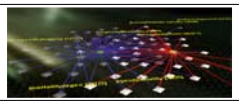
# Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'

Katy Börner & the InfoVis Lab
School of Library and Information Science
INDIANA UNIVERSITY
BLOOMINGTON
katy@indiana.edu

Networks and Complex Systems Talk, IUB, February 28th, 2005.



## 'Mapping Science' Talks

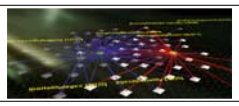
**Mapping Knowledge Domains**
SLIS talk in Jan 2003. Slides/audio are at http://vw.indiana.edu/talks03/

**Modeling the Simultaneous Evolution of Author and Paper Networks**
SLIS talk in Nov 2003. Slides/audio are at http://vw.indiana.edu/talks04/

**Analyzing and Communicating the Structure and Evolution of Science**
Colloquium Talk in Nov 2004, Department of History and Philosophy of Science, IUB. Slides are at http://ella.slis.indiana.edu/~katy/events/

**InfoVis Cyberinfrastructure**
SLIS talk in Nov 2004. Slides are at http://ella.slis.indiana.edu/~katy/events/

**Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'**
Networks and Complex Systems Talk. Slides/video will be at http://vw.indiana.edu/talks-spring05/



## Motivation for Newcomers to Mapping Science

**Mankind's Challenges**

Exploit human knowledge & expertise to

- Reduce pollution and global warming,
- Win the global fight to control disease,
- Provide shelter, food, and clean water for everybody,
- Live together peacefully,
- Etc.

**All these challenges have a very strong correlation with population density.**

Mastering those challenges requires effective access to high quality data(streams), services, compute power, scholarly knowledge & expertise, etc. but also better means 'to allocate our attention efficiently among the overabundance of information sources that might consume it' *(Herbert Simon)* and to 'stand on the shoulders of giants' *(Isaac Newton).*



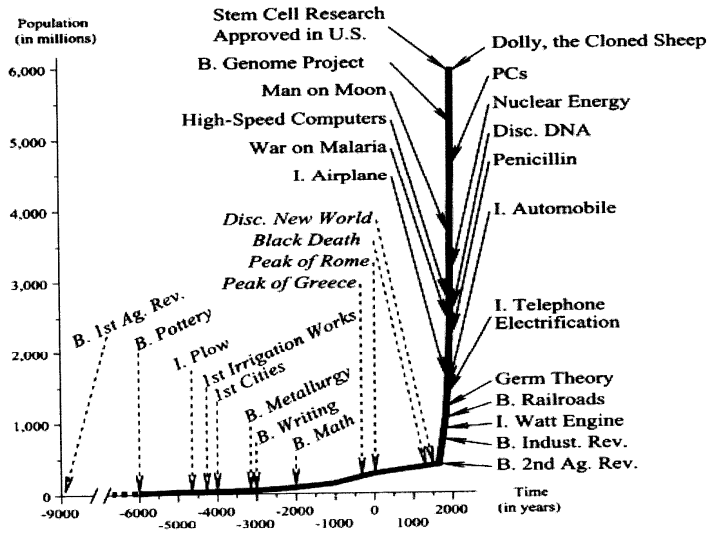
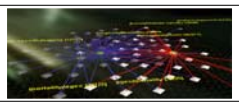



**Figure 2.1 The Growth of World Population and Some Major Events in the History of Technology.**

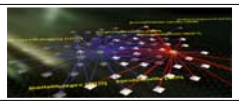
*Sources:* Cipolla 1974; Clark 1961; Fagan 1977; McNeill 1971; Piggott 1965; Derry and Williams 1960; Trewartha 1969. See also Allen 1992, 1994; Slicher van Bath 1963; Wrigley 1987.

*Note:* There is usually a lag between the invention (I) of a process or a machine and its general application to production. "Beginning" (B) usually means the earliest stage of this diffusion process.



The problem is **not** how one person can access knowledge but how we can collectively access and manage humanity's knowledge & expertise and utilize it to solve the diverse challenges.





**Part I: Is a 'Global Brain' emerging?**

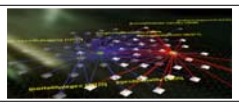
## InfoVis Contest Data Set

## Visualization of Growing Co-Author Networks

Ke, Weimao, Börner, Katy and Viswanath, Lalitha. (2004). Analysis and Visualization of the IV 2004 Contest Dataset. Poster Compendium, IEEE Information Visualization Conference, pp. 49-50, 2004. This entry won first place in the InfoVis 2004 Contest.

## Studying the Emerging Global Brain

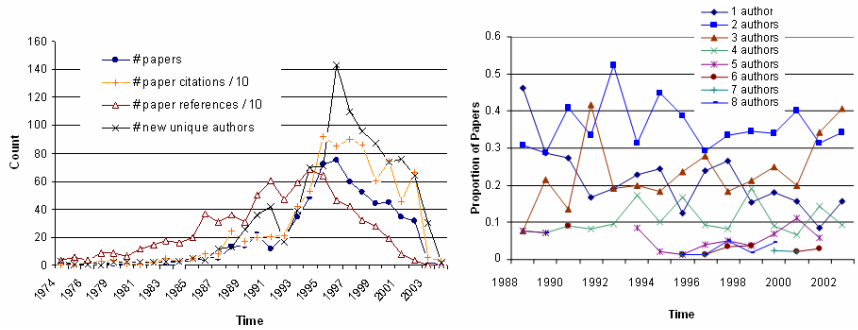
Börner, Katy, Dall'Asta, Luca, Ke, Weimao and Vespignani, Alessandro. (in press) Studying the Emerging Global Brain: Analyzing and Visualizing the Impact of Co-Authorship Teams. *Complexity*, special issue on *Understanding Complex Systems*.



## The InfoVis Contest 2004 dataset

Available at http://www.cs.umd.edu/hcil/iv04contest/

- Was retrieved from the ACM digital library.
- Covers 614 papers published in the area of information visualization by 1,036 unique authors between 1974 and 2004.



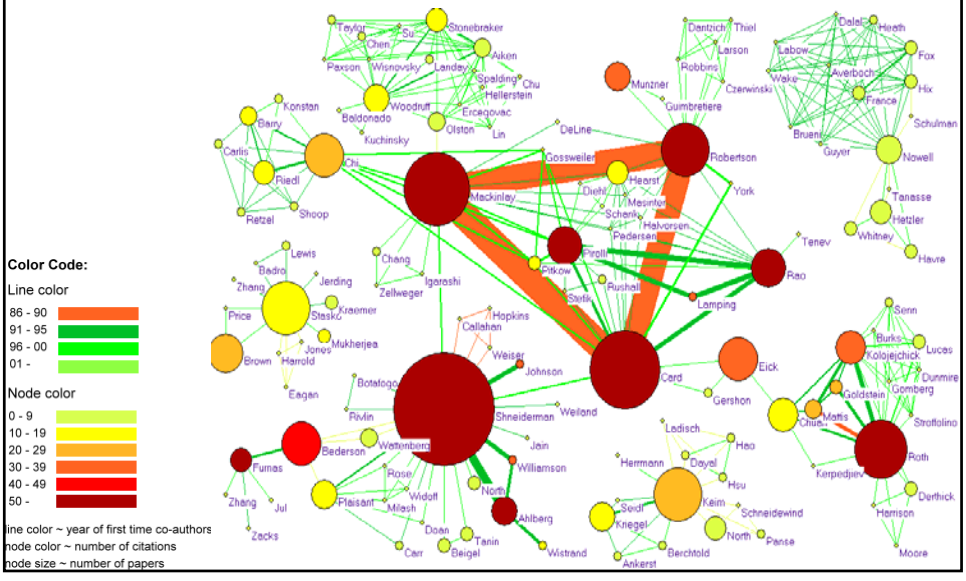

- Is unique in that it documents the birth and growth of InfoVis.

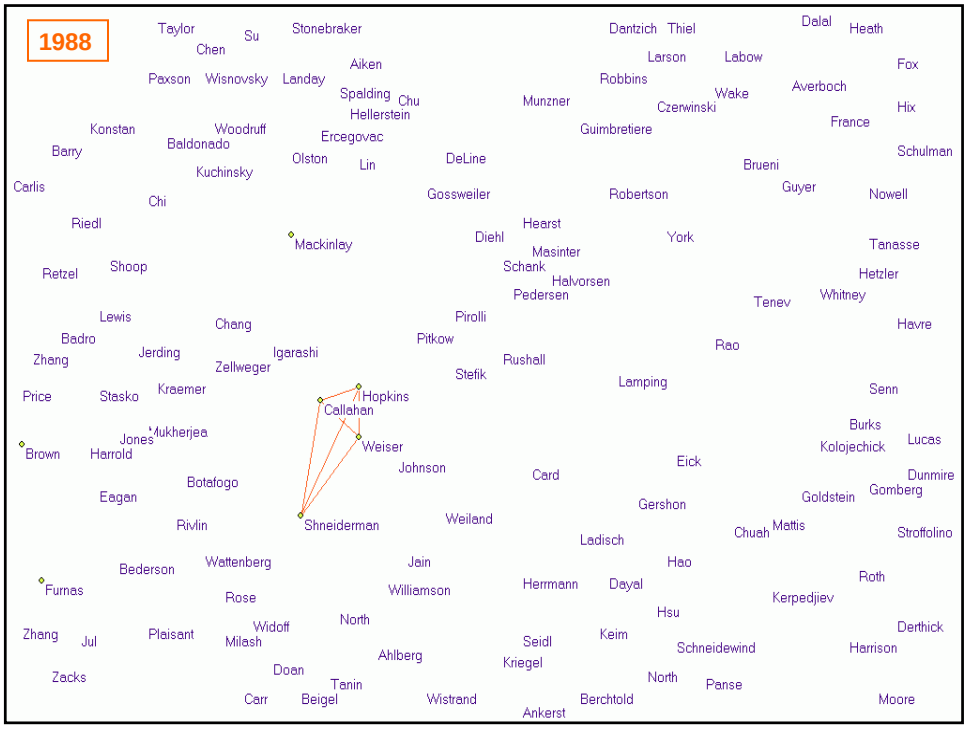
Katy Börner, Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'. February 28th, 2005.

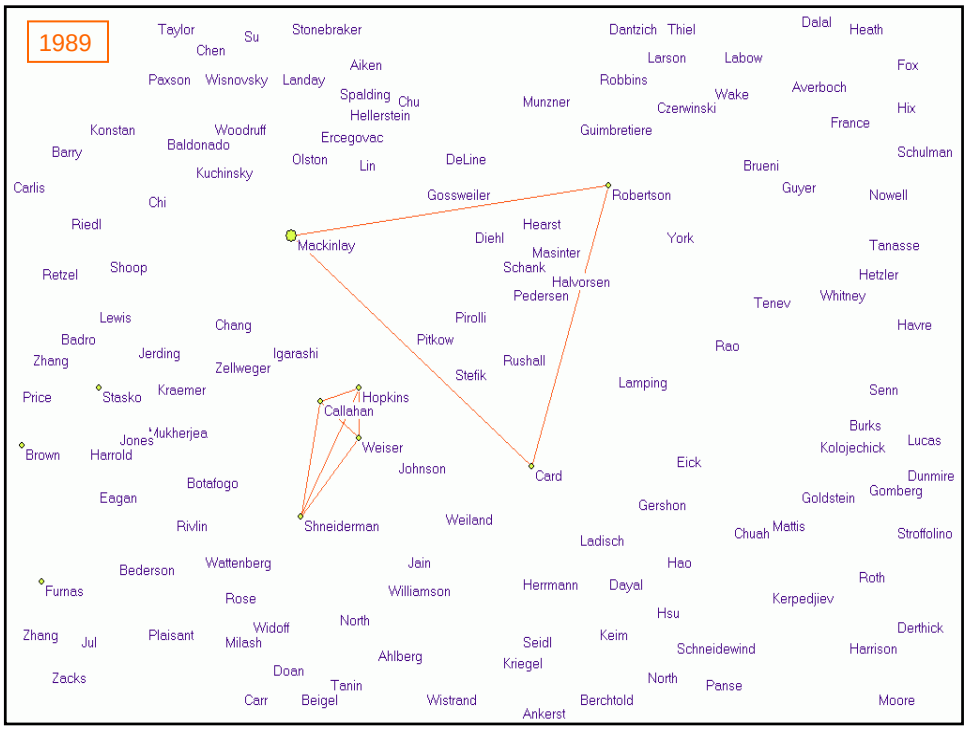
## Visualization of Growing Co-Author Networks

