”, “systems” and “development” are still in use today. Overall there are observed a larger corpus of words found in the paper titles and this may reflect a greater diversity of topics in recent years. We are nevertheless left with many open questions. What has changed and is the change reflected in the titles of the papers? Is there a (at least partly) distinct Scandinavian School within IS research? Are there larger groups that have made research in the same areas, methods or approaches, such as collaborations between Gothenburg-Oslo and Copenhagen-Malmö-Lund or VITS network within Sweden? Alas, the keyword-list from papers titles is too limited to draw conclusions on so many questions. Similar methods that were used to incorporate keywords of the title in Commetrix could be used in future research to formulate tables of co-occurrences of citations. That is cited names that appear in the same articles together could be analyzed as keyword terms are in the present study. A preliminary requirement, would be of course to obtain the full text of all the past conference papers.

Finally, one limitation of the current approach to network visualizations is that it is only another form of representation and analytical access to large networks. That is such analysis rarely adds hard statistical facts to our knowledge. As such, visualizations always need to be supported with thorough statistical analysis.

Concluding Remarks

Our analysis points out the importance of a research forum such as the annual IRIS conference. It indicates that a significant network of researcher collaboration has emerged through a very social and productive annual gathering over 29 years. We conclude that IRIS has influenced the research direction in Scandinavia. However, further use of Commetrix and analysis of the IRIS collaborative network is desirable. For example, further visualizations could explore citation relationships in the conference papers, or even the subset of IRIS-authors cited in top journals. While data collection is a daunting task, such an approach would further demonstrate the relatedness and impact of the work begun at IRIS on a larger global community.

References

Barabási, A.L, Jeong, H., Néda, Z., Ravasz, E., Schubert, A. and T. Vissek, (2002). "Evolution of the Social Network of Scientific Collaborations," *Physica A,* 311, 590-614.

Brandes, U. (2001). "A Faster Algorithm for Betweenness Centrality," Journal of Mathematical Sociology 25(2):163-177.

Burt, R.S. (1997). "The Contingent Value of Social Capital," *Administration Science Quarterly*, 42(2):339-365.

DeCastro, R. and J.W. Grossman. (1999). Famous trails to Paul Erdos. *Mathematical Intelligencer*, 21(3):51--63.

Freeman, L.C. (1979). "Centrality in social networks: I. Conceptual clarification." *Social Networks*, 1, 215-239.

Goldkuhl, G., (2006). "Collaborative researching – from ISAC to VITS through Humor", In Bubenko J et al (eds) *ICT for people*. 40 years of academic development in Stockholm, DSV, Stockholm University, pp.1-8.

Grossman, J. W. (2004). *Erdös number*. (Access: http://www.oakland.edu/enp/).

Kerola, P. (1979). "On the extension of the PSC systemeering model" (in Jarvi and Nurminen (eds*.), Report of the Scandinavian research seminar* on systemeering models, 50--54).

Kerola, P. and P. Järvinen, (1979). "The PSC systemeering model" (in Kerola, Klemola, Kamarainen and Lyytinen (eds.), *Summary report of the systemeering research seminar*, 30--38).

Kerola, P. and W. Taggart (1980), Human information processing styles and the information systems architect in PSC-systemeering model (in Lyytinen and Peltola (eds.), *Report of the third Scandinavian research seminar* on systemeering models, 95--125).

Lundeberg, M., Goldkuhl, G. and A. Nilsson (1981). *Information systems development – A systematic approach,* Prentice-Hall, Englewoods Cliffs.

Molka-Danielsen, J. and Søvik, B.L.B., (2005). "Direction of Links in Collaborative Research Networks and Other Structural Characteristics", *Proceedings of the 28th Information Systems Research Seminar in Scandinavia*, eds. Eli Hustad, Bjørn Erik Munkvold, Knut Rolland, and Leif Skiftenes Flak, paper number 32, Kristiansand, Norway, August 6-9. ISBN: 82-7117-554-8.

Newman, M. (2001). "Clustering and Preferential Attachment in Networks,", *Physical Review*, E64, 25102.

Nurminen, M.I., (1997). "Nurminen's Column on IRIS: Part 1," *Scandinavian Journal of Information Systems*, 9(2):47–52.

Taggart, W. (1980). "Human information processing styles and the information systems architect in the PSC systemeering model", *Proceedings of the 17th annual computer personnel research conference*, p.63-78, June 26-27, Miami, Florida, United States.

Trier, M. (2005). "IT-supported Visualization of Knowledge Community Structures", *Proceedings of 38th IEEE Hawaii International Conference of Systems Sciences HICCS38,* p.194-202, January, Big Island, Hawaii, United States.

Wasserman, S., and K. Faust (1994). *Social Network Analysis: Methods and Applications*. Cambridge, England: Cambridge University Press.

Appendix

A description of measures used by Commetrix

Metric (per author)	Commetrix	Wasserman/Faust 1994
Direct Contacts	$DC(n_i) = d_i$ d_i := number of direct contacts of actor i Co-authorships of actor i	$DC(n_i) = d_i$ d_i := number of direct contacts of actor i
Betweenness Centrality	$C'_B(n_i) = \frac{\sum_{j=1}^{g} g_{jk}(n_i)}{\sum_{j=1}^{g} g_{jk}}$ → standardized value of $C_B(n_i)$ There might be more than one shortest path between actors j and k. $g_{jk}(n_i)$:= Boolean value if there is a shortest path between actors j and k over actor i (j and k distinct from i). {0,1} g_{jk}:= Boolean value if there is a shortest path between actors j and k. {0,1} Number of shortest pathes between all pairs of actors j and k (not including actor i) over actor i. Standardized by the number of all shortest pathes in the network. Only one shortest path between	$C'_B(n_i) = \frac{\sum_{j=1}^{g} \frac{g_{jk}(n_i)}{g_{jk}}}{\frac{(g-1)(g-2)}{2}}$ → standardized value of $C_B(n_i)$ There might be more than one shortest path between actors j and k. $\frac{g_{jk}(n_i)}{g_{jk}}$:= probability of shortest path $g_{jk}(n_i)$. All pathes equally likely. If only one shortest path exists value is 1. Number of shortest pathes between all pairs of actors j and k (not including actor i) over actor i. Standardized by the number of pairs of actors not including n_i. All shortest pathes between

	actors j and k is taken into account. Shortest pathes are calculated by the Floyd Warshal Algorithm.	actors j and k are taken into account.
Degree Centrality	$C'_D(n_i) = \frac{d(n_i)}{g-1}$ → standardized value of $C_D(n_i)$ $d(n_i)$:= number of direct contacts of node i g := number of actors → $g-1$:= maximum number of direct contacts Standardized, relative measure of direct contacts	$C'_D(n_i) = \frac{d(n_i)}{g-1}$ → standardized value of $C_D(n_i)$ $d(n_i)$:= number of direct contacts of node i g := number of actors Standardized, relative measure of direct contacts