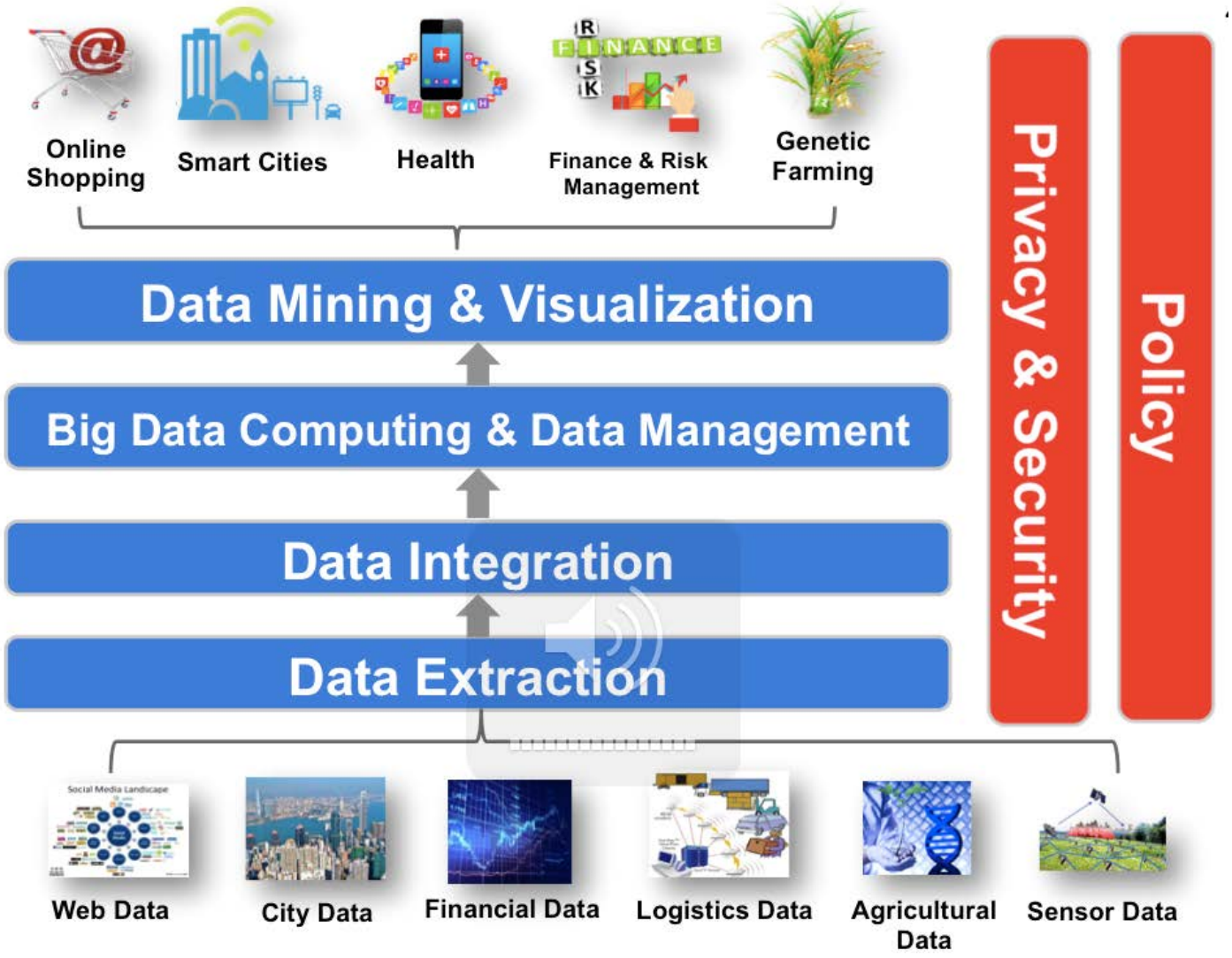
Data Visualization and Data Management

Huamin Qu

Hong Kong University of Science and Technology

INNOVATING TODAY
IMAGINING TOMORROW
敢・創・未來 HKUST

HKUST's Big Data Platform



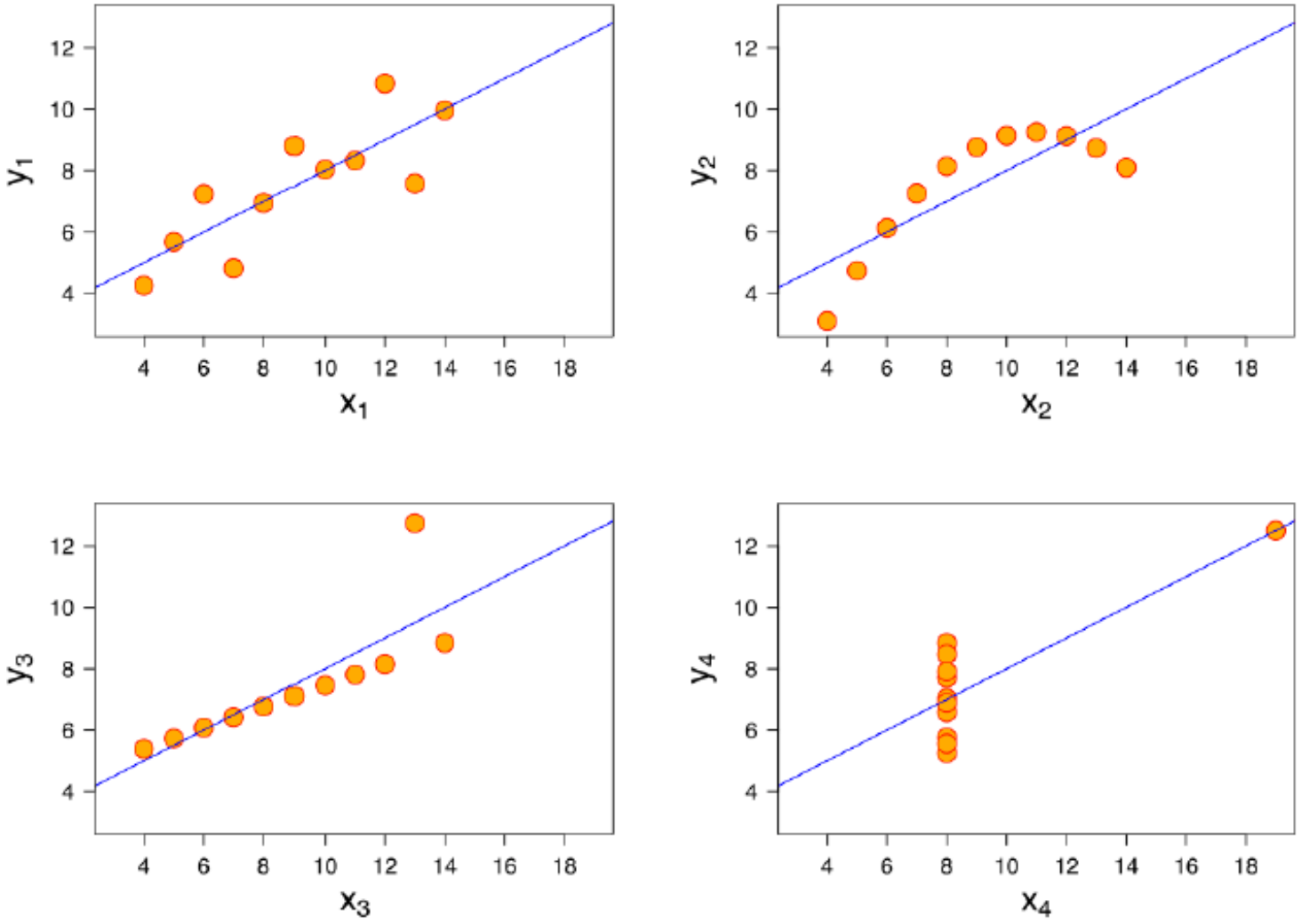
Anscombe's Quartet: Four datasets

Anscombes quarte

I		II		III		IV	
x	*y*	*x*	*y*	*x*	*y*	*x*	*y*
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

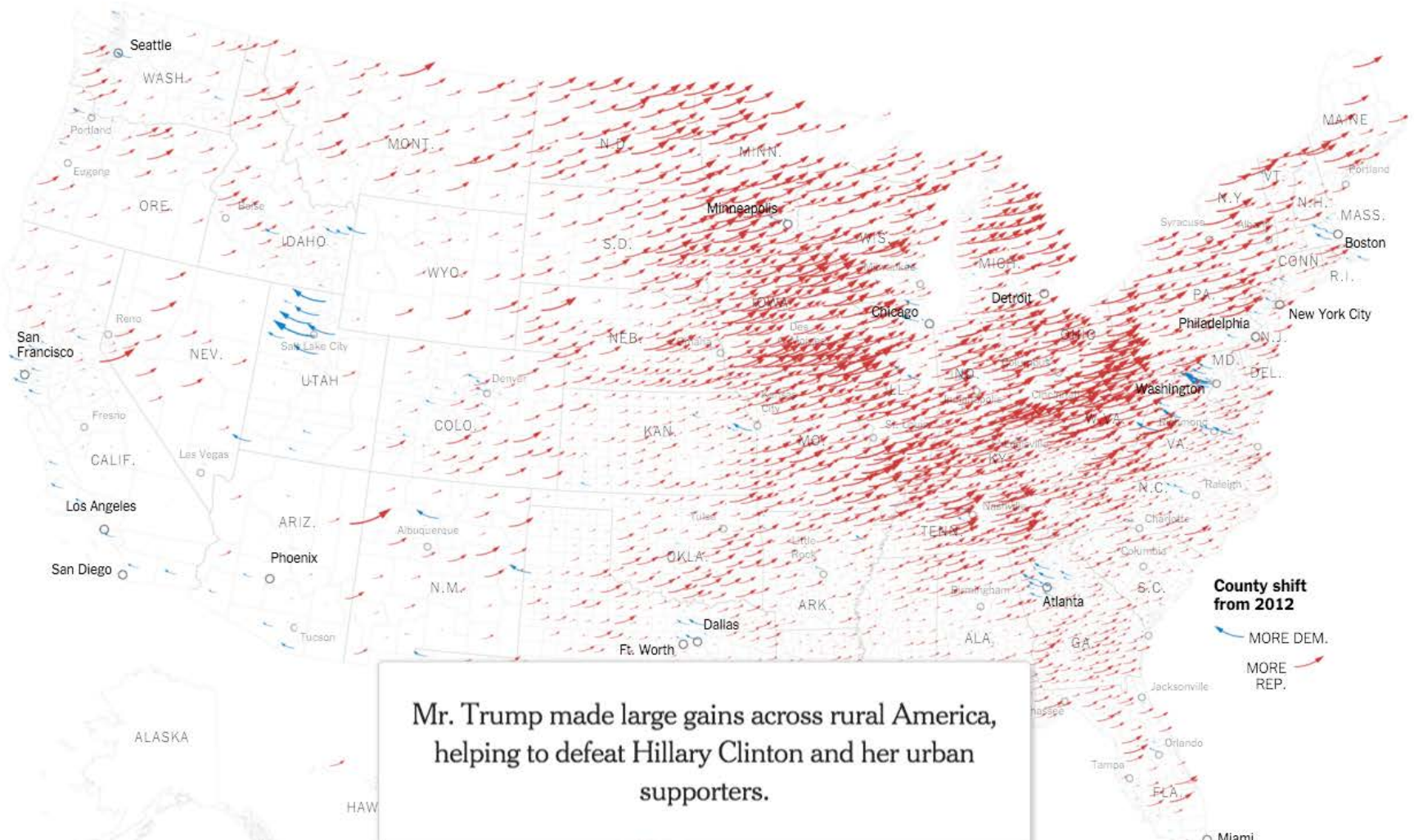
Table 1.1: Anscombe's quartet: four different datasets.

Anscombe's Quartet: Visualizations



What is data visualization?