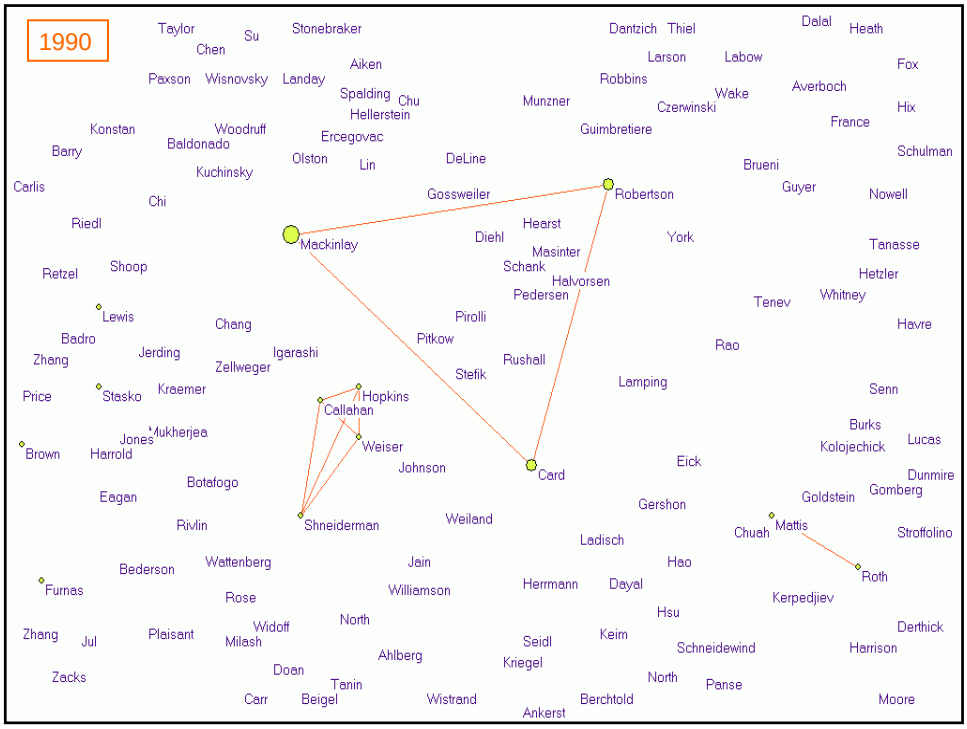
Won 1st price at the IEEE InfoVis Contest

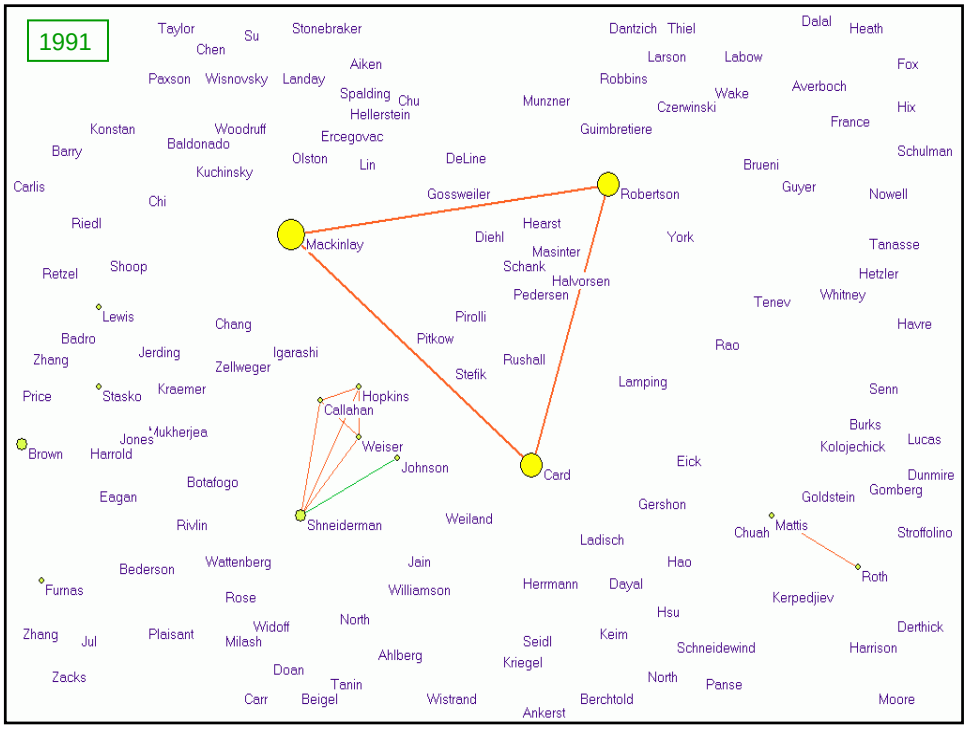
*(Ke, Visvanath & Börner, 2004)*

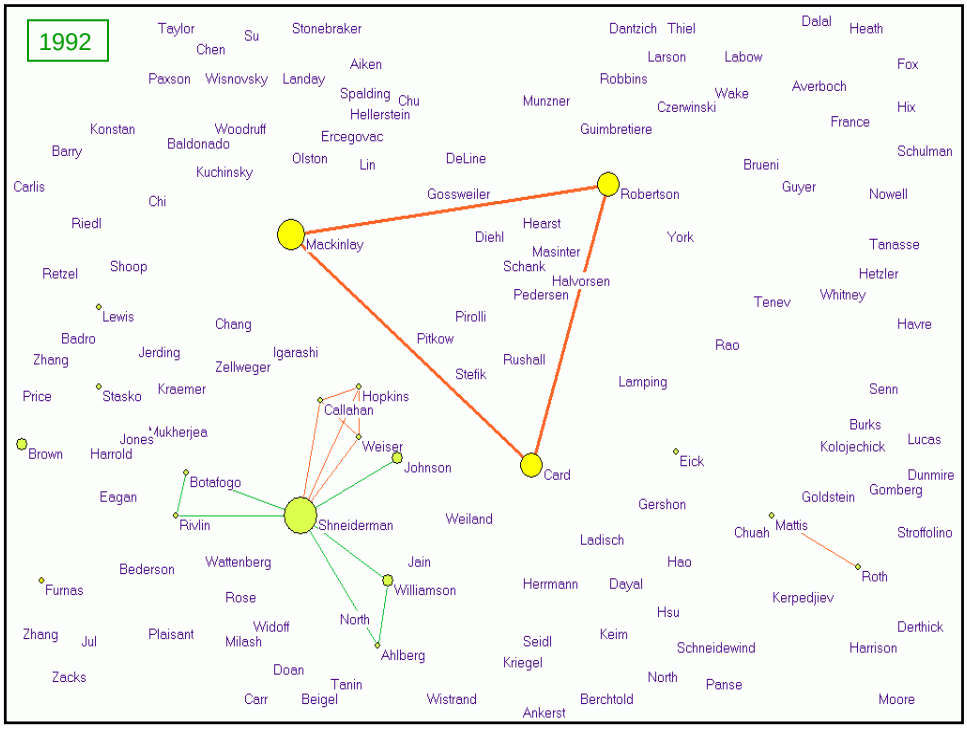


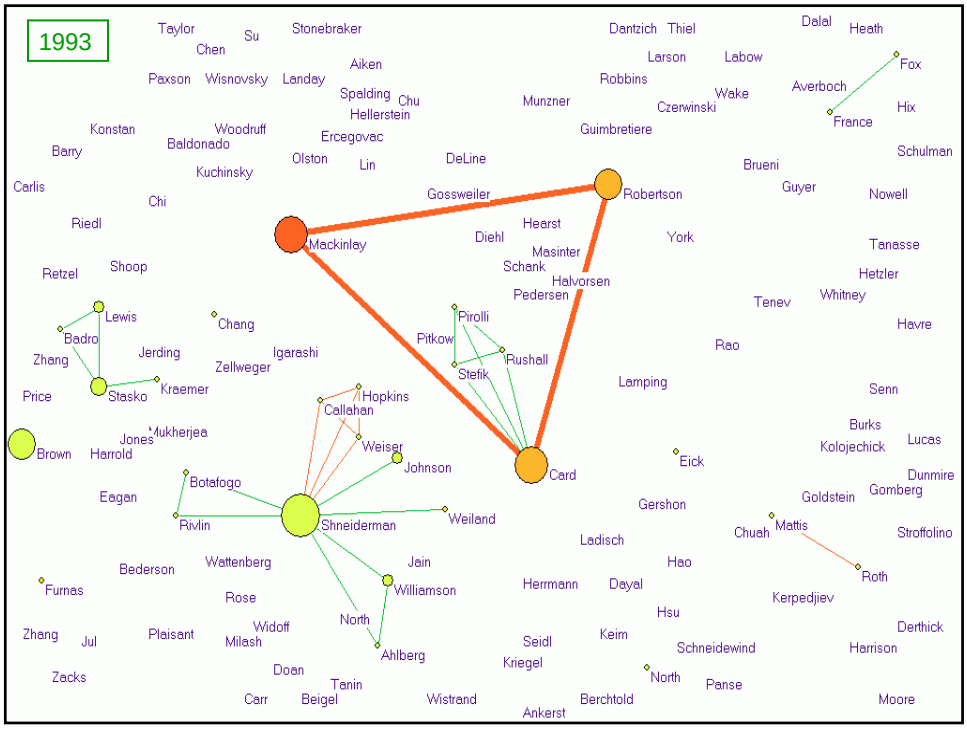

1988
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Konstan
Woodruff
Baldonado
Ercegovac
Barry
Olston
Lin
DeLine
Kuchinsky
Carlis
Chi
Gossweiler
Riedl
Mackinlay
Diehl
Hearst
Masinter
Schank
Retzel
Shoop
Halvorsen
Pedersen
Lewis
Chang
Pirolli
Badro
Pitkow
Zhang
Jerding
Igarashi
Zellweger
Rushall
Stefik
Price
Stasko
Kraemer
Hopkins
Callahan
Weiser
Jones
Mukherjea
Brown
Harrold
Johnson
Botafogo
Eagan
Rivlin
Shneiderman
Weiland
Bederson
Wattenberg
Jain
Furnas
Rose
Williamson
Widoff
North
Zhang
Jul
Plaisant
Milash
Ahlberg
Doan
Zacks
Tanin
Carr
Beigel
Wistrand
Dantzich
Thiel
Dalal
Heath
Larson
Labow
Fox
Robbins
Averboch
Munzner
Wake
Czerwinski
Hix
Guimbretiere
France
Schulman
Brueni
Guyer
Robertson
Nowell
York
Tanasse
Hetzler
Whitney
Tenev
Havre
Rao
Lamping
Senn
Burks
Kolojechick
Lucas
Eick
Dunmire
Gomberg
Gershon
Goldstein
Mattis
Chuah
Stroffolino
Ladisch
Hao
Roth
Herrmann
Dayal
Kerpedjiev
Hsu
Derthick
Keim
Seidl
Schneidewind
Harrison
Kriegel
North
Panse
Berchtold
Moore
Ankerst

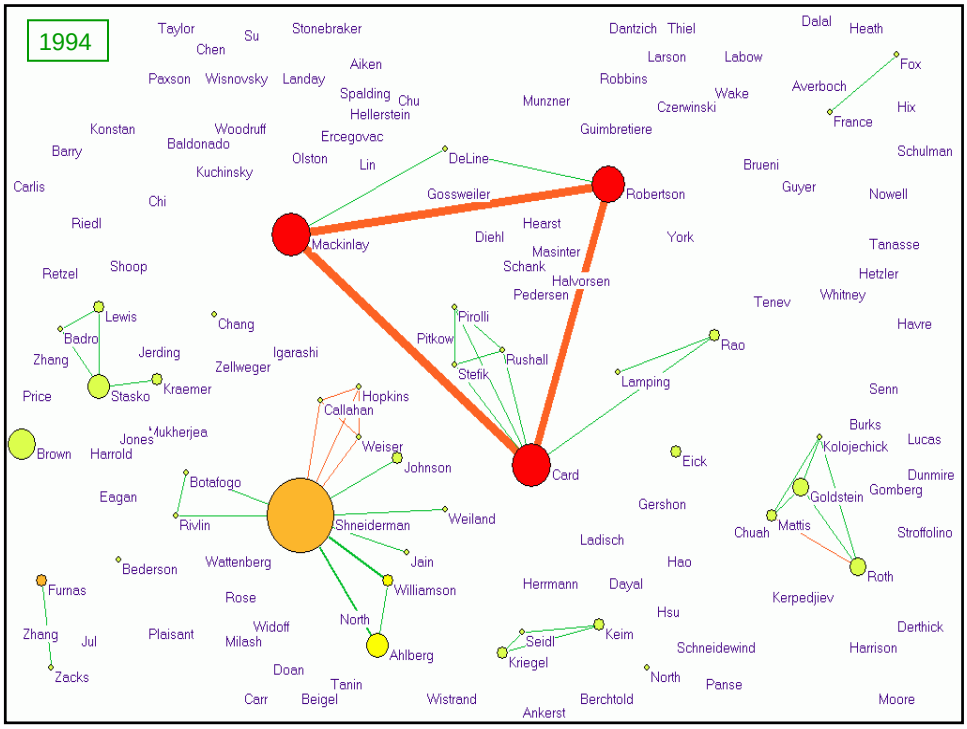

1989
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Konstan
Woodruff
Baldonado
Ercegovac
Barry
Olston
Lin
DeLine
Kuchinsky
Carlis
Chi
Gossweiler
Robertson
Riedl
Mackinlay
Diehl
Hearst
Masinter
Schank
Retzel
Shoop
Halvorsen
Pedersen
Lewis
Chang
Pirolli
Badro
Pitkow
Zhang
Jerding
Igarashi
Zellweger
Rushall
Stefik
Price
Stasko
Kraemer
Hopkins
Callahan
Weiser
Jones
Mukherjea
Brown
Harrold
Johnson
Card
Botafogo
Eagan
Rivlin
Shneiderman
Weiland
Bederson
Wattenberg
Jain
Furnas
Rose
Williamson
Widoff
North
Zhang
Jul
Plaisant
Milash
Ahlberg
Doan
Zacks
Tanin
Carr
Beigel
Wistrand
Dantzich
Thiel
Dalal
Heath
Larson
Labow
Fox
Robbins
Averboch
Munzner
Wake
Czerwinski
Hix
Guimbretiere
France
Schulman
Brueni
Guyer
Nowell
York
Tanasse
Hetzler
Whitney
Tenev
Havre
Rao
Lamping
Senn
Burks
Kolojechick
Lucas
Eick
Dunmire
Gomberg
Gershon
Goldstein
Mattis
Chuah
Stroffolino
Ladisch
Hao
Roth
Herrmann
Dayal
Kerpedjiev
Hsu
Derthick
Keim
Seidl
Schneidewind
Harrison
Kriegel
North
Panse
Berchtold
Moore
Ankerst

