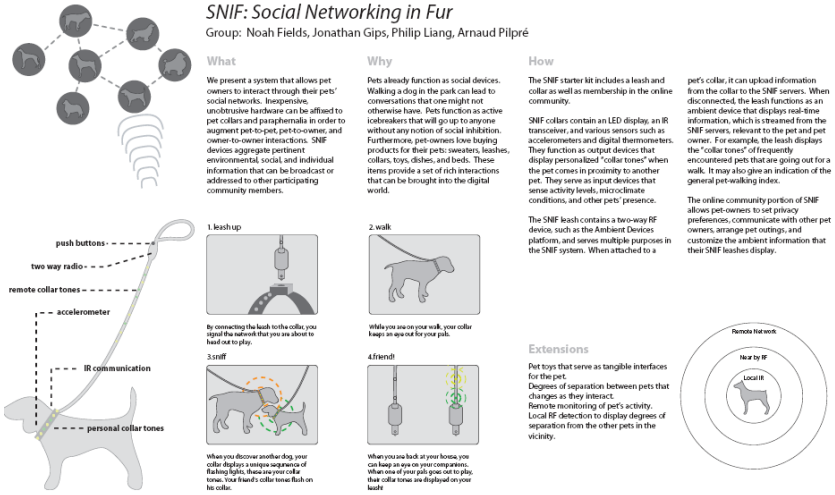
Some Assembly Required: Organizing in the 21st Century



Noshir Contractor

Jane S. & William J. White Professor of Behavioral Sciences

Northwestern University

Twitter: @noshir

Supported by following National Science Foundation grants:

IIS-0838564, IIS-0841583, BCS- 0940851, OCI-0904356, CNS-1010904, IIS-1249137

Army Research Institute W91WAW-08-C-0106/P00004;

Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053

& NIH/NCRR & NIH/NCRR grant to Northwestern University

Clinical and Translational Sciences Institute (NUCATS)

NORTHWESTERN UNIVERSITY

SONIC advancing the science of networks in communities

Building the Team That Built Watson

Ozier Muhammad/The New York Times

David Ferrucci led the team behind Watson, the victorious "Jeopardy" computer. "For the scientist in me," he says, "it was an irresistible challenge.".

By DAVID A. FERRUCCI
Published: January 7, 2012

THE assignment was one of the biggest challenges in the field of artificial intelligence: build a computer smart enough to beat grand champions at the game of "Jeopardy."

Related

Smarter Than You Think: What Is I.B.M.'s Watson? (June 20, 2010)

Computer Wins on 'Jeopardy!': Trivial, It's Not (February 17, 2011)

When I stepped up to lead the team at I.B.M. that would create this computer, called Watson, I knew the task would be formidable. The computer would have to answer an unpredictable variety of complex questions with confidence, precision and speed. And we would put it to the test in a publicly televised "human versus machine" competition against the best players of all time.

It was not easy finding people to join the Watson team in the mid-1990s. Most scientists I approached favored their own individual projects and career tracks. And who could blame them? This was an effort that, at best, would mingle the contributions of many. At its worst it would fail miserably, undermining the credibility of all involved.