- Input: data Output: visual form Goal: insight



Obama's big data plan

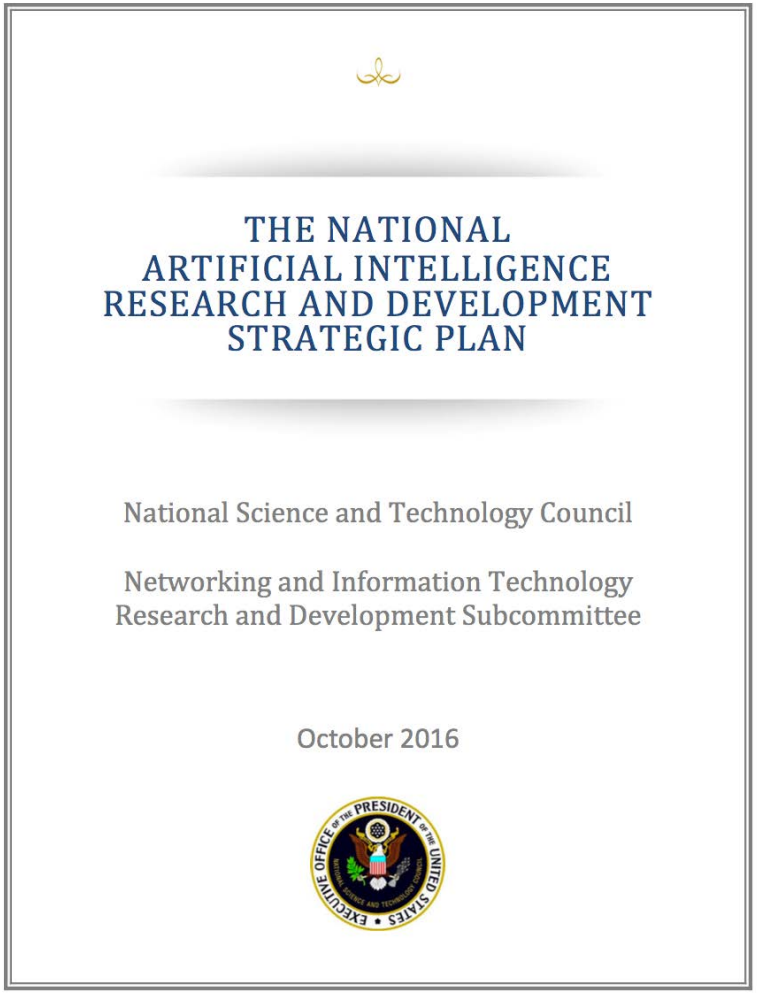
Office of Science and Technology Policy
Executive Office of the President
New Executive Office Building
Washington, DC 20502

FOR IMMEDIATE RELEASE
March 29, 2012

Contact: Rick Weiss 202 456-6037 rweiss@ostp.eop.gov
Lisa-Joy Zgorski 703 292-8311 lisajoy@nsf.gov

OBAMA ADMINISTRATION UNVEILS "BIG DATA" INITIATIVE:
ANNOUNCES $200 MILLION IN NEW R&D INVESTMENTS

- Issuing a $2 million award for a research training group to support training for undergraduates to use graphical and visualization techniques for complex data.



Strategy 2: Develop Effective Methods for Human-AI Collaboration

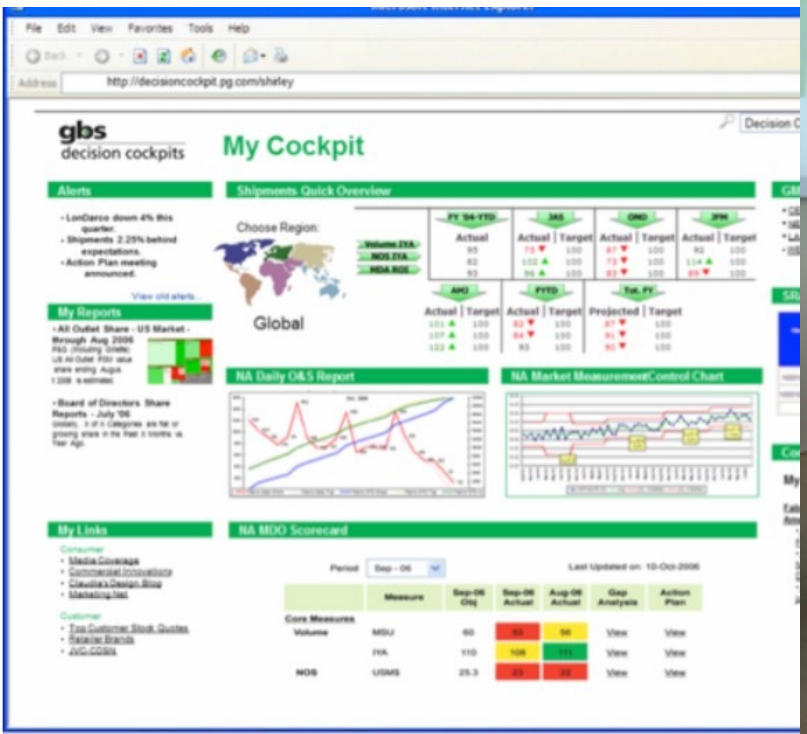
Developing techniques for visualization and AI-human interfaces

Better visualization and user interfaces are additional areas that need much greater development to help humans understand large-volume modern datasets and information coming from a variety of sources. V

HBR Blog Network

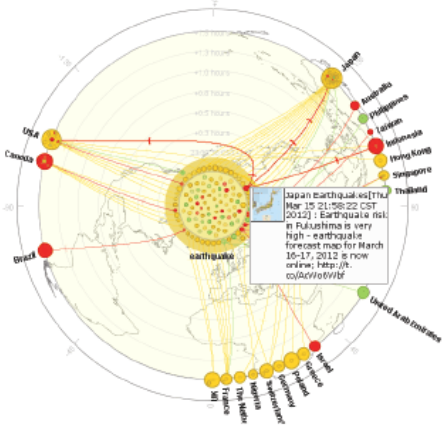
How P&G Presents Data to Decision-Makers

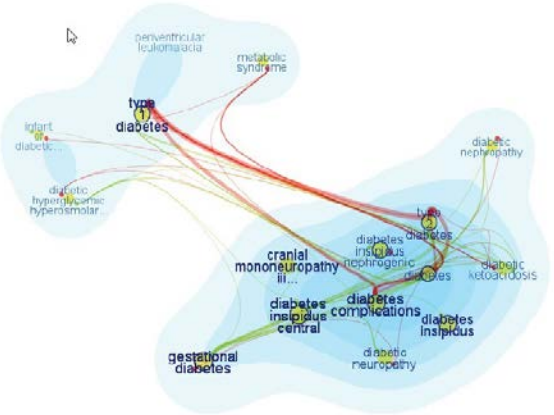
by Tom Davenport | 3:00 PM April 4, 2013

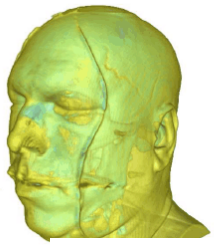


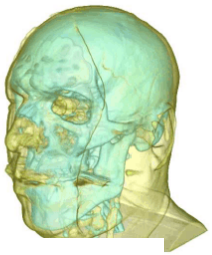
VIS@HKUST

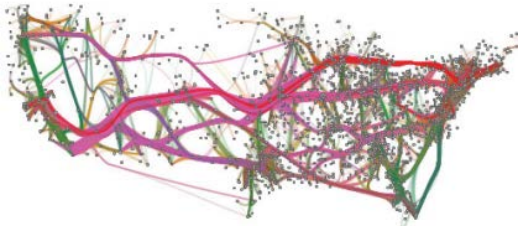

system
visualization
virtual
research
interfaces
computer
graphics
interactive
software
model
data
applications
environment
user
information
simulation
method
process
object
based
approach
analysis
algorithm
rendering
pointing

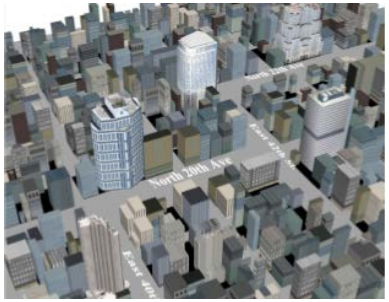

USA
Canada
Brazil
Australia
Philippines
Japan
Indonesia
Hong Kong
Singapore
Thailand
United Arab Emirates
Spain
Greece
Poland
Germany
Switzerland
The Netherlands
France
UK
earthquake
Japan Earthquakes[Thu Mar 15 21:58:22 CST 2012] : Earthquake risk in Fukushima is very high - earthquake forecast map for March 16-17, 2012 is now online; http://t.co/AcWo6WbF

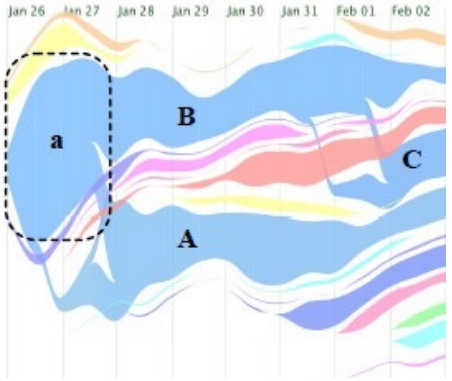

periventricular leukomalacia
metabolic syndrome
type 1 diabetes
infant diabetic...
diabetic nephropathy
diabetic hyperglycemic hyperosmolar...
cranial mononeuropathy
diabetes insipidus nephrogenic
diabetes
diabetic ketoacidosis
diabetes insipidus central
diabetes complications
diabetes insipidus
gestational diabetes
diabetic neuropathy