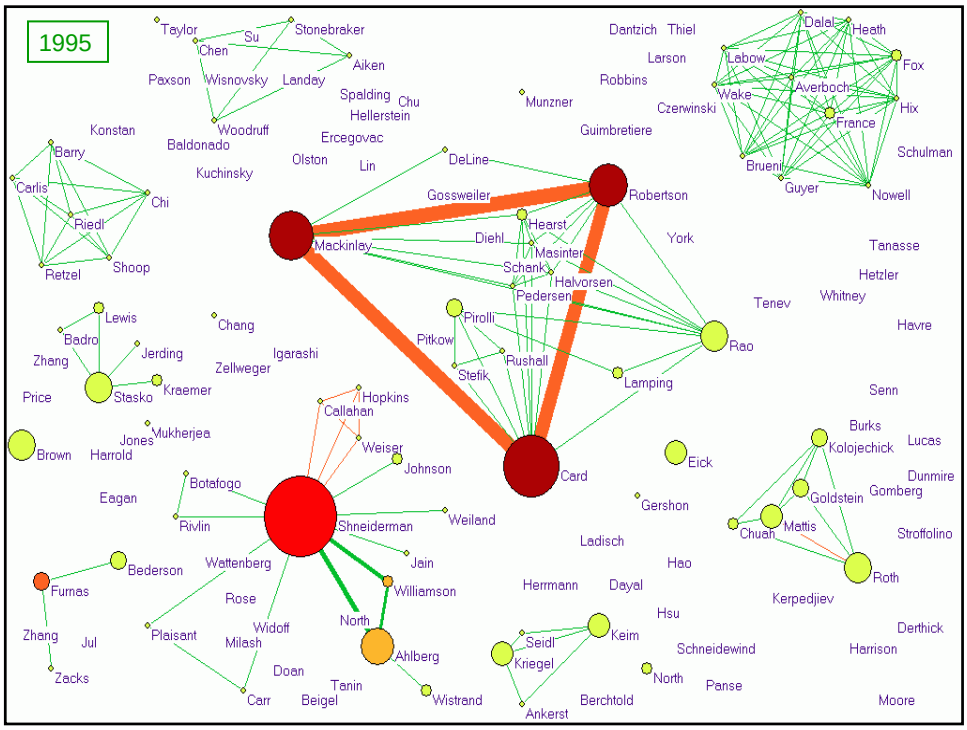

1990
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Konstan
Woodruff
Baldonado
Ercegovac
Barry
Olston
Lin
DeLine
Kuchinsky
Carlis
Chi
Gossweiler
Robertson
Riedl
Mackinlay
Diehl
Hearst
Retzel
Shoop
Schank
Masinter
Halvorsen
Pedersen
Lewis
Chang
Pirolli
Badro
Pitkow
Zhang
Jerding
Igarashi
Zellweger
Rushall
Stefik
Price
Stasko
Kraemer
Hopkins
Callahan
Weiser
Jones
Mukherjea
Brown
Harrold
Johnson
Card
Botafogo
Eagan
Rivlin
Shneiderman
Weiland
Bederson
Wattenberg
Jain
Furnas
Rose
Williamson
Herrmann
Dayal
North
Widoff
Plaisant
Milash
Seidl
Keim
Zacks
Doan
Ahlberg
Kriegel
Tanin
Carr
Beigel
Wistrand
Berchtold
Ankerst
Dantzich
Thiel
Dalal
Heath
Larson
Labow
Fox
Robbins
Averboch
Munzner
Wake
Hix
Czerwinski
Guimbretiere
France
Schulman
Brueni
Guyer
Nowell
York
Tanasse
Hetzler
Whitney
Tenev
Havre
Rao
Lamping
Senn
Burks
Kolojechick
Lucas
Eick
Dunmire
Gomberg
Goldstein
Gershon
Mattis
Chuah
Stroffolino
Ladisch
Hao
Roth
Kerpedjiev
Hsu
Derthick
Schneidewind
Harrison
North
Panse
Moore

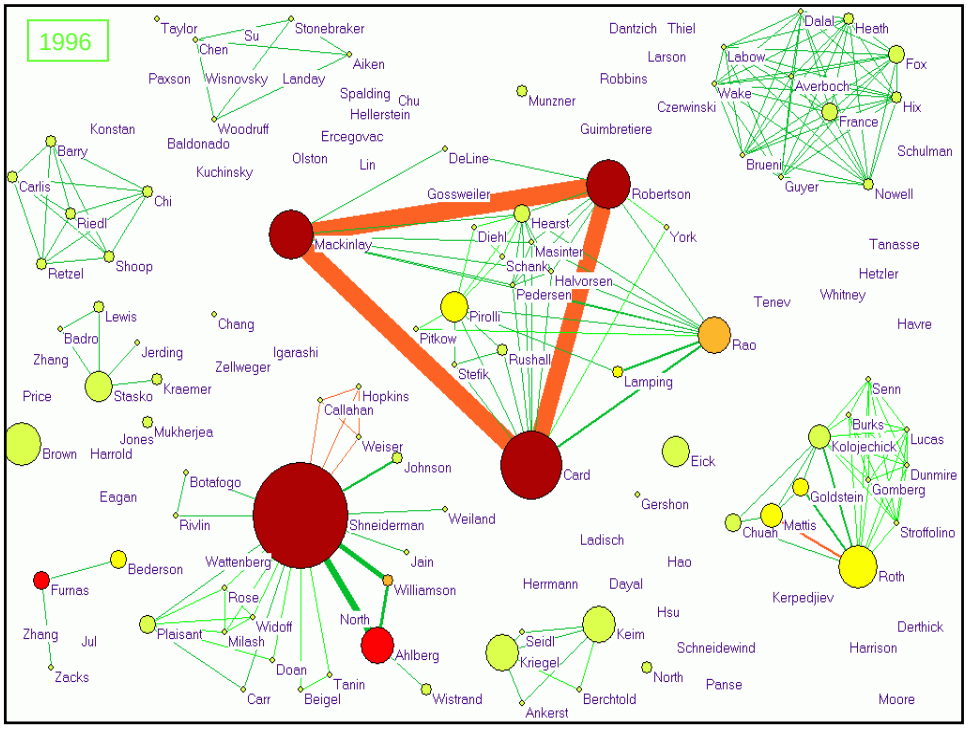

1991
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Konstan
Woodruff
Baldonado
Ercegovac
Barry
Olston
Lin
DeLine
Kuchinsky
Carlis
Chi
Gossweiler
Robertson
Riedl
Mackinlay
Diehl
Hearst
Retzel
Shoop
Schank
Masinter
Halvorsen
Pedersen
Lewis
Chang
Pirolli
Badro
Pitkow
Zhang
Jerding
Igarashi
Zellweger
Rushall
Stefik
Price
Stasko
Kraemer
Hopkins
Callahan
Weiser
Jones
Mukherjea
Brown
Harrold
Johnson
Card
Botafogo
Eagan
Rivlin
Shneiderman
Weiland
Bederson
Wattenberg
Jain
Furnas
Rose
Williamson
Herrmann
Dayal
North
Widoff
Plaisant
Milash
Seidl
Keim
Zacks
Doan
Ahlberg
Kriegel
Tanin
Carr
Beigel
Wistrand
Berchtold
Ankerst
Dantzich
Thiel
Dalal
Heath
Larson
Labow
Fox
Robbins
Averboch
Munzner
Wake
Hix
Czerwinski
Guimbretiere
France
Schulman
Brueni
Guyer
Nowell
York
Tanasse
Hetzler
Whitney
Tenev
Havre
Rao
Lamping
Senn
Burks
Kolojechick
Lucas
Eick
Dunmire
Gomberg
Goldstein
Gershon
Mattis
Chuah
Stroffolino
Ladisch
Hao
Roth
Kerpedjiev
Hsu
Derthick
Schneidewind
Harrison
North
Panse
Moore

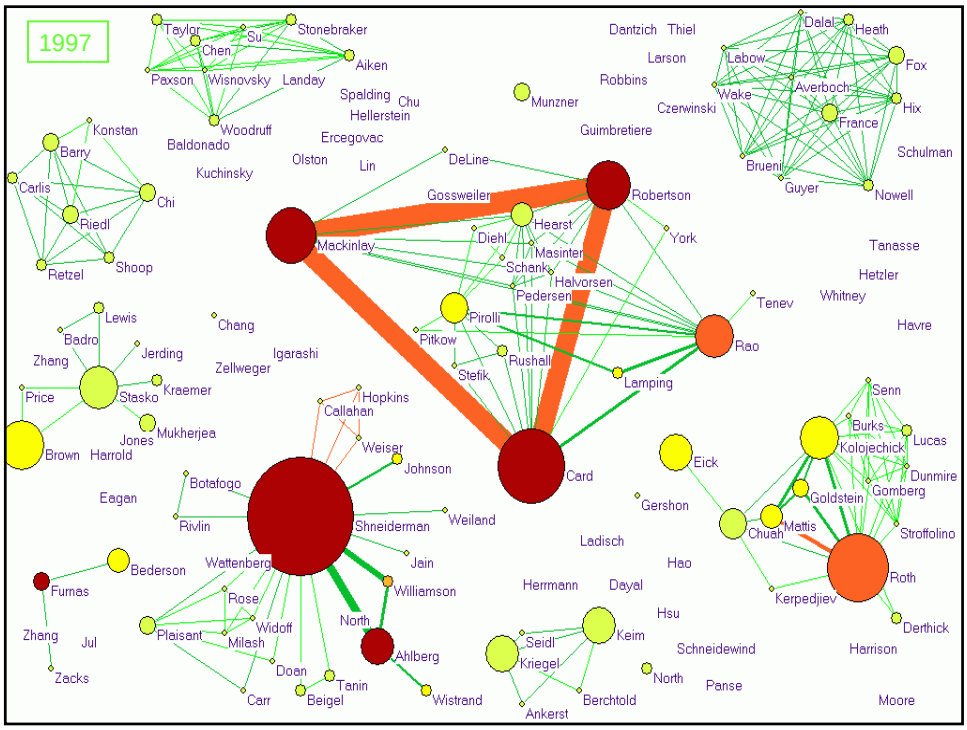

1992
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Konstan
Woodruff
Baldonado
Ercegovac
Barry
Olston
Lin
DeLine
Kuchinsky
Carlis
Chi
Gossweiler
Robertson
Riedl
Mackinlay
Diehl
Hearst
Masinter
Schank
Halvorsen
Pedersen
Retzel
Shoop
Lewis
Chang
Pirolli
Badro
Pitkow
Jerding
Igarashi
Zellweger
Rushall
Zhang
Stefik
Price
Stasko
Kraemer
Hopkins
Callahan
Weiser
Jones
Mukherjea
Brown
Harrold
Johnson
Card
Botafogo
Eagan
Rivlin
Shneiderman
Weiland
Wattenberg
Bederson
Jain
Furnas
Williamson
Rose
North
Widoff
Zhang
Jul
Plaisant
Milash
Ahlberg
Doan
Zacks
Tanin
Carr
Beigel
Wistrand
Dantzich
Thiel
Dalal
Heath
Larson
Labow
Fox
Robbins
Averboch
Munzner
Wake
Czerwinski
Hix
Guimbretiere
France
Schulman
Brueni
Guyer
Nowell
York
Tanasse
Hetzler
Whitney
Tenev
Havre
Rao
Lamping
Senn
Burks
Kolojechick
Lucas
Eick
Dunmire
Gomberg
Goldstein
Gershon
Mattis
Chuah
Stroffolino
Ladisch
Roth
Hao
Herrmann
Dayal
Kerpedjiev
Hsu
Derthick
Keim
Seidl
Schneidewind
Harrison
Kriegel
North
Panse
Berchtold
Ankerst
Moore