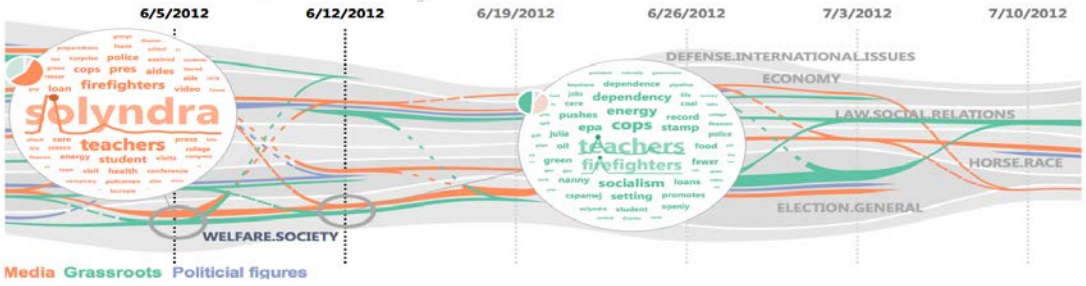







North 10th Ave


Jan 26
Jan 27
Jan 28
Jan 29
Jan 30
Jan 31
Feb 01
Feb 02
a
B
C
A

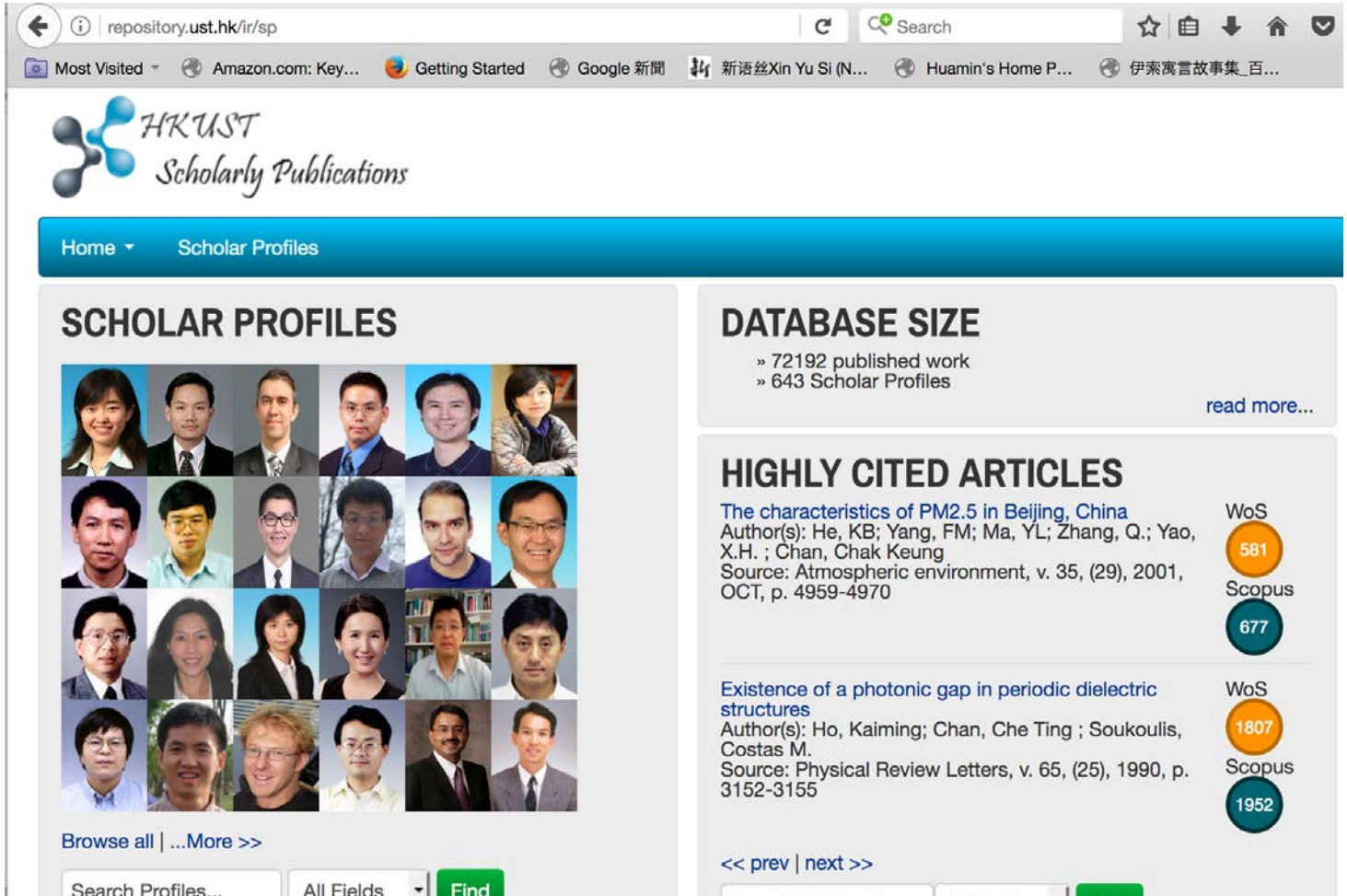

6/5/2012
6/12/2012
6/19/2012
6/26/2012
7/3/2012
7/10/2012
solyndra
police
cops
pres
aides
firefighters
loan
video
teachers
student
energy
health
DEFENSE.INTERNATIONAL.ISSUES
ECONOMY
LAW.SOCIAL.RELATIONS
HORSE.RACE
ELECTION.GENERAL
WELFARE.SOCIETY
dependency
energy
pushes
epa
cops
stamp
record
food
fewer
socialism
setting
promotes
loans
Media
Grassroots
Politicial figures

Visualization for Students

- Choose an advisor
- Know paths of PG students
- Present a research paper

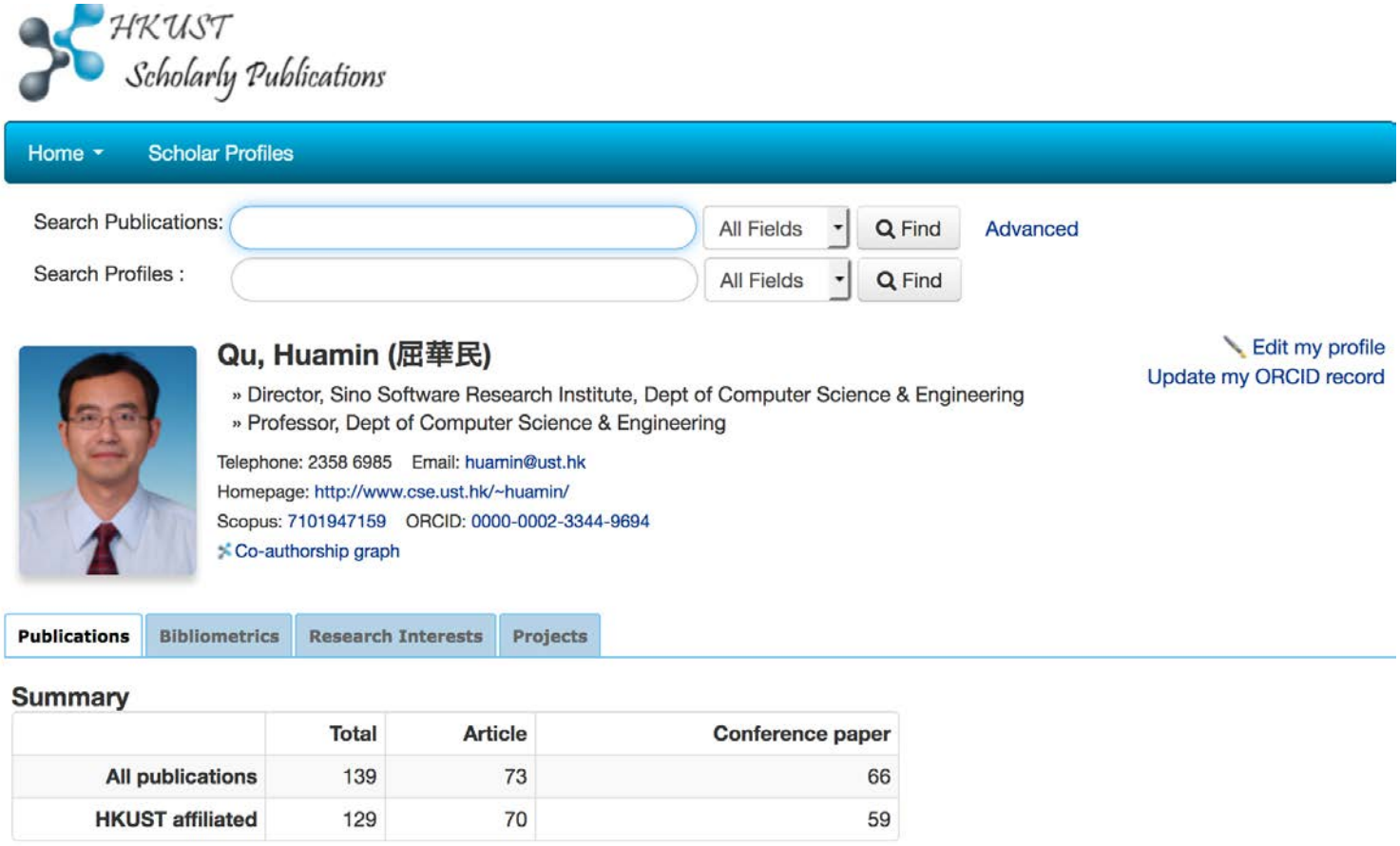
Choosing an Advisor

- http://repository.ust.hk/ir/sp



Choosing an Advisor

- http://repository.ust.hk/ir/AuthorProfile/qu-huamin

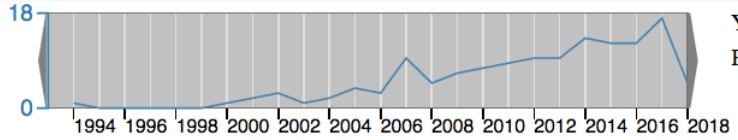


	Total	Article	Conference paper
All publications	139	73	66
HKUST affiliated	129	70	59

HKUST Co-authorship Graph

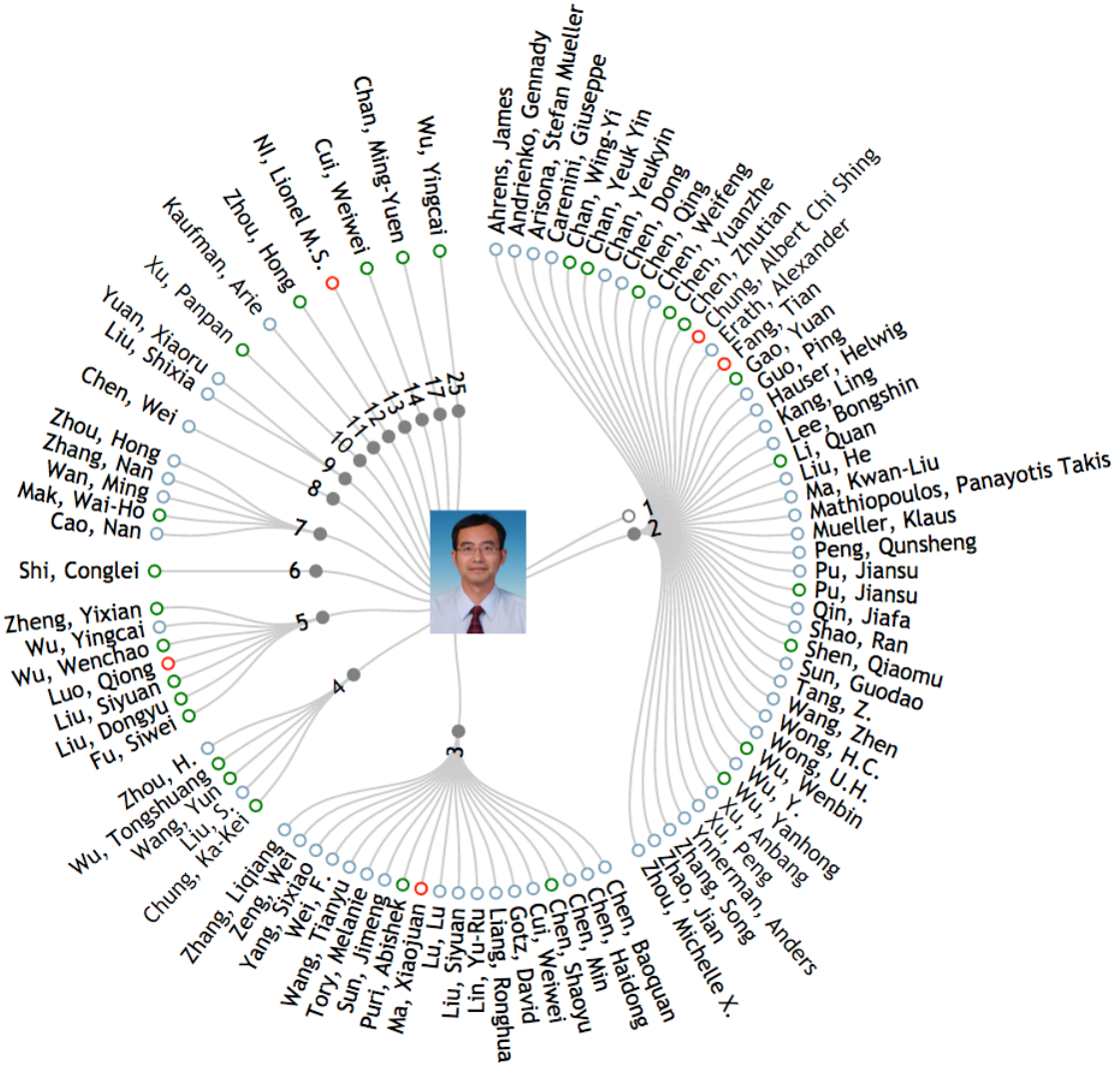
Powered by: HKUST VisLab

Qu, Huamin **Department:** CSE **Total co-authors:** 270 **Total publications:** 139 *Help*



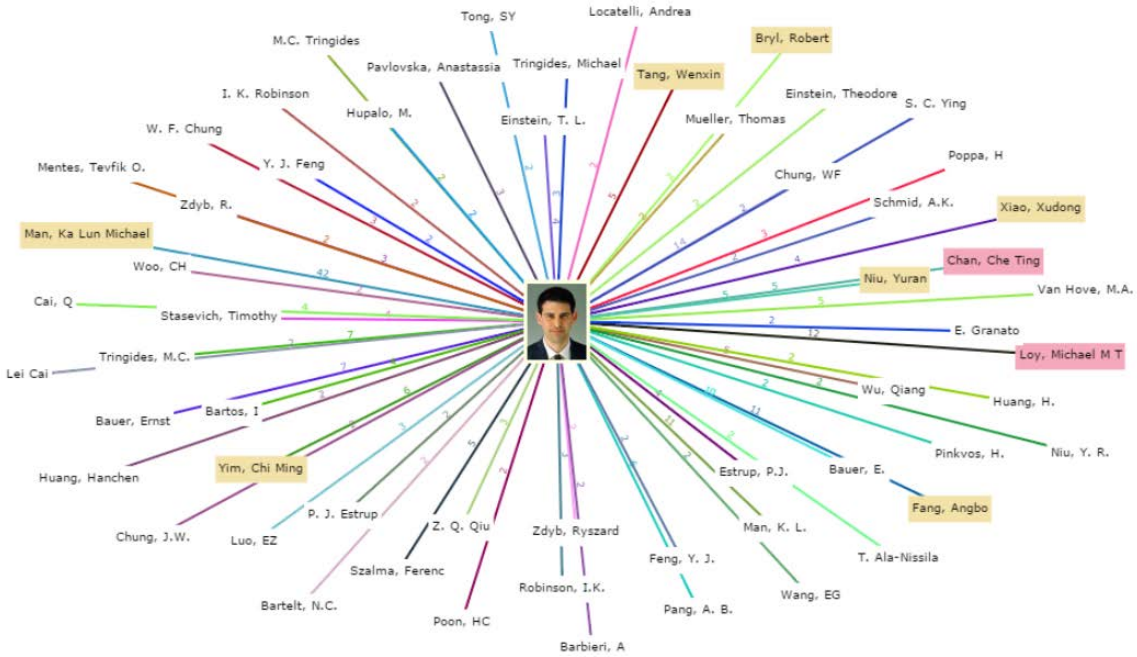
Year: 1993 - 2017
Publication: 139

Co-author List by: Name Publication



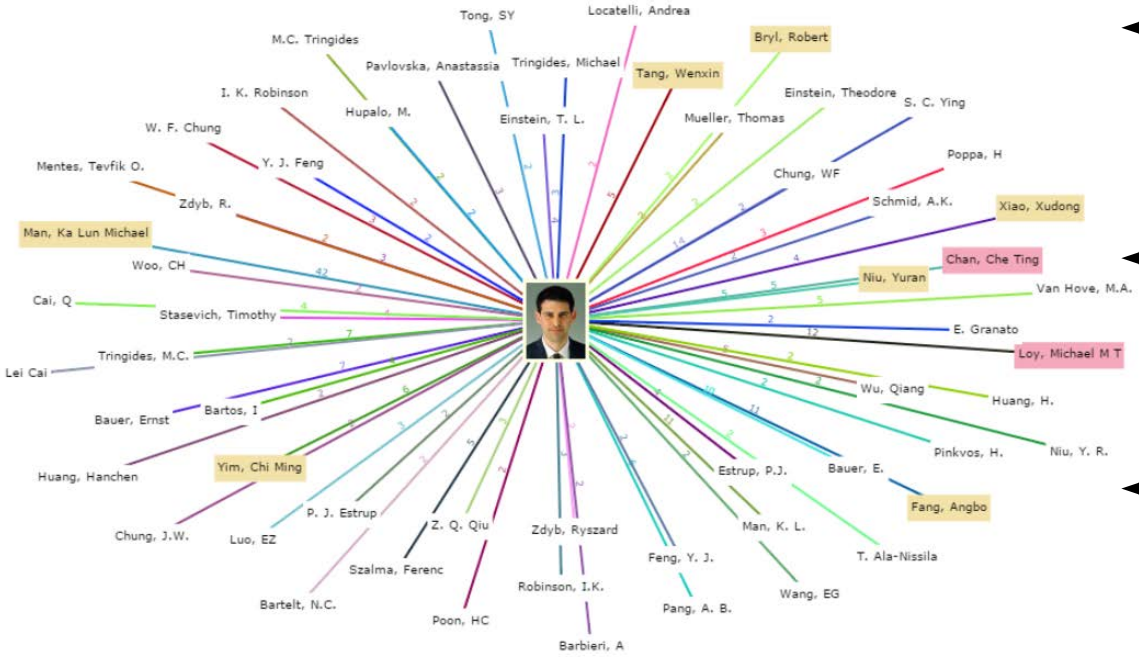
HKUST Faculty Members Other HKUST Members Others Last update: 2017-04-11 00:19:46

VISUALIZATION OF EGO-NETWORK



Michael S. Altman's ego-network visualization from HKUST Scholarly Publications

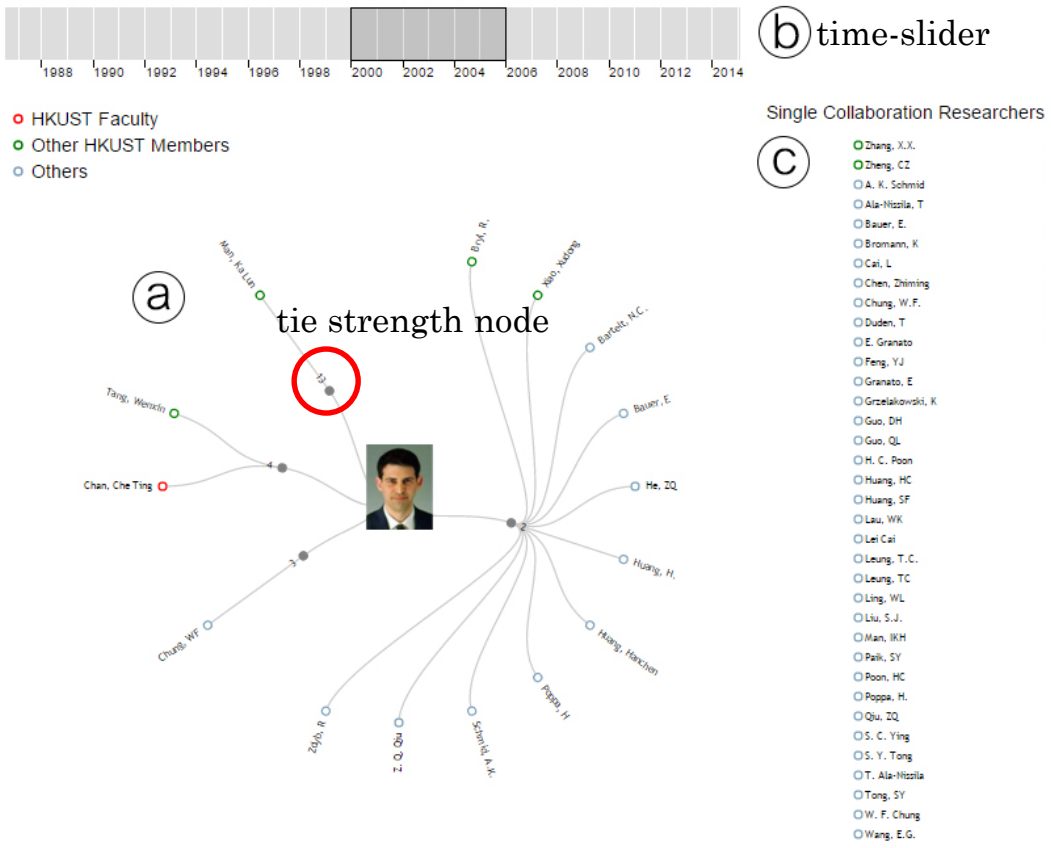
VISUALIZATION OF EGO-NETWORK



Michael S. Altman's ego-network visualization from HKUST Scholarly Publications

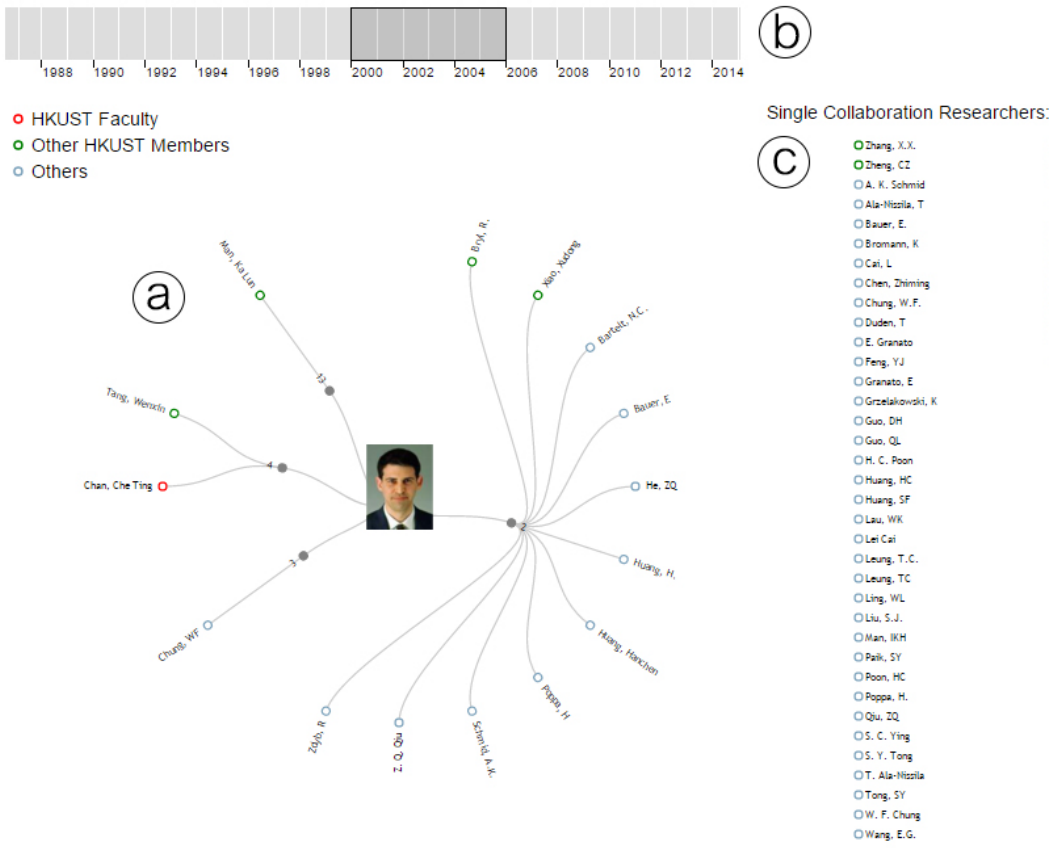
- Do not utilize benefits from animation and color
- Tie strength value on each link is hard to read and compare
- Do not include temporal information

TIME TREE DESIGN



- Alters are grouped by tie strength nodes according to the tie strength value
- Alter nodes shown here correspond to a specified time range from the time-slider
- If the layout is too dense, alters whose tie strength is 1 are listed in the right pane
- Node colors represent each type of alters