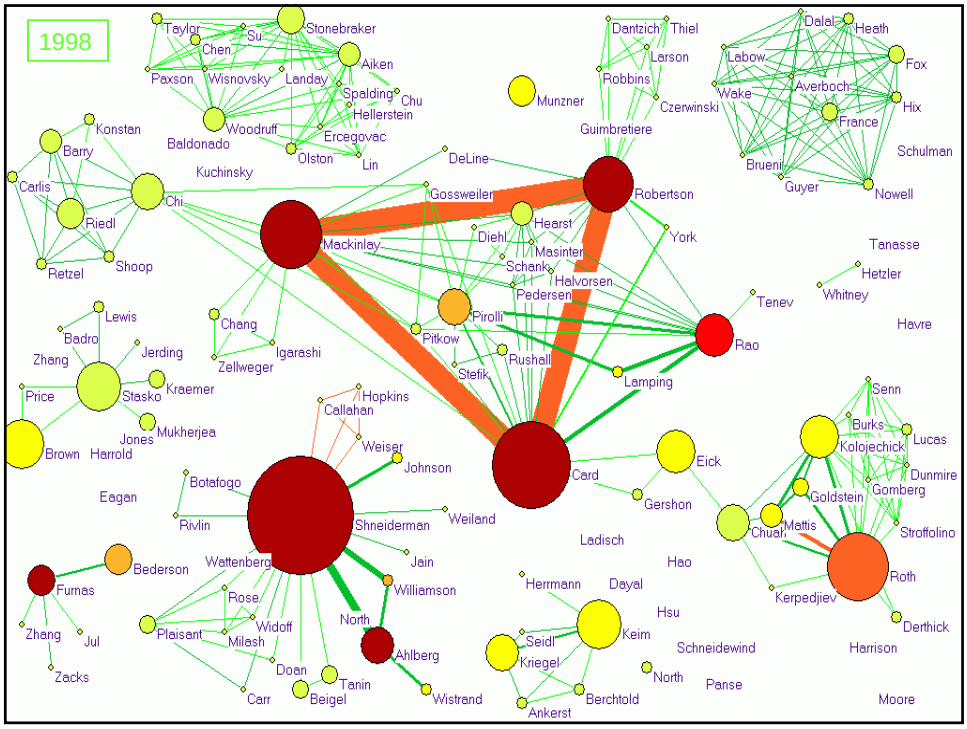

1993
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Konstan
Woodruff
Baldonado
Ercegovac
Barry
Olston
Lin
DeLine
Kuchinsky
Carlis
Chi
Gossweiler
Robertson
Riedl
Mackinlay
Diehl
Hearst
Masinter
Schank
Halvorsen
Pedersen
Retzel
Shoop
Lewis
Chang
Pirolli
Badro
Pitkow
Jerding
Igarashi
Zellweger
Rushall
Zhang
Stefik
Price
Stasko
Kraemer
Hopkins
Callahan
Weiser
Jones
Mukherjea
Brown
Harrold
Johnson
Card
Botafogo
Eagan
Rivlin
Shneiderman
Weiland
Wattenberg
Bederson
Jain
Furnas
Williamson
Rose
North
Widoff
Zhang
Jul
Plaisant
Milash
Ahlberg
Doan
Zacks
Tanin
Carr
Beigel
Wistrand
Dantzich
Thiel
Dalal
Heath
Larson
Labow
Fox
Robbins
Averboch
Munzner
Wake
Czerwinski
Hix
France
Guimbretiere
Schulman
Brueni
Guyer
Nowell
York
Tanasse
Hetzler
Whitney
Tenev
Havre
Rao
Lamping
Senn
Burks
Kolojechick
Lucas
Eick
Dunmire
Gomberg
Goldstein
Gershon
Mattis
Chuah
Stroffolino
Ladisch
Roth
Hao
Herrmann
Dayal
Kerpedjiev
Hsu
Derthick
Keim
Seidl
Schneidewind
Harrison
Kriegel
North
Panse
Berchtold
Ankerst
Moore

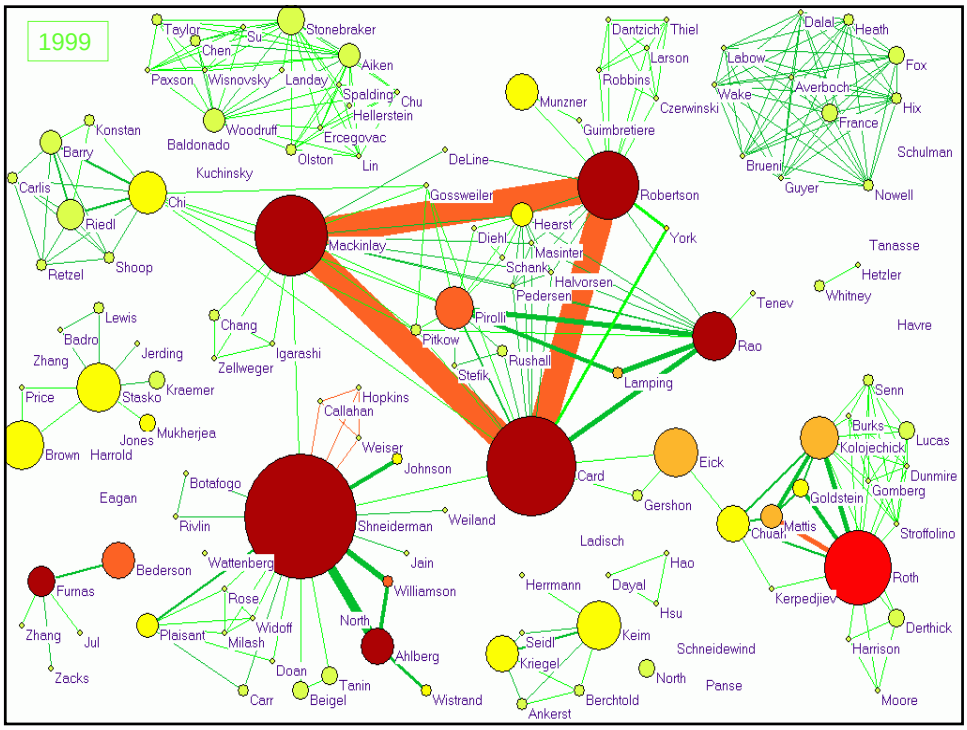

1994
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Konstan
Woodruff
Baldonado
Ercegovac
Barry
Olston
Lin
Kuchinsky
Carlis
Chi
Riedl
Retzel
Shoop
Lewis
Chang
Badro
Jerding
Igarashi
Zhang
Zellweger
Kraemer
Price
Stasko
Hopkins
Callahan
Mukherjea
Jones
Brown
Harrold
Weiser
Johnson
Botafogo
Eagan
Rivlin
Shneiderman
Weiland
Jain
Bederson
Wattenberg
Furnas
Rose
Williamson
Widoff
North
Plaisant
Milash
Ahlberg
Zacks
Doan
Tanin
Carr
Beigel
Wistrand
DeLine
Gossweiler
Mackinlay
Olston
Lin
Robertson
Munzner
Guimbretiere
Dantzich
Thiel
Larson
Labow
Robbins
Czerwinski
Wake
Averboch
Fox
Hix
France
Dalal
Heath
Brueni
Guyer
Nowell
Schulman
Diehl
Hearst
Masinter
Schank
Halvorsen
Pedersen
York
Tanasse
Hetzler
Whitney
Tenev
Havre
Pirolli
Pitkow
Rushall
Stefik
Rao
Lamping
Senn
Card
Eick
Burks
Kolojechick
Lucas
Dunmire
Goldstein
Gomberg
Gershon
Mattis
Chuah
Stroffolino
Ladisch
Roth
Hao
Herrmann
Dayal
Kerpedjiev
Hsu
Seidl
Keim
Derthick
Schneidewind
Harrison
Kriegel
North
Panse
Berchtold
Moore
Ankerst

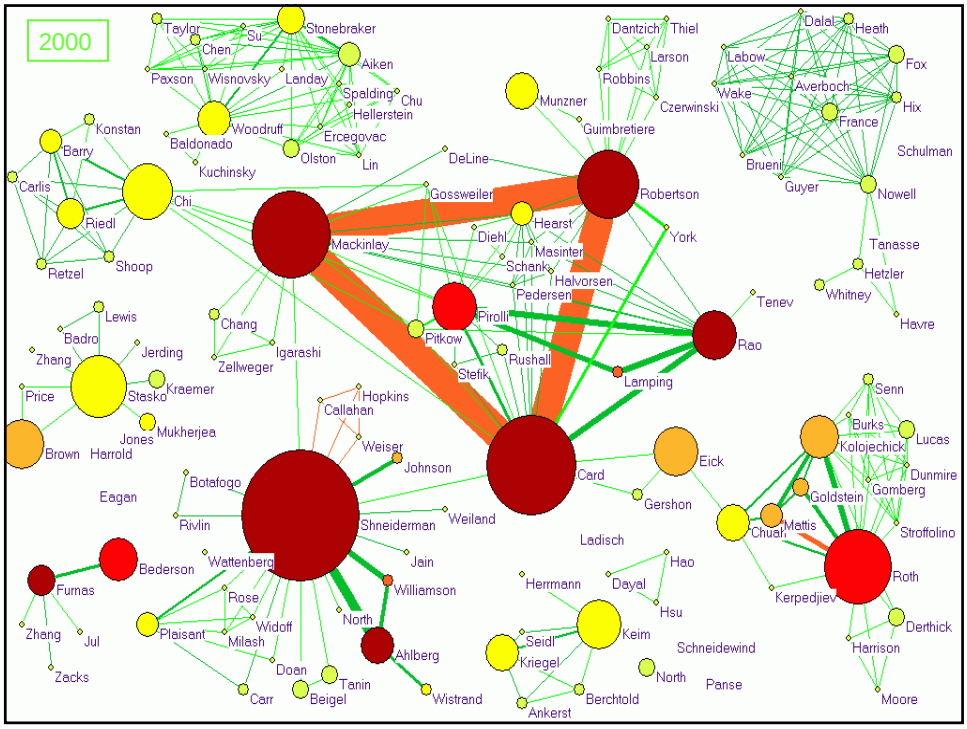

1995
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Konstan
Woodruff
Baldonado
Ercegovac
Barry
Olston
Lin
Kuchinsky
Carlis
Chi
Riedl
Retzel
Shoop
Lewis
Chang
Badro
Jerding
Igarashi
Zhang
Zellweger
Kraemer
Price
Stasko
Hopkins
Callahan
Mukherjea
Jones
Brown
Harrold
Weiser
Johnson
Botafogo
Eagan
Rivlin
Shneiderman
Weiland
Jain
Bederson
Wattenberg
Furnas
Rose
Williamson
Widoff
North
Plaisant
Milash
Ahlberg
Zacks
Doan
Tanin
Carr
Beigel
Wistrand
DeLine
Gossweiler
Mackinlay
Olston
Lin
Robertson
Munzner
Guimbretiere
Dantzich
Thiel
Larson
Labow
Robbins
Czerwinski
Wake
Averboch
Fox
Hix
France
Dalal
Heath
Brueni
Guyer
Nowell
Schulman
Diehl
Hearst
Masinter
Schank
Halvorsen
Pedersen
York
Tanasse
Hetzler
Whitney
Tenev
Havre
Pirolli
Pitkow
Rushall
Stefik
Rao
Lamping
Senn
Card
Eick
Burks
Kolojechick
Lucas
Dunmire
Goldstein
Gomberg
Gershon
Mattis
Chuah
Stroffolino
Ladisch
Roth
Hao
Herrmann
Dayal
Kerpedjiev
Hsu
Seidl
Keim
Derthick
Schneidewind
Harrison
Kriegel
North
Panse
Berchtold
Moore
Ankerst

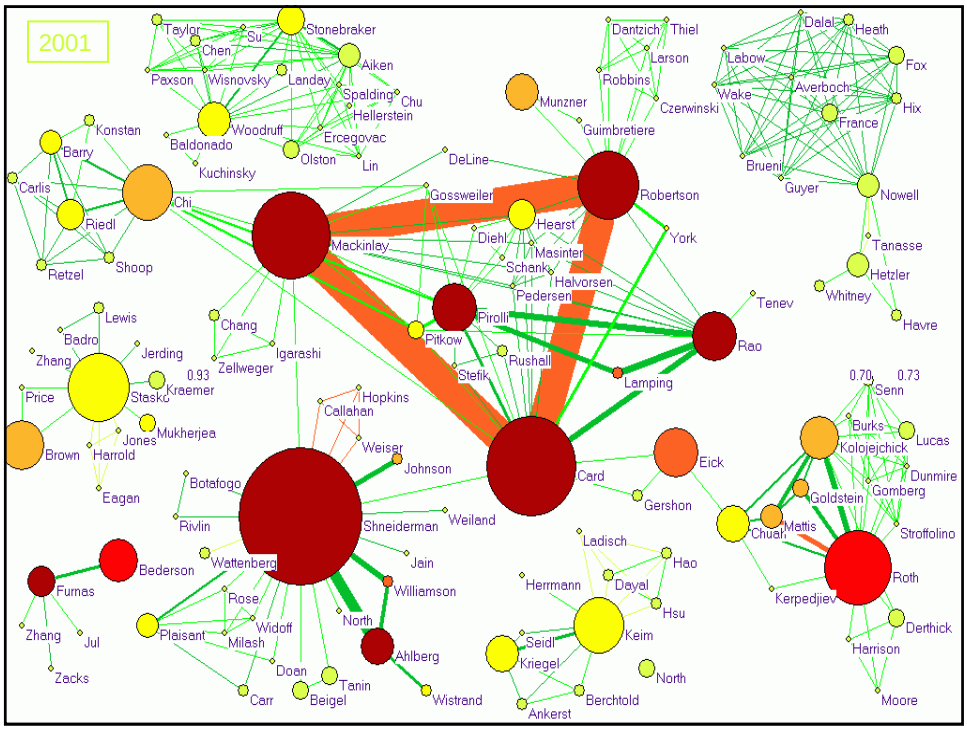

1996
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Woodruff
Konstan
Baldonado
Ercegovac
Barry
Olston
Lin
Kuchinsky
Carlis
Chi
Riedl
Retzel
Shoop
Dantzich
Thiel
Larson
Robbins
Munzner
Czerwinski
Guimbretiere
Dalal
Heath
Labow
Fox
Wake
Averboch
Hix
France
Bruen
Guyer
Nowell
Schulman
DeLine
Gossweiler
Robertson
Hearst
Diehl
York
Mackinlay
Masinter
Schank
Halvorsen
Pedersen
Tanasse
Hetzler
Tenev
Whitney
Havre
Pirolli
Pitkow
Rao
Rushall
Stefik
Lamping
Lewis
Chang
Badro
Jerding
Zhang
Igarashi
Zellweger
Kraemer
Price
Stasko
Hopkins
Callahan
Jones
Mukherjea
Brown
Harrold
Weiser
Johnson
Card
Eick
Botafogo
Eagan
Rivlin
Shneiderman
Weiland
Gershon
Senn
Burks
Lucas
Kolojechick
Dunmire
Goldstein
Gomberg
Chuah
Mattis
Stroffolino
Roth
Kerpedjiev
Ladisch
Hao
Herrmann
Dayal
Hsu
Derthick
Schneidewind
Harrison
Panse
Moore
Wattenberg
Jain
Bederson
Furnas
Rose
Williamson
Widoff
Plaisant
Milash
North
Ahlberg
Zacks
Jul
Doan
Tanin
Carr
Beigel
Wistrand
Seidl
Keim
Kriegel
Berchtold
Ankerst