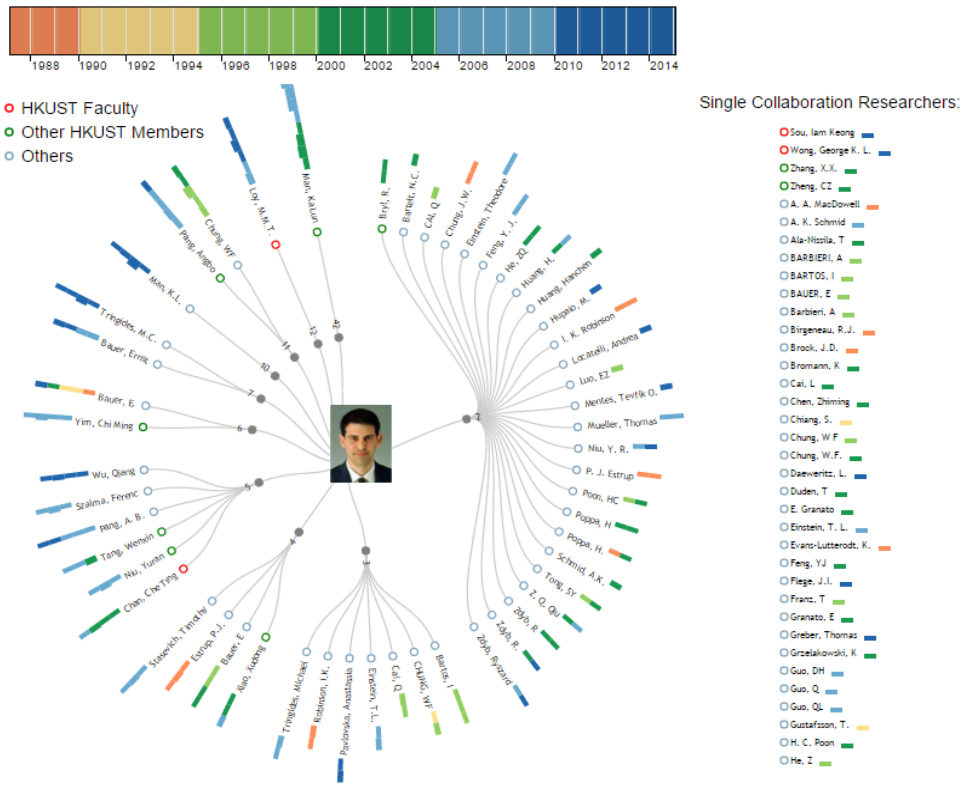
Time tree design



Limitations

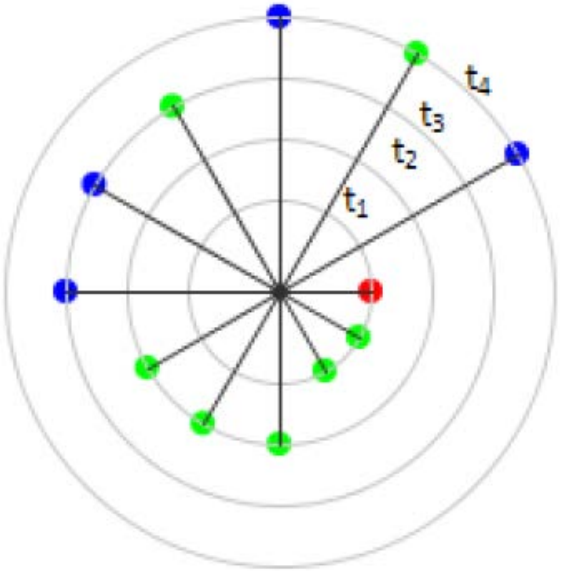
- Do not show overall trends of alters and tie strength value
- Hard to compare between two time steps

Time Tree with Color Bars



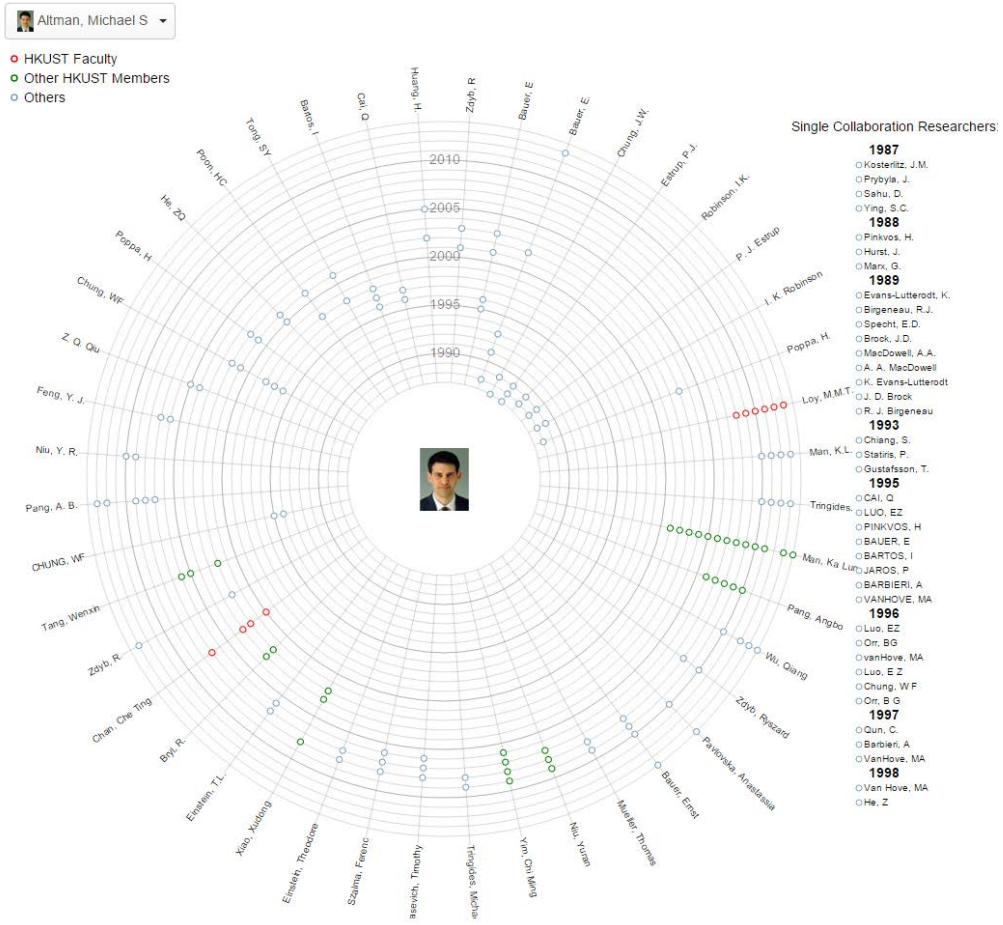
- Colors are assigned for every group of 5 years
- Color bars are added next to alter nodes
- Color bar represents a year that alters interact with the ego
- Height of color bar encodes tie strength respect to that year
- Node colors represent each type of alters

Concentric circle design



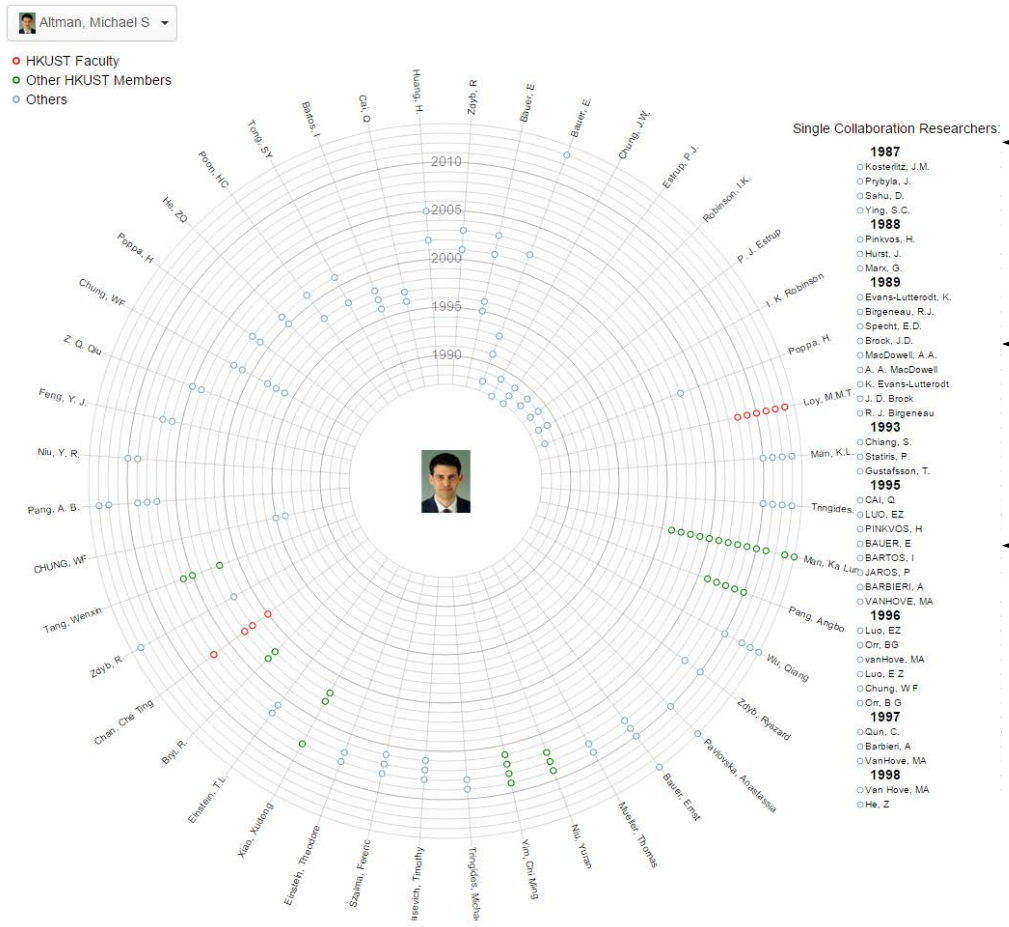
Tree ring ego network layout principle from [2]

Concentric circle design



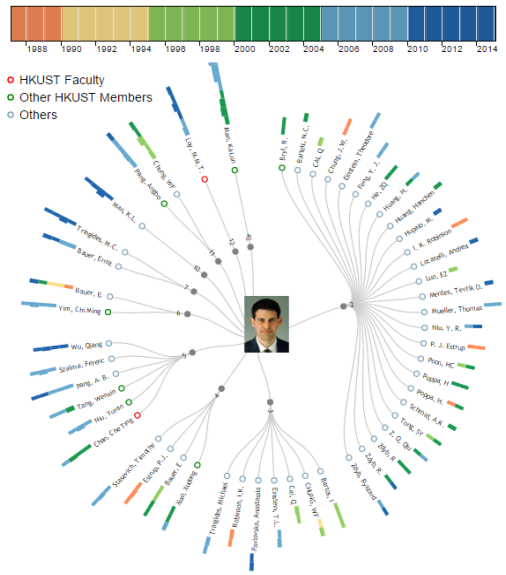
- Each ring represents a one year period
- Alters have their own specific angular positions and are placed according to the time they interacted with the ego
- Radius of each circle encodes tie strength value.
- If the layout is too dense, alters whose tie strength is 1 are listed in the right pane

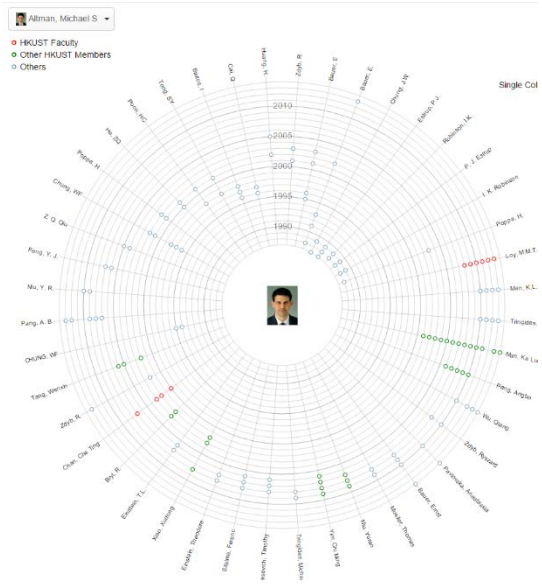
Concentric circle design



- Encode temporal information by concentric circle
- Trends can be observed from nodes appearing on each circle
- Allow comparison between multiple time steps

Comparison

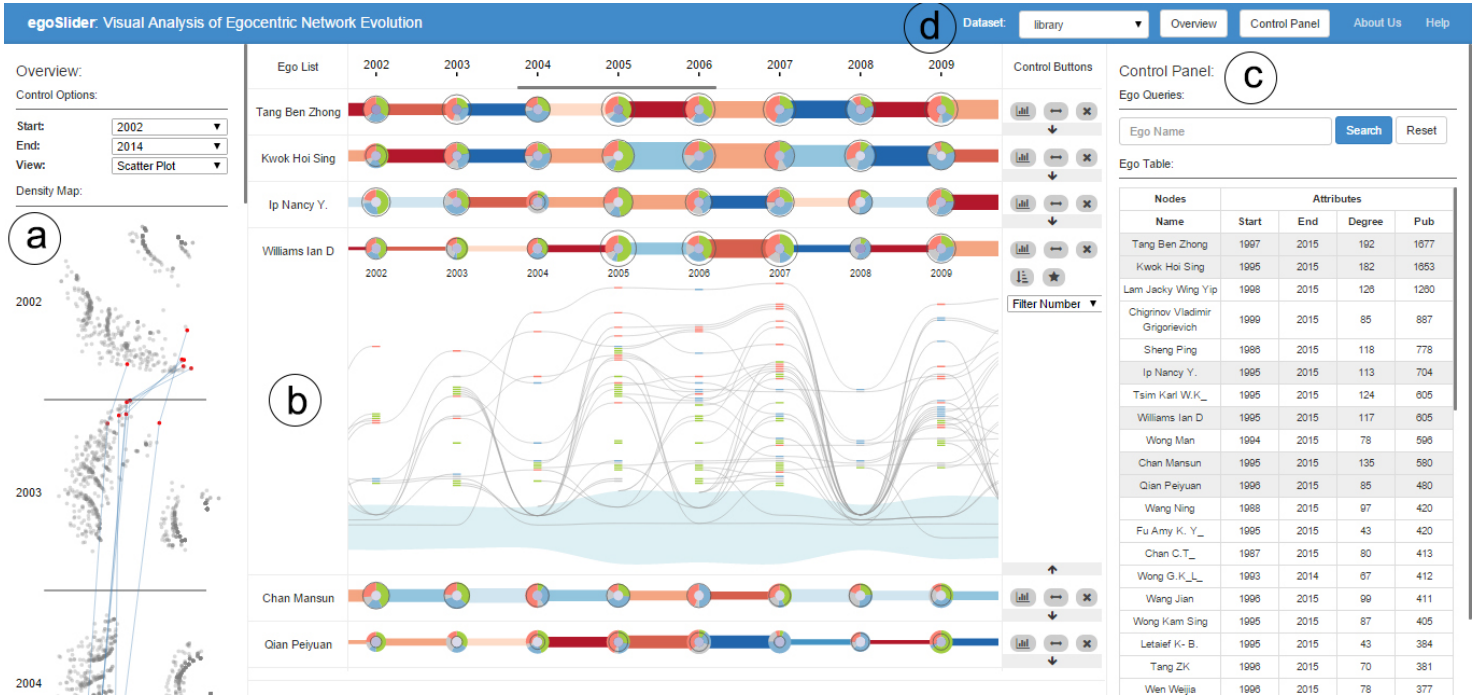




- More compact layout
- Suitable for analysis on specific alters

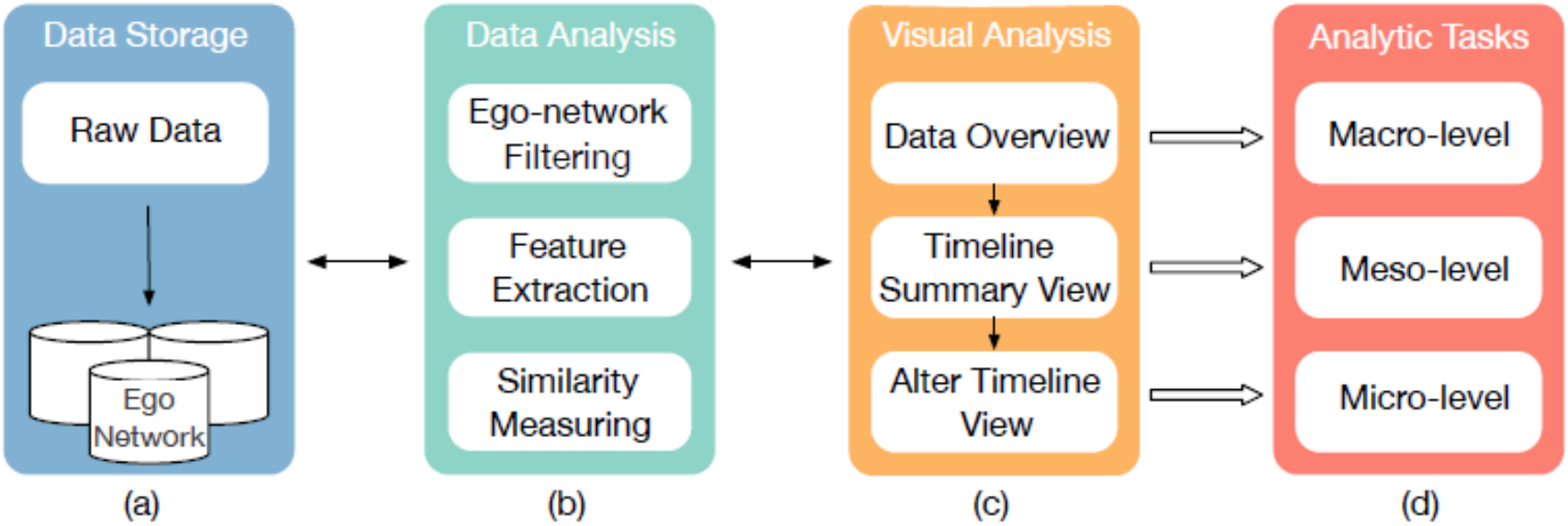
- Require more screen space
- Better comparison between time steps

VISUAL ANALYTICS OF EGO-NETWORK



egoSlider [4], a visual analysis tool for exploring and analyzing dynamic ego-networks.

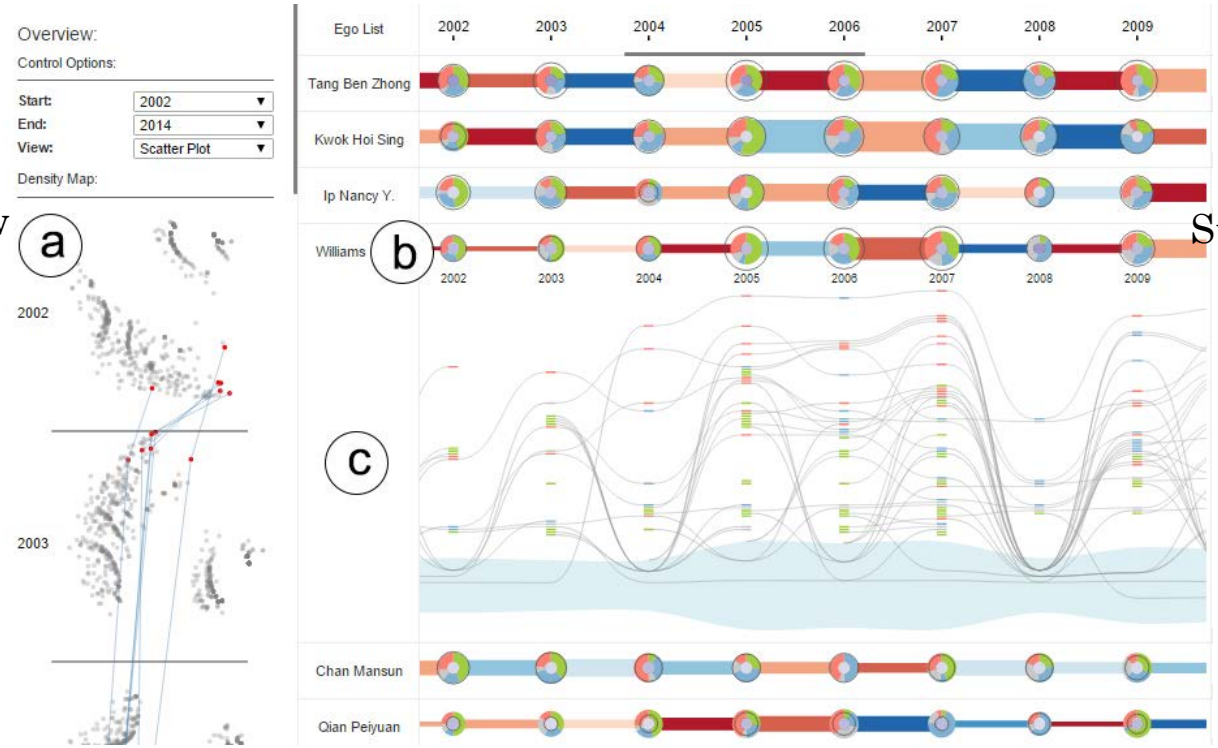
System overview



Macroscopic level tasks focus on the overview of all ego-networks
Mesoscopic level tasks focus on overall comparison of multiple ego-networks
Mesoscopic level tasks focus on details of specific ego-network