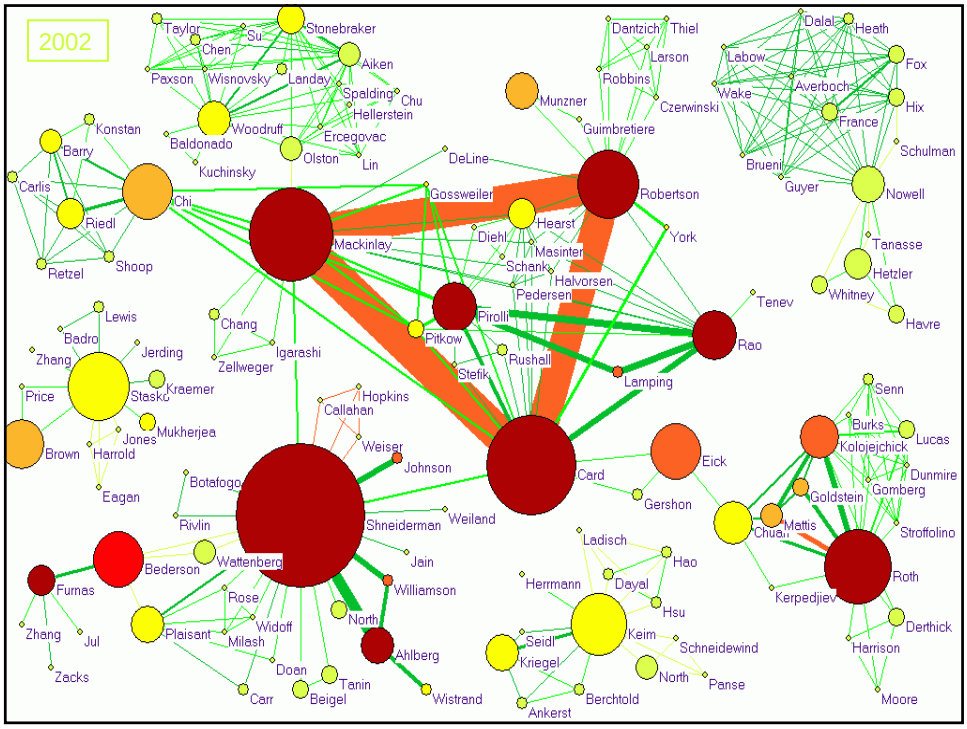

1997
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Woodruff
Konstan
Baldonado
Ercegovac
Barry
Olston
Lin
Kuchinsky
Carlis
Chi
Riedl
Retzel
Shoop
Dantzich
Thiel
Larson
Robbins
Munzner
Czerwinski
Guimbretiere
Dalal
Heath
Labow
Fox
Wake
Averboch
Hix
France
Bruen
Guyer
Nowell
Schulman
DeLine
Gossweiler
Robertson
Hearst
Diehl
York
Mackinlay
Masinter
Schank
Halvorsen
Pedersen
Tanasse
Hetzler
Tenev
Whitney
Havre
Pirolli
Pitkow
Rao
Rushall
Stefik
Lamping
Lewis
Chang
Badro
Jerding
Zhang
Igarashi
Zellweger
Kraemer
Price
Stasko
Hopkins
Callahan
Jones
Mukherjea
Brown
Harrold
Weiser
Johnson
Card
Eick
Botafogo
Eagan
Rivlin
Shneiderman
Weiland
Gershon
Senn
Burks
Lucas
Kolojechick
Dunmire
Goldstein
Gomberg
Chuah
Mattis
Stroffolino
Roth
Kerpedjiev
Ladisch
Hao
Herrmann
Dayal
Hsu
Derthick
Schneidewind
Harrison
Panse
Moore
Wattenberg
Jain
Bederson
Furnas
Rose
Williamson
Widoff
Plaisant
Milash
North
Ahlberg
Zacks
Jul
Doan
Tanin
Carr
Beigel
Wistrand
Seidl
Keim
Kriegel
Berchtold
Ankerst

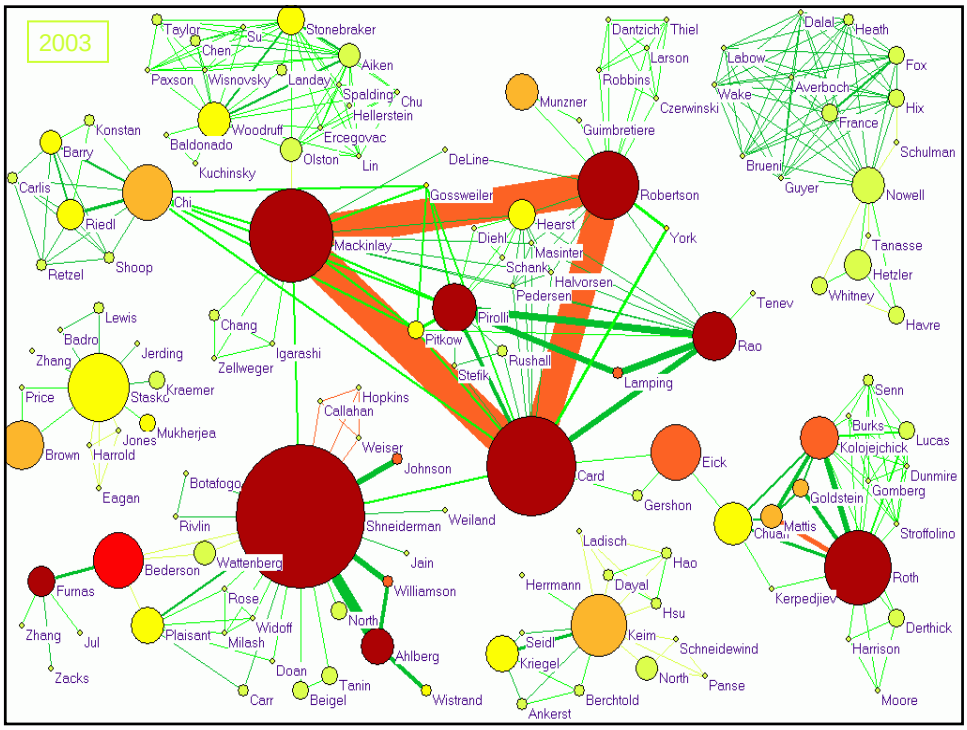

1998
Taylor
Stonebraker
Su
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Woodruff
Ercegovac
Baldonado
Olston
Lin
Konstan
Barry
Kuchinsky
DeLine
Carlis
Chi
Gossweiler
Riedl
Mackinlay
Shoop
Retzel
Munzner
Dantzich
Thiel
Larson
Robbins
Czerwinski
Guimbretiere
Robertson
Hearst
Diehl
York
Masinter
Schank
Halvorsen
Pedersen
Pirolli
Pitkow
Rushall
Stefik
Lamping
Rao
Tenev
Dalal
Heath
Labow
Fox
Wake
Averboch
Hix
France
Schulman
Bruen
Guyer
Nowell
Tanasse
Hetzler
Whitney
Havre
Lewis
Badro
Jerding
Zhang
Chang
Igarashi
Zellweger
Price
Stasko
Kraemer
Jones
Mukherjea
Brown
Harrold
Hopkins
Callahan
Weiser
Johnson
Eagan
Botafogo
Rivlin
Shneiderman
Weiland
Wattenberg
Jain
Bederson
Furnas
Rose
Williamson
Plaisant
Widoff
Milash
North
Ahlberg
Zacks
Jul
Doan
Tanin
Carr
Beigel
Wistrand
Card
Eick
Gershon
Ladisch
Hao
Herrmann
Dayal
Hsu
Seidl
Keim
Kriegel
Schneidewind
North
Panse
Berchtold
Ankerst
Senn
Burks
Lucas
Kolojechick
Dunmire
Goldstein
Gomberg
Chuah
Mattis
Stroffolino
Roth
Kerpedjiev
Derthick
Harrison
Moore

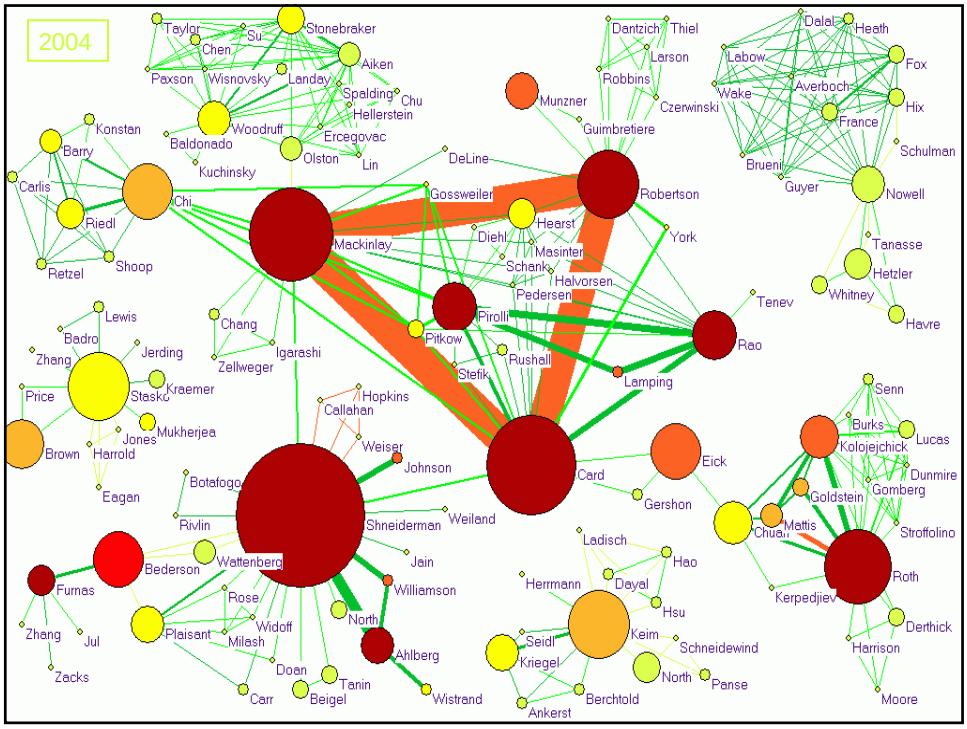

1999
Taylor
Stonebraker
Su
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Woodruff
Ercegovac
Baldonado
Olston
Lin
Konstan
Barry
Kuchinsky
DeLine
Carlis
Chi
Gossweiler
Riedl
Mackinlay
Shoop
Retzel
Munzner
Dantzich
Thiel
Larson
Robbins
Czerwinski
Guimbretiere
Robertson
Hearst
Diehl
York
Masinter
Schank
Halvorsen
Pedersen
Pirolli
Pitkow
Rushall
Stefik
Lamping
Rao
Tenev
Dalal
Heath
Labow
Fox
Wake
Averboch
Hix
France
Schulman
Bruen
Guyer
Nowell
Tanasse
Hetzler
Whitney
Havre
Lewis
Badro
Jerding
Zhang
Chang
Igarashi
Zellweger
Price
Stasko
Kraemer
Jones
Mukherjea
Brown
Harrold
Hopkins
Callahan
Weiser
Johnson
Eagan
Botafogo
Rivlin
Shneiderman
Weiland
Wattenberg
Jain
Bederson
Furnas
Rose
Williamson
Plaisant
Widoff
Milash
North
Ahlberg
Zacks
Jul
Doan
Tanin
Carr
Beigel
Wistrand
Card
Eick
Gershon
Ladisch
Hao
Herrmann
Dayal
Hsu
Seidl
Keim
Kriegel
Schneidewind
North
Panse
Berchtold
Ankerst
Senn
Burks
Lucas
Kolojechick
Dunmire
Goldstein
Gomberg
Chuah
Mattis
Stroffolino
Roth
Kerpedjiev
Derthick
Harrison
Moore

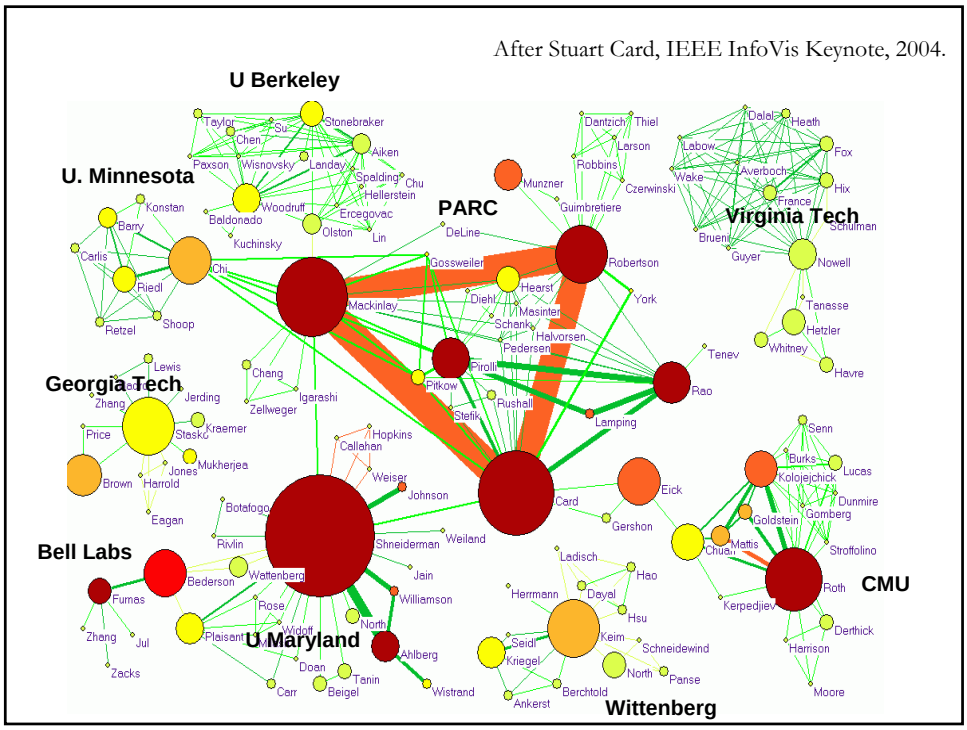

2000
Taylor
Stonebraker
Su
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Woodruff
Ercegovac
Baldonado
Olston
Lin
Kuchinsky
Konstan
Barry
Carlis
Chi
Riedl
Retzel
Shoop
Lewis
Badro
Zhang
Jerding
Kraemer
Price
Stasko
Jones
Mukherjea
Brown
Harrold
Eagan
Chang
Zellweger
Igarashi
Hopkins
Callahan
Weiser
Johnson
Botafogo
Rivlin
Shneiderman
Weiland
Jain
Wattenberg
Bederson
Furnas
Rose
Plaisant
Widoff
Milash
North
Williamson
Ahlberg
Zhang
Jul
Zacks
Doan
Carr
Beigel
Tanin
Wistrand
DeLine
Gossweiler
Mackinlay
Munzner
Guimbretiere
Dantzich
Thiel
Larson
Robbins
Czerwinski
Robertson
Hearst
Diehl
Masinter
Schank
Halvorsen
Pedersen
York
Pirolli
Pitkow
Rushall
Stefik
Lamping
Rao
Tenev
Dalal
Heath
Labow
Fox
Wake
Averboch
Hix
France
Bruen
Guyer
Schulman
Nowell
Tanasse
Hetzler
Whitney
Havre
Senn
Burks
Lucas
Kolojejchick
Dunmire
Goldstein
Gomberg
Mattis
Chuah
Stroffolino
Roth
Kerpedjiev
Derthick
Harrison
Moore
Eick
Card
Gershon
Ladisch
Hao
Herrmann
Dayal
Hsu
Keim
Seidl
Kriegel
Schneidewind
North
Berchtold
Panse
Ankerst

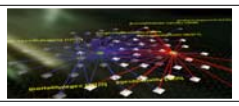

2001
Taylor
Stonebraker
Su
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Woodruff
Ercegovac
Baldonado
Olston
Lin
Kuchinsky
Konstan
Barry
Carlis
Chi
Riedl
Retzel
Shoop
Lewis
Badro
Zhang
Jerding
0.93
Kraemer
Price
Stasko
Jones
Mukherjea
Brown
Harrold
Eagan
Chang
Zellweger
Igarashi
Hopkins
Callahan
Weiser
Johnson
Botafogo
Rivlin
Shneiderman
Weiland
Jain
Wattenberg
Bederson
Furnas
Rose
Plaisant
Widoff
Milash
North
Williamson
Ahlberg
Zhang
Jul
Zacks
Doan
Carr
Beigel
Tanin
Wistrand
DeLine
Gossweiler
Mackinlay
Munzner
Guimbretiere
Dantzich
Thiel
Larson
Robbins
Czerwinski
Robertson
Hearst
Diehl
Masinter
Schank
Halvorsen
Pedersen
York
Pirolli
Pitkow
Rushall
Stefik
Lamping
Rao
Tenev
Dalal
Heath
Labow
Fox
Wake
Averboch
Hix
France
Bruen
Guyer
Nowell
Tanasse
Hetzler
Whitney
Havre
0.70
0.73
Senn
Burks
Lucas
Kolojejchick
Dunmire
Goldstein
Gomberg
Mattis
Chuah
Stroffolino
Roth
Kerpedjiev
Derthick
Harrison
Moore
Eick
Card
Gershon
Ladisch
Hao
Herrmann
Dayal
Hsu
Keim
Seidl
Kriegel
North
Berchtold
Ankerst


2002
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Woodruff
Baldonado
Ercegovac
Olston
Lin
Kuchinsky
Konstan
Barry
Carlis
Riedl
Chi
Retzel
Shoop
DeLine
Gossweiler
Mackinlay
Munzner
Guimbretiere
Dantzich
Thiel
Larson
Robbins
Czerwinski
Robertson
Hearst
Diehl
Masinter
Schank
Halvorsen
Pedersen
York
Pirolli
Pitkow
Rushall
Stefik
Lamping
Rao
Tenev
Dalal
Heath
Labow
Fox
Wake
Averboch
Hix
France
Bruenl
Guyer
Schulman
Nowell
Tanasse
Hetzler
Whitney
Havre
Lewis
Badro
Zhang
Jerding
Zellweger
Chang
Igarashi
Price
Stasko
Kraemer
Jones
Mukherjea
Brown
Harrold
Eagan
Hopkins
Callahan
Weiser
Johnson
Botafogo
Rivlin
Shneiderman
Weiland
Wattenberg
Bederson
Furnas
Zhang
Jul
Zacks
Rose
Plaisant
Widoff
Milash
Jain
Williamson
North
Ahlberg
Doan
Carr
Beigel
Tanin
Wistrand
Card
Eick
Gershon
Ladisch
Hao
Herrmann
Dayal
Hsu
Keim
Schneidewind
Seidl
Kriegel
North
Panse
Berchtold
Ankerst
Senn
Burks
Lucas
Kolojejchick
Dunmire
Goldstein
Gomberg
Chuah
Mattis
Stroffolino
Roth
Kerpedjiev
Derthick
Harrison
Moore


2003
Taylor
Su
Stonebraker
Chen
Aiken
Paxson
Wisnovsky
Landay
Spalding
Chu
Hellerstein
Woodruff
Baldonado
Ercegovac
Olston
Lin
Kuchinsky
Konstan
Barry
Carlis
Riedl
Chi
Retzel
Shoop
DeLine
Gossweiler
Mackinlay
Munzner
Guimbretiere
Dantzich
Thiel
Larson
Robbins
Czerwinski
Robertson
Hearst
Diehl
Masinter
Schank
Halvorsen
Pedersen
York
Pirolli
Pitkow
Rushall
Stefik
Lamping
Rao
Tenev
Dalal
Heath
Labow
Fox
Wake
Averboch
Hix
France
Bruenl
Guyer
Schulman
Nowell
Tanasse
Hetzler
Whitney
Havre
Lewis
Badro
Zhang
Jerding
Zellweger
Chang
Igarashi
Price
Stasko
Kraemer
Jones
Mukherjea
Brown
Harrold
Eagan
Hopkins
Callahan
Weiser
Johnson
Botafogo
Rivlin
Shneiderman
Weiland
Wattenberg
Bederson
Furnas
Zhang
Jul
Zacks
Rose
Plaisant
Widoff
Milash
Jain
Williamson
North
Ahlberg
Doan
Carr
Beigel
Tanin
Wistrand
Card
Eick
Gershon
Ladisch
Hao
Herrmann
Dayal
Hsu
Keim
Schneidewind
Seidl
Kriegel
North
Panse
Berchtold
Ankerst
Senn
Burks
Lucas
Kolojejchick
Dunmire
Goldstein
Gomberg
Chuah
Mattis
Stroffolino
Roth
Kerpedjiev
Derthick
Harrison
Moore


2004


After Stuart Card, IEEE InfoVis Keynote, 2004.
U Berkeley
U. Minnesota
PARC
Virginia Tech
Georgia Tech
Bell Labs
U Maryland
CMU
Wittenberg



## Studying the Emerging Global Brain

**Research question:**

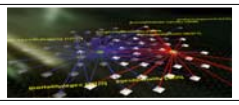
Is science driven by prolific single experts or by high-impact co-authorship teams?

**Contributions of this study:**

1. New approach to allocate citational credit.
2. Novel weighted graph representation.
3. Visualization of the growth of weighted co-author network.
4. Centrality measures to identify author impact.
5. Global statistical analysis of paper production and citations in correlation with co-authorship team size over time.
6. Local, author-centered entropy measure.

Börner, Katy, Dall'Asta, Luca, Ke, Weimao and Vespignani, Alessandro. (in press) Studying the Emerging Global Brain: Analyzing and Visualizing the Impact of Co-Authorship Teams. *Complexity*, special issue on *Understanding Complex Systems*.





## 1) Allocation of citational credit

**Allocation of citational credit**

Common approaches used in information science (Blaise, 1994):

- Only the first author receives credit,
- Each author receives full credit, and
- Fractions are awarded to each co-author.

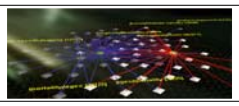
We award citational credit to co-author relations so that the collective success of co-authorship teams – as opposed to the success of single authors – can be studied.

**Weighted co-authorship networks**

Prior work by M. Newman (2004) focused on an evaluation of the strength of the connection in terms of the continuity and time share of a collaboration.

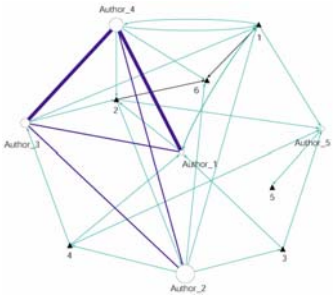
Our focus is on the productivity (number of papers) and the impact (number of papers and citations) of co-authorship teams.

**Note:**
**The number of papers and their citations are only a partial indicator of the scientific impact and influence of a collaboration.**



## 2) Representing author-paper networks as weighted graphs

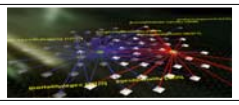
Author-paper networks are tightly coupled and cannot be studied in isolation. We suggest to project important features of one network (e.g., the number of papers produced by a co-author team or the number of citations received by a paper) onto a second network (e.g., the network of co-authors that produced the set of papers).




**Assumptions:**

- The existence of a paper *p* is denoted with a unitary weight of 1, representing the production of the paper itself. (This way, papers that do not receive any citations do not completely disappear from the network.)
- The *impact* of a paper grows linearly with the number of citations $c_p$ the paper receives.
- Single author papers do not contribute to the co-authorship network weight or topology.
- The impact generated by a paper is equally shared among all co-authors.

Katy Börner, Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'. February 28th, 2005.



## Defining the 'impact' weight of a co-authorship edge

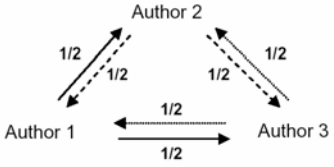
*The impact weight* of a co-authorship edge equals the sum of the *normalized impact* of the paper(s) that resulted from the co-authorship. Formally, the *impact weight* $w_{ij}$ associated with an edge *(i,j)* is defined as

$$w_{ij} = \sum_p \frac{(1+c_p)}{n_p(n_p-1)},$$

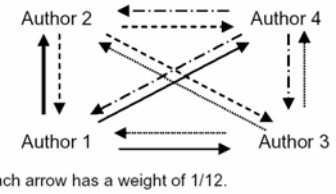
were index *p* runs over all papers co-authored by the authors *i* and *j*, and $n_p$ is the number of authors and $c_p$ the number of citations of paper *p*, respectively. The normalization factor $n_p(n_p-1)$ ensures that the sum over all the edge weights per author equals the number of citations divided by the number of authors.

**Exemplification of the impact weight definition:**

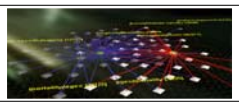
Weights added by a paper with three authors and two citations: $1+c_p = 3$, $n_p(n_p-1) = 6$, $w_{ij} = 1/2$




Weights added by a paper with four authors and no citations: $1+c_p = 1$, $n_p(n_p-1) = 12$, $w_{ij} = 1/12$




Each arrow has a weight of 1/12.