Main interface

Data Overview

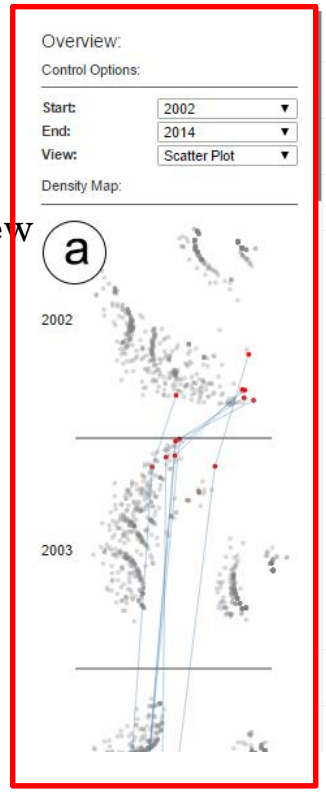


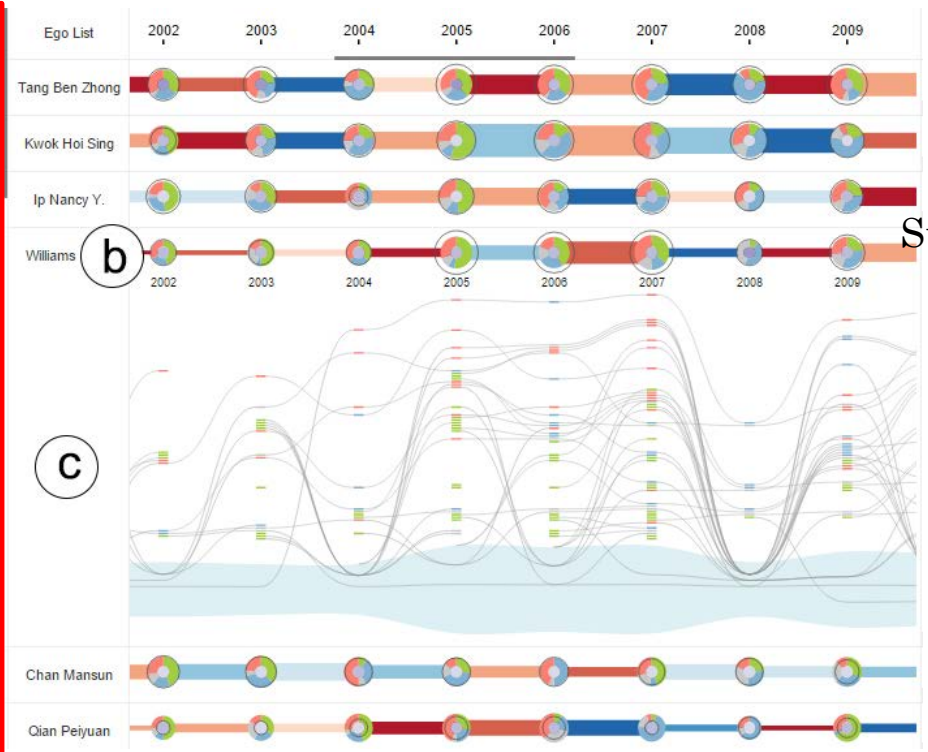
Summary Timeline View

Alter Timeline View

Main interface

Data Overview



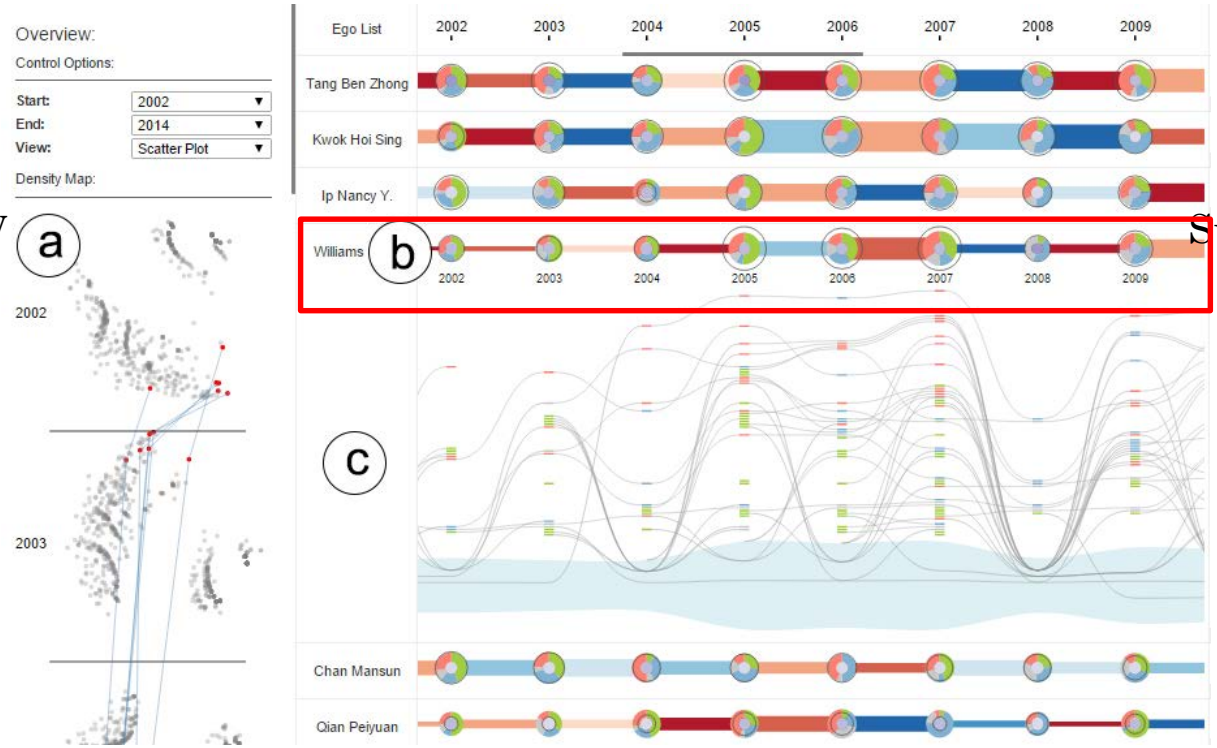


Summary Timeline View

Alter Timeline View

Main interface

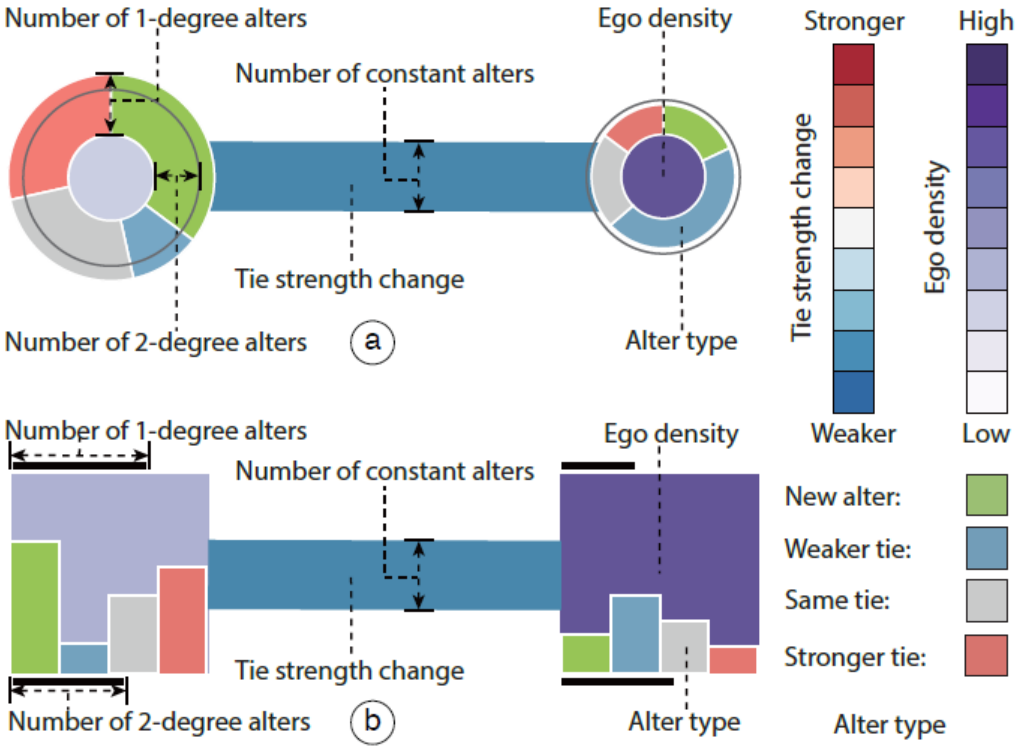
Data Overview



Summary Timeline View

Alter Timeline View

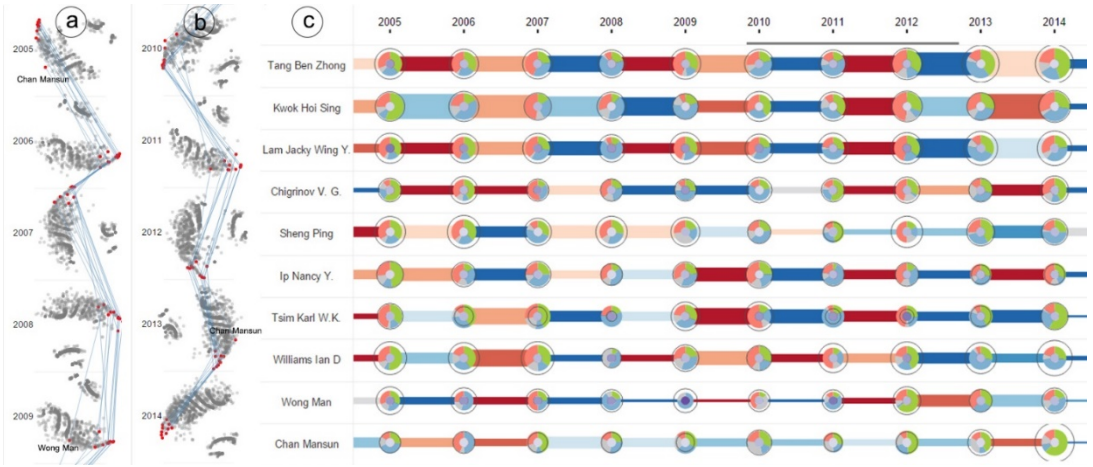
Summary Timeline View



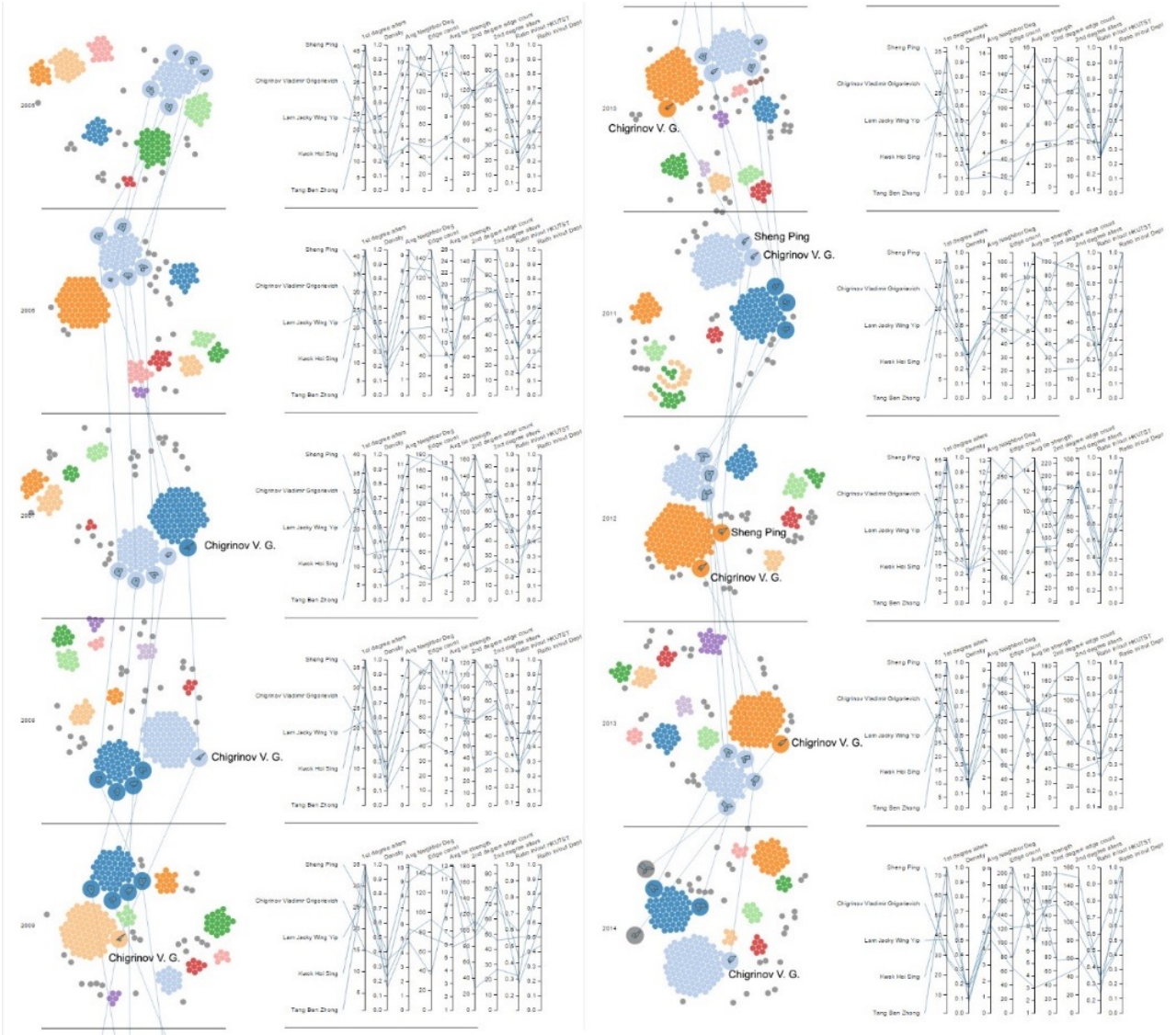
Snapshot glyph and transition glyph

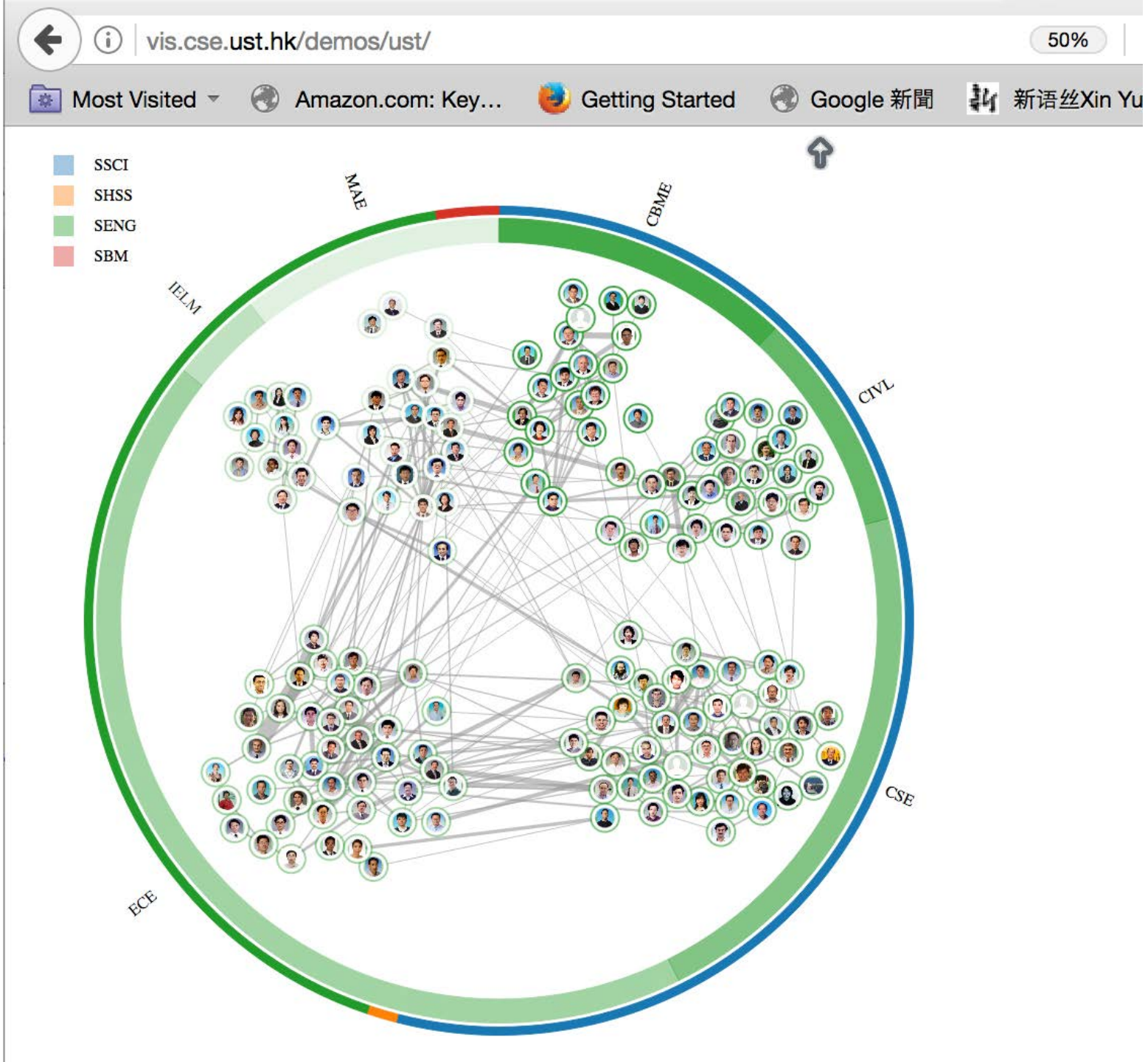
Case Study:
HKUST Collaboration Network

- Retrieve dataset from HKUST Scholarly Publications
- 57383 publications from 1983-2014
- 4189 distinct ego-networks are constructed.
- Tie strength between two authors is defined as the number of papers they co-author in a year.