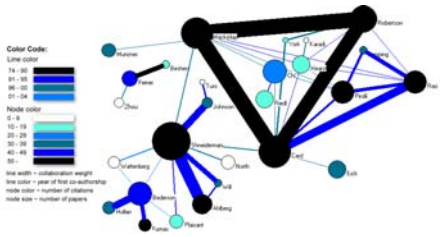


## 3) Visualization of network evolution

To see structure and dynamics of co-authorship relations

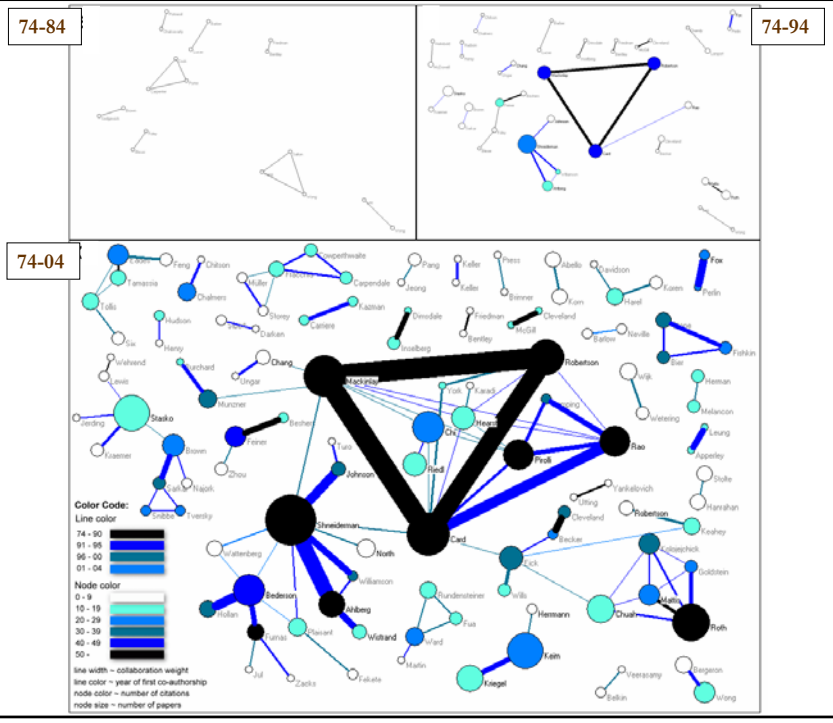
**Visual Encoding**

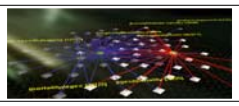
- Nodes represent authors
- Edges denote co-authorship relations
- Node area size reflects the number of single-author and co-authored papers published in the respective time period.
- Node color indicates the cumulative number of citations received by an author.
- Edge color reflects the year in which the co-authorship was started.
- Edge width corresponds to the impact weight.



**Exemplification of the visualization of the evolution of the InfoVis Contest** weighted co-author networks for three different time slices: 1974-1984, 1974-1994 and 1974-2004 laid out using Kamada and Kawai layout algorithm in Pajek.

Katy Börner, Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'. February 28th, 2005.





## 4) Measures to identify author impact

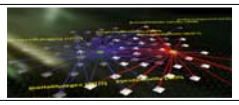
**Degree k:** equals the number of edges attached to the node.
e.g., number of unique co-authors an author has acquired.

**Citation Strength $S_c$** of a node *i* is defined as $s_c(i) = \sum_j w_{ij}$
e.g., number of papers an author team produced and the citations these papers attracted.

**Productivity Strength $S_p$** of a node i is defined as $s_p(i) = s_c(i)|_{c_p=0}$
e.g., number of papers an author team produced.

**Betweenness** of a node *i*, is defined to be the fraction of shortest paths between pairs of nodes in the network that pass through *i*.
e.g., the extent to which a node (author) lies on the paths between other authors.

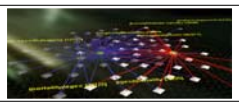




## Exemplification of impact measures using the InfoVis Contest dataset:

**Table 1.** Author ranking based on degree (# co-authors), productivity strength (# produced papers), citation strength (# received citations), and betweenness (# of shortest paths that pass through this author).

| Degree k | # | Productivity Strength $S_p$ | # | Citation Strength $S_c$ | # | Betweenness | # |
|---|---|---|---|---|---|---|---|
| B._Shneiderman | 23 | B._Shneiderman | 7.62 | S._K._Card | 88 | B._Shneiderman | 10893 |
| J._D._Mackinlay | 17 | S._K._Card | 5.71 | J._D._Mackinlay | 67 | S._K._Card | 10618 |
| S._K._Card | 17 | J._D._Mackinlay | 4.37 | B._Shneiderman | 66 | J._D._Mackinlay | 8357 |
| G._Robertson | 16 | Daniel_A._Keim | 4.11 | G._Robertson | 64 | Stephen_G._Eick | 7420 |
| Allison_Woodruff | 15 | Steven F. Roth | 3.96 | Christopher Ahlberg | 36 | Chris_Olston | 5165 |
| Lucy_T._Nowell | 15 | John_T._Stasko | 3.92 | R._Rao | 34 | Ben_Bederson | 4791 |
| Roberto_Tamassia | 15 | Stephen_G._Eick | 3.67 | Ben_Bederson | 25 | Mei_C._Chuah | 4718 |
| Ben_Bederson | 15 | G._Robertson | 3.46 | Peter_Pirolli | 21 | G._Robertson | 3187 |
| Harpreet_S._Sawhney | 14 | Ben_Bederson | 3.40 | Steven_F._Roth | 20 | Steven_F._Roth | 2063 |
| M._Stonebraker | 14 | Marc_H._Brown | 3.33 | Brian_Johnson | 17 | E._H._-H._Chi | 1718 |



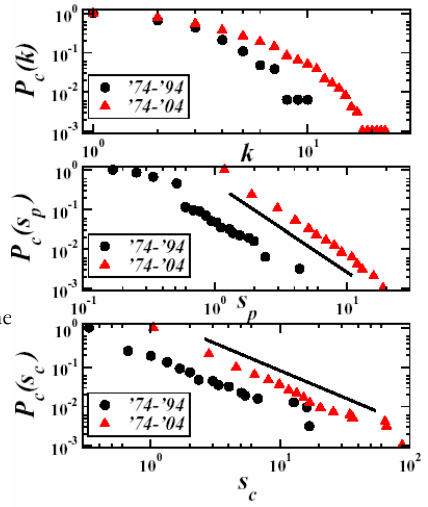
## 5) Global statistical analysis of paper production & citation

**Comparison of cumulative distributions *Pc(x)* of**

- Degree k
- Citation strength $S_c$
- Productivity strength $S_p$

for two time periods: 74-94 and 74-04.

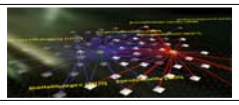
Solid line is a reference to the eye corresponding to a heavy-tail with power-law behavior $P(x) = x^{-\gamma}$ with $\gamma$= 2.0 (for $S_c$) and 1.4 (for $S_p$).




**Discussion:**

- Distributions are progressively broadening in time developing heavy tails.
- We are moving from a situation with very few authors of large impact and a majority of peripheral authors to a scenario in which impact is spread over a wide range of values with large fluctuations for the distribution.

Katy Börner, Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'. February 28th, 2005.



Publication strength $S_p$ and the citation strength $S_c$ of authors versus the degree of authors (number of co-authors) for the 74-04 time slice.

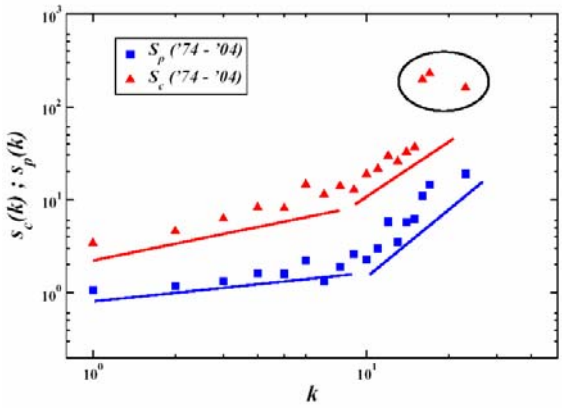
Solid lines are a guide to the eye indicating the presence of two different regimes as a function of the co-authorship degree *k*.

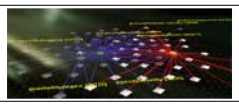
**Discussion:**

Two definite slopes.

Impact and productivity grow faster for authors with a large number of co-authorships.

The three high degree nodes represent S._K._Card, J._D._Mackinlay, and B._Shneiderman.





## Size and Distribution of Connected Components

**Size of connected component is calculated in four different ways**

$G_N$ is the relative size measured as the percentage of nodes within the largest component.

$E_g$ is the relative size in terms of edges.

$G_{sp}$ is the size measured by the total strength in papers of authors in the largest component.

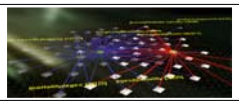
$G_{sc}$ is the size measured by the relative strength in citations of the authors contained in the largest component.

**Exemplification using InfoVis Contest Dataset:**

| | 1974-1994 | 1974-1999 | 1974-2004 |
|---|---|---|---|
| $G_N$ | 8.30% | 12.50% | 15.50% |
| $E_g$ | 14.40% | 16.50% | 20.20% |
| $G_{sp}$ | 10.10% | 21.80% | 24.10% |
| $G_{sc}$ | 19.30% | 38.80% | 40.60% |

There is a steady increase of the giant component in terms of all four measures for the three time slices. Giant component has 15% of authors but 40% of citation impact.

Katy Börner, Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'. February 28th, 2005.



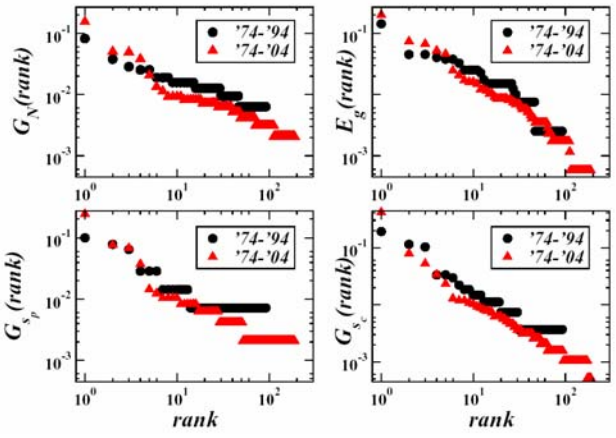
## Zipf plot of the relative sizes of graph components

**Zipf plot** is obtained by ranking all components of the co-authorship graphs in decreasing order of size and then plotting the size and the corresponding rank of each cluster on a double logarithmic scale.

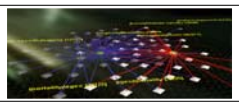
**Discussion:**
Largest component is steadily increasing both in size and impact.

All four curves cross -> the few best ranked components increase at the expense of the smaller ones.

The second largest component is much smaller than the largest one.




Katy Börner, Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'. February 28th, 2005.



## 6) Local, author-centered entropy measure

**Measures the homogeneity of co-authorship weights per author to answer:**

Is the impact of an author spread evenly over all her/his co-authors or are there 'high impact co-authorship edges' that act as strong communication channels and high impact collaborations?

**Novel local entropy-like measure:** $H_{Sx}(i) = \frac{-1}{\log k_i} \sum_j \left( \frac{w_{ij}}{s_x(i)} \right) \log \left( \frac{w_{ij}}{s_x(i)} \right)$

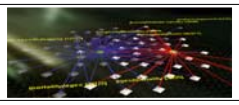
where *x* can be replaced by *p* or *c* denoting the productivity strength or citation strength respectively, *k* is the degree of node *i* and $w_{ij}$ is the impact weight.

This quantity is bounded by definition between 0 and 1. It measures the level of disorder with which the weights are distributed in the neighborhood of each node.

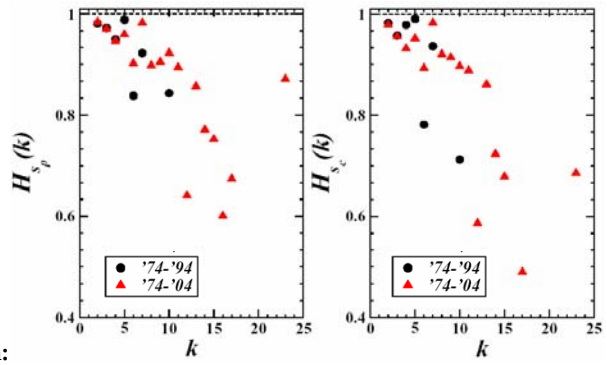
***Homogeneous situation:*** All weights equal, i.e., *wij=w* and *si=ki w*. Entropy equals 1.

***Inhomogeneous situation:*** A small set of connection accumulates a disproportionate weight at the expenses of all others. Entropy goes towards 0.





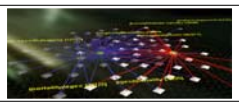
## Entropy spectrum for InfoVis Contest dataset




**Discussion:**

- Entropy is decreases as *k* increases.
- Highly connected authors develop a few collaborations that have a very high strength compared to all other edges.

Katy Börner, Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'. February 28th, 2005.



## Summary

I presented a novel weighted graph representation of coupled networks that can be applied to study the local and global properties of co-authorship networks and to dynamically visualize the changing impact of co-authorship relations.

Based on this graph representation diverse measures were defined to determine:

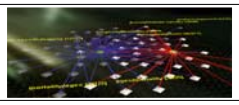
- Global co-author degree and strength distributions,
- The size of the largest connected component and its growth over time,
- The homogeneity of impact weights per author using a novel entropy-like measure, and

Weighted graph representation and measures were applied to analyze a 31-year publication dataset. Major results comprise

- The identification of key authors,
- A first glimpse of the richness of the co-author ecology in which each author has a different role and characteristics,
- A change from a situation in which very few authors have high impact to a scenario in which impact is spread over a wide range of impact values, a steady increase of the size of the largest connected component, and an inhomogeneous distribution of high impact edges per author.

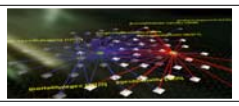
Results obtained for this particular dataset point towards the emergence of a 'global brain', i.e., a more interdisciplinary, globally connected science as opposed to science driven by single experts.