Case Study


Chigrinov V. G.
Sheng Ping
Chigrinov V. G.
Sheng Ping
Chigrinov V. G.
Chigrinov V. G.
Chigrinov V. G.
Chigrinov V. G.

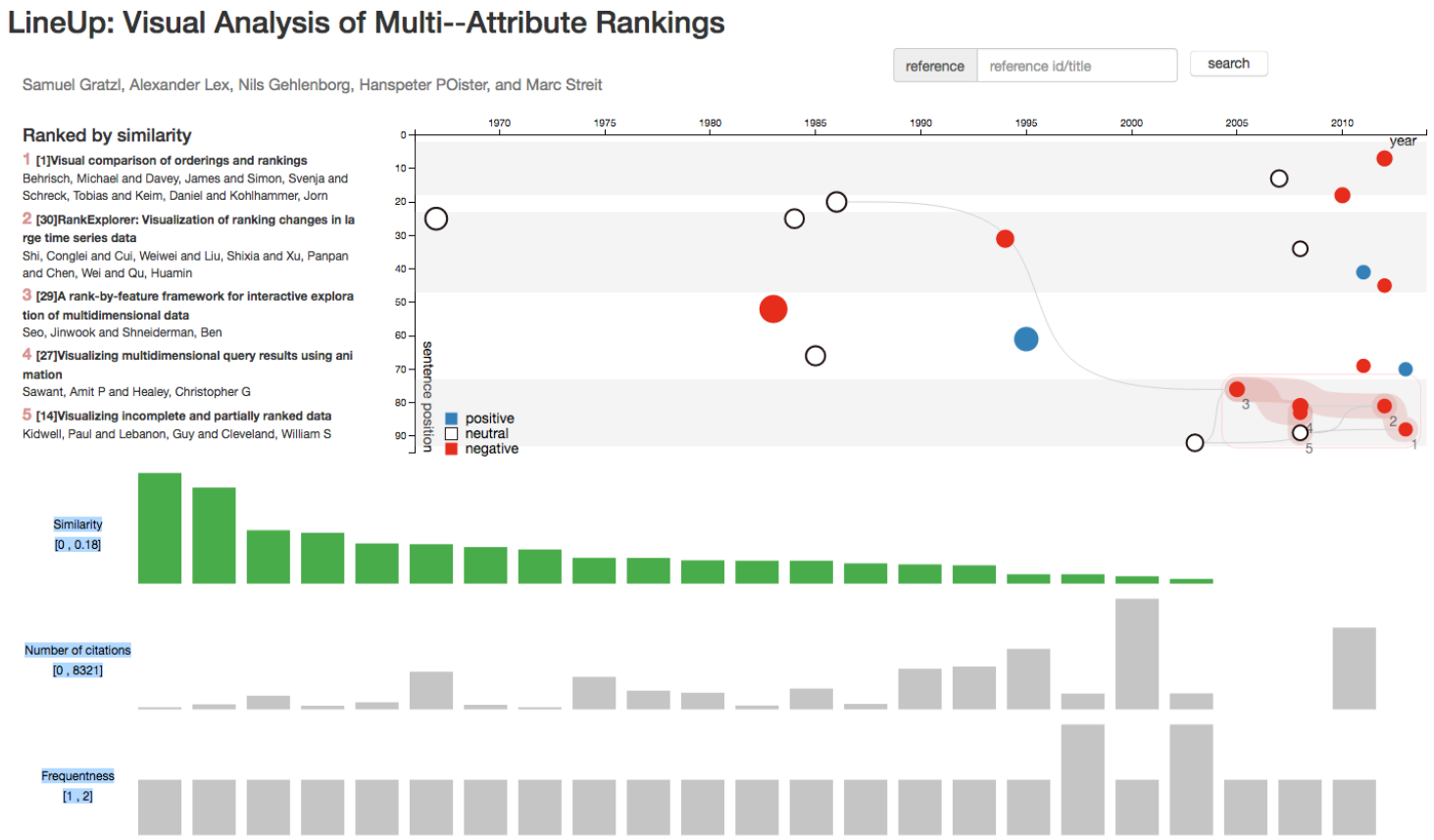

vis.cse.ust.hk/demos/ust/
50%
Most Visited
Amazon.com: Key…
Getting Started
Google 新聞
新语丝Xin Yu
SSCI
SHSS
SENG
SBM
MAE
CBME
IELM
CIVL
CSE
ECE

Paths of PG students

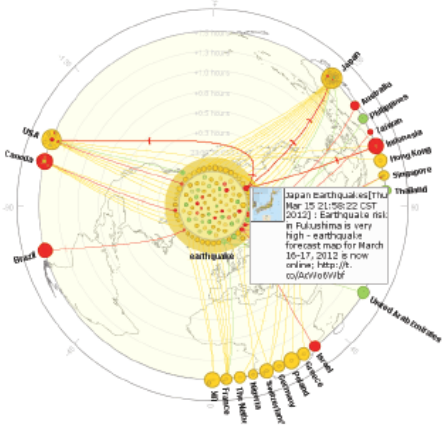
- http://vis.cse.ust.hk/pqeDefenceVis/

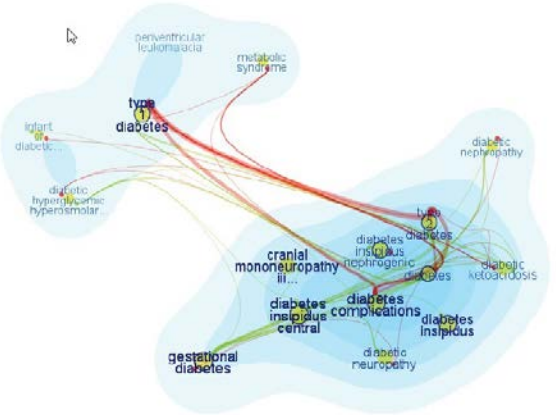
Presenting a Paper

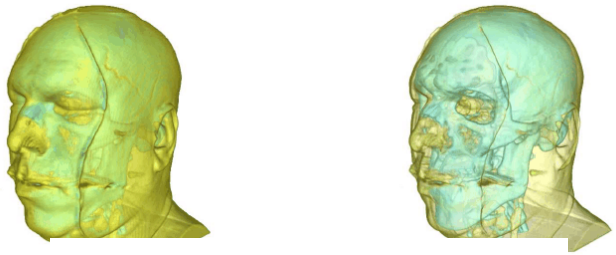
- http://vis.cse.ust.hk/vispaper/#!/

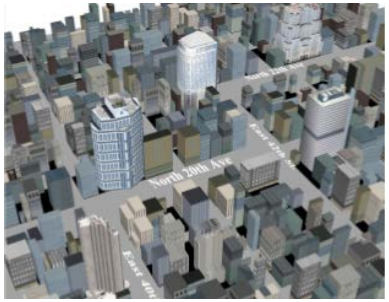


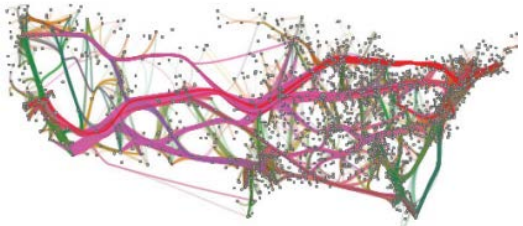
Thank You!

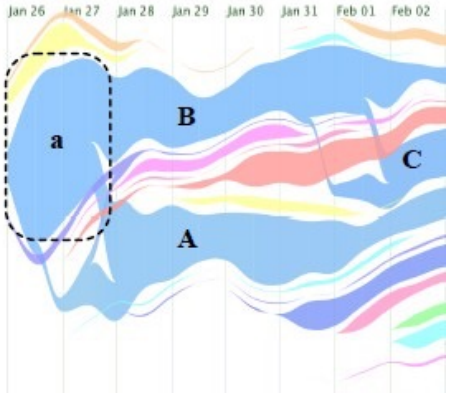


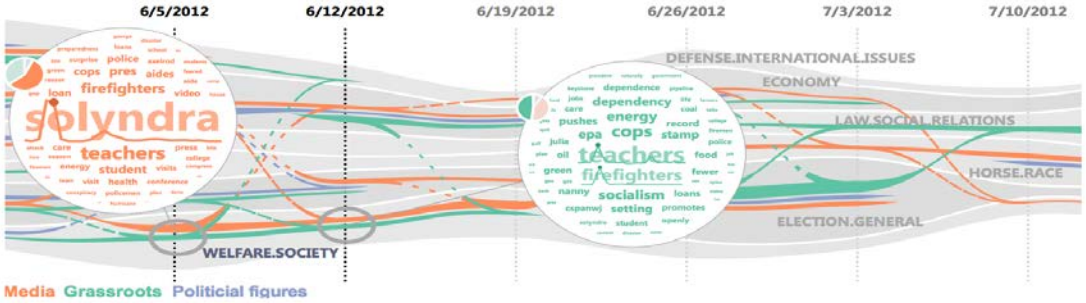












INNOVATING TODAY
IMAGINING TOMORROW
敢・創・未來 HKUST