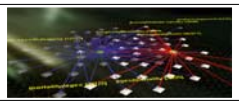
## Future work

- Application of the weighted graph representation and measurement set to diverse publication datasets.
- The normalization of citations for young papers is a serious issue for the analysis of recent developments.
- Issues due to the 'cut' of a dataset (e.g., papers in the very first few years of a dataset receive very few citation edges from papers within the set; citations from outside papers need to be accounted for) need more thorough examination.
- A closer mathematical and empirical examination of the correlation among the four centrality and impact measures of authors and their relation to prior work in bibliometrics is expected to lead to new insights into the co-authorship dynamics.
- We are interested to study the utility of the proposed graph representation and measures to develop more robust and scalable methods for network pruning and visualization.



## Part II: Mapping the Backbone of Science

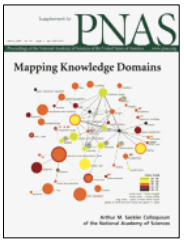
Katy Börner, Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'. February 28th, 2005.



## Mapping Knowledge Domains

Knowledge domain visualizations help answer questions such as:

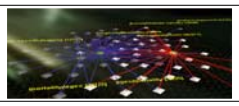
- What are the major research areas, experts, institutions, regions, nations, grants, publications, journals in xx research?
- Which areas are most insular?
- What are the main connections for each area?
- What is the relative speed of areas?
- Which areas are the most dynamic/static?
- What new research areas are evolving?
- Impact of xx research on other fields?
- How does funding influence the number and quality of publications?




Answers are needed by funding agencies, companies, and researchers.

*Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). Mapping Knowledge Domains. Proceedings of the National Academy of Sciences of the United States of America, 101(Suppl_1).*

*Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). Visualizing Knowledge Domains. In Blaise Cronin (Ed.), Annual Review of Information Science & Technology, Volume 37, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, chapter 5, pp. 179-255.*



## Opportunities and challenges for studying the structure and evolution of science

Opportunities:

- Today, many scientific publications are available in digital form (some full text journal data sets go as far back as 110 years).
- We do have algorithms and computing resources to analyze and map science on a large scale.
- We do need better means to access and manage what we collectively know. 'Google' like search is great for fact finding but does not support global understanding and decision making.

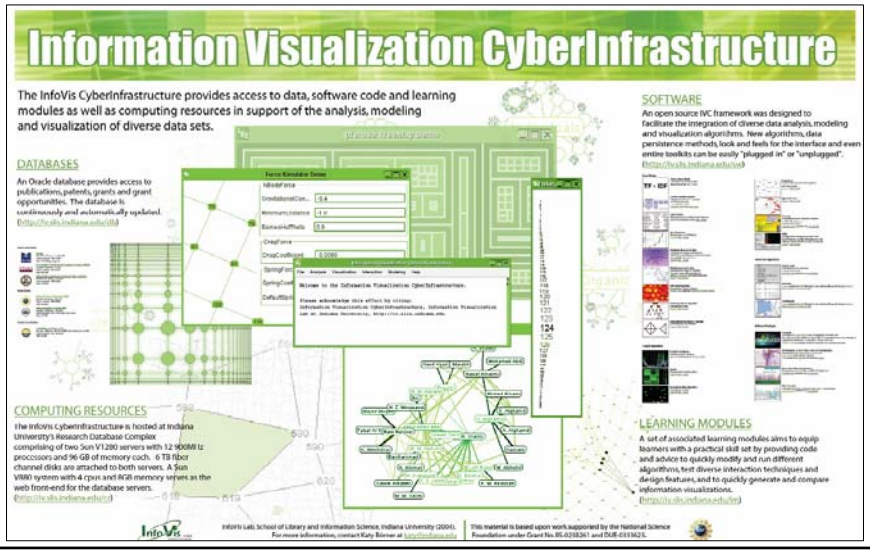
Challenges:

- Data access is difficult. Preservation is a big problem.
- Data integration, i.e., merging data from different databases, is a "hot" research topic as are scalable data analysis and visualization algorithms.
- Code repositories & standardization are needed.
- Data analysis and mapping to generate readable maps is unresolved.
- Educate people about better means to access/use humanity's knowledge & expertise

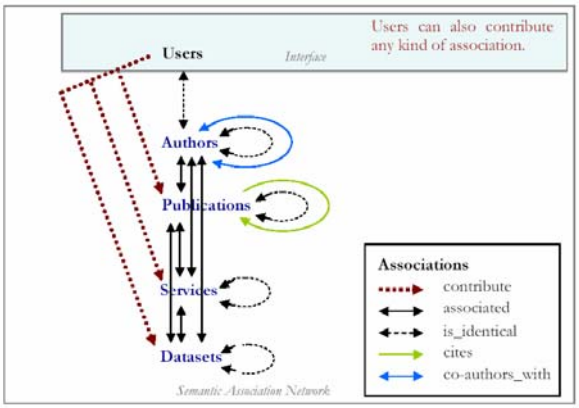
Katy Börner, Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'. February 28th, 2005.

**InfoVis Cyberinfrastructure at IUB** (*http://iv.slis.indiana.edu*) aims to address

- Data access, integration and preservation.
- Code repositories & standardization.
- Education.

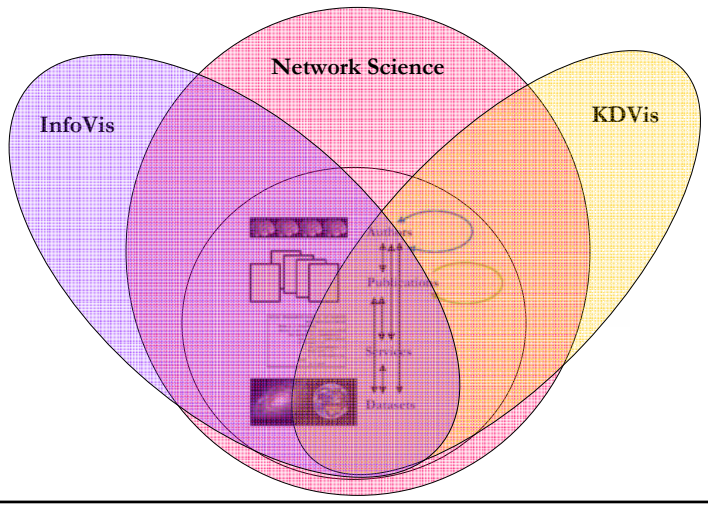


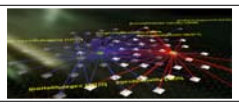
**InfoVis Cyberinfrastructure at IUB** (*http://iv.slis.indiana.edu*) aims to interconnect 'Datasets', 'Services', 'Publications', 'Authors' and 'Users' via diverse associations to facilitate novel means of data access & management.




*Katy Börner. (in press) Semantic Association Networks: Using Semantic Web Technology to Improve Scholarly Knowledge and Expertise Management. In Vladimir Geroimenko & Chaomei Chen (eds.) Visualizing the Semantic Web, Springer Verlag, 2nd Edition, chapter 11.*

**InfoVis Cyberinfrastructure at IUB** (*http://iv.slis.indiana.edu*) **and**
**Network Science Cyberinfrastructure at IUB**
that aims to support large-scale network analysis, modeling and visualization in biomedical, social science and physics research (see NSF SEIII proposal by Katy Börner, Albert-Laszlo Barabasi, Santiago Schnell, Alessandro Vespignani & Stanley Wasserman, Craig Stewart).
**are different views of 'Dataset', 'Service', 'Publication', 'Author' and 'User' data.**





**Generation of accurate and readable KDVis**

Recent work by Kevin W. Boyack, Richard Klavans & Katy Börner aims to evaluate diverse data analysis and mapping techniques to generate citation maps that have a high local [1] and global [2] accuracy.

[1] Klavans, R. & Boyack, K. W. (in press) Identifying a better measure of relatedness for mapping science, *Journal of the American Society for Information Science and Technology.*

[2] Kevin W. Boyack, Richard Klavans & Katy Börner: Mapping the backbone of science. Submitted.

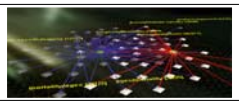


Map of science generated using the *IC-Jaccard* similarity measure. The map is comprised of 7,121 journals from year 2000. Large font size labels identify major areas of science. Small labels denote the disciplinary topics of nearby large clusters of journals.

**Kevin W. Boyack, Richard Klavans & Katy Börner: Mapping the backbone of science. Submitted.**

Map of the backbone of science with 212 clusters comprising 7000 journals. Clusters are denoted by circles that are labeled with their dominant ISI category names. Circle sizes (area) denote the number of journals in each cluster. Circle color depicts the independence of each cluster, with darker colors depicting greater independence. Arrows show all relationships where the citing cluster gives more than 7.5% of its total citations to the cited cluster, with darker arrows indicating a greater fraction of citations given by the citing cluster.

**Kevin W. Boyack, Richard Klavans & Katy Börner: Mapping the backbone of science. Submitted.**



**Kevin W. Boyack, Sandia National Laboratories will present this work**

**<u>Tuesday, March 1st, 2005 in Indianapolis:</u>**

**Lecture: Mapping Science at the Paper Level**

Time: 12:00 – 1:00pm (Lunch will be provided)

Place: Regenstrief Institute, 5th floor classroom

**Tutorial: Please RSVP to Ashley York (ayork@regenstrief.org)**

Time: 2:00 – 3:30pm

Place: Regenstrief Institute, 5th floor classroom

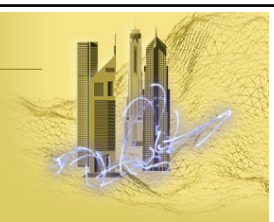
**<u>Thursday, March 3rd, 2005, in the InfoVis Lab meeting in Bloomington:</u>**

Time: 10:30-11:30am

Place: Main Library 036

Katy Börner, Studying the emergent 'Global Brain' in large-scale co-author networks and mapping the 'Backbone of Science'. February 28th, 2005.

## Mapping Humanity's Knowledge and Expertise in the Digital Domain



At the 101st Annual Meeting of the Association of American Geographers Denver, CO: April 5-9, 2005.

### Session Organizers

Katy Börner, Indiana University
André Skupin, University of New Orleans

### Sponsors

Cartography and GIS specialty groups

### Description

This session will bring together leading researchers and practitioners that aim to develop techniques, tools, and infrastructures to map humanity's knowledge and expertise for the improvement of science and education.

Knowledge and expertise is typically extracted from digitally available literature, news, computer mediated communication data as well as from information about the producers and consumers of those data sets. Advanced data analysis techniques in combination with spatial metaphors, geographic principles, and cartographic methods are applied to organize, visualize, and communicate the semantic relationships inherent in the data.

The ultimate goal of this work might be an interactive cartographic map of all of science, with continents representing the major research areas such as, e.g., biology or physics, dots denoting major authors, PIs, papers or news, dynamically evolving research frontiers, blinking 'hot' papers and topics, etc. This map could be used to teach and understand the evolving structure of all of science, to identify major experts, to find and read the most relevant papers and news, to see the effects of resource allocation decisions, to study social networks, etc. Last but not least, it would provide a unique bird's eye view of major experts in specific areas and mankind's knowledge in general.

Some of the leading-edge research on this topic is found where geography intersects with information/library science, computer science, and cognitive science. We invite papers on the broad foundations, computational methods, software systems, and evaluation of such data analyses and visualizations, as they have emerged in this interdisciplinary endeavor.

http://vw.indiana.edu/aag05

## The fourth International Symposium on Knowledge Domain Visualization (KDViz'05)

Knowledge Domain Visualization (KDViz) aims to improve our understanding of the development of a knowledge domain through the study of a wide variety of quantitative and qualitative properties of a knowledge domain. KDViz emphasizes the great potential of an approach that integrates techniques such as information visualization, exploratory data analysis, information retrieval, and information science.

**Aims**
International Symposium on KDViz aims to provide an inter-disciplinary forum for researchers and practitioners from a wide variety of disciplines to address theories, methodologies, techniques, applications, evaluations and case studies in relation to KDViz. The symposium also aims to promote the cross-disciplinary awareness between disciplines such as information visualization and information science. For the purpose of this symposium, a knowledge domain is broadly defined as a dynamic, evolving intellectual structure of a given subject matter. Knowledge domain visualization aims to reveal the dynamics of a knowledge domain by utilizing a wide variety of techniques involving visual thinking, visual discovery, visual exploration, and visual analysis.

**Scope**
The symposium will seek original papers concerning, but not limited to, the following topics. Submitted papers must clearly demonstrate a connection between information visualization and the study of a knowledge domain:

- **Fundamentals of KDViz**
- **Case Studies**
- **Citation Analysis**
- **Domain Analysis and Modeling**
- **Historical, Sociological, or Philosophical Approaches**
- **Knowledge Discovery, Knowledge Representation, and Knowledge Diffusion**
- **Invisible Colleges, Scientific Networks, Social Networks, Scientific Paradigms**
- **Qualitative and Quantitative Methodologies**
- **Scientometrics**
- **Dynamic Models of Scientific Disicplines**
- **Growth Models of Science and Technology**

**Highlights**
A major goal of the symposium is to demonstrate and compare different techniques, algorithms, and approaches that can be utilized to analyze and visualize knowledge domains. In order to facilitate this goal a large-scale data set from the information visualization domain will be made available and participants will be encouraged to utilize this data set to demonstrate new approaches and algorithms.

http://www.graphicslink.demon.co.uk/IV05/KDViz.htm