David Ferucci,
New York Times
1/7/2012

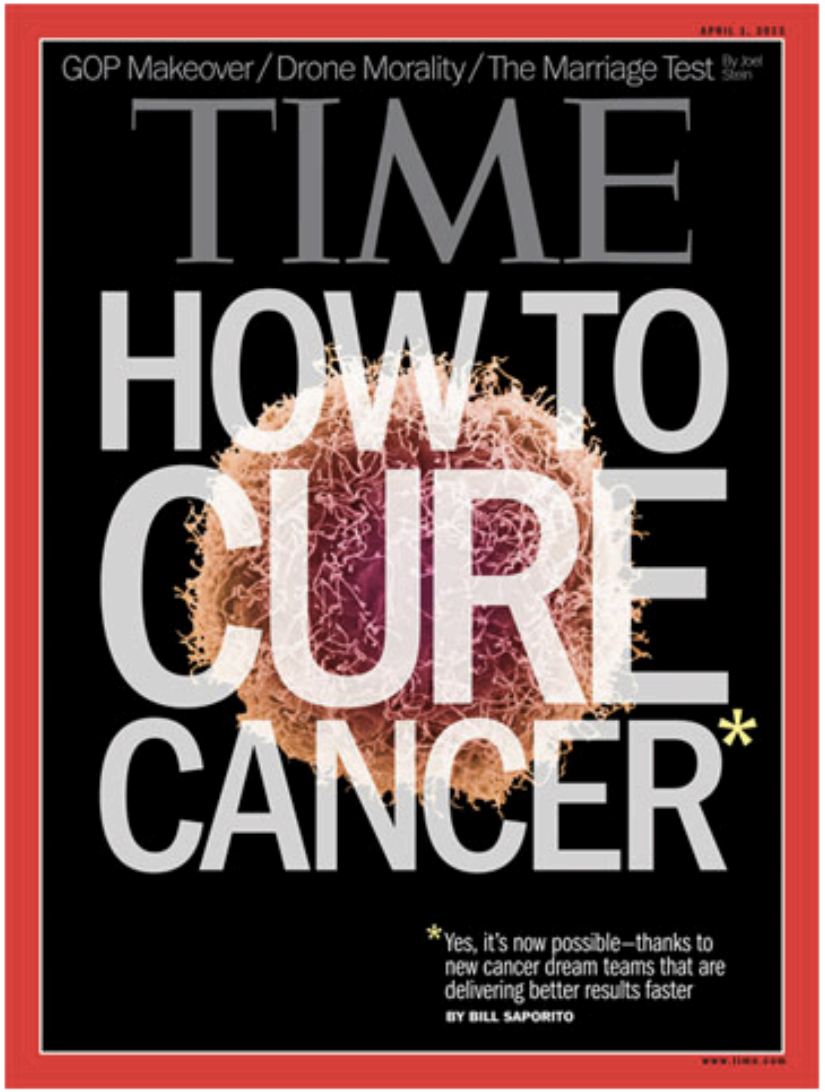
The Conspiracy To End Cancer

By Bill Saporito | Monday, Apr. 01, 2013

The hero scientist who defeats cancer will likely never exist.

No exalted individual, no victory celebration, no Marie Curie or Jonas Salk, who in 1955, after he created the first polio vaccine, was asked, So what's next? Cancer?--as if a doctor finished with one disease could simply shift his attention to another, like a chef turning from the soup to the entrée.

Cancer doesn't work that way. It's not just one disease; it's hundreds, potentially thousands. And not all cancers are caused by just one agent--a virus or bacterium that can be flushed and crushed. Cancer is an intricate and potentially...



Battiere Effect

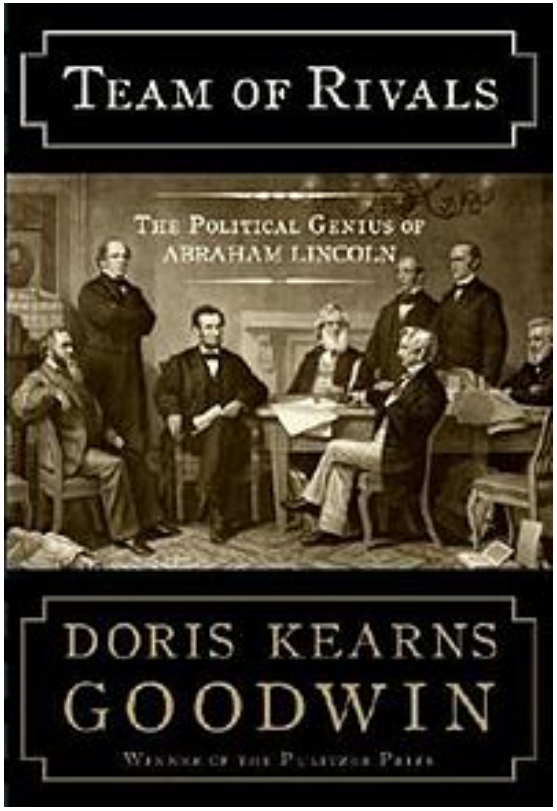
The No-Stats All-Star

Robert Seale for The New York Times

Statistical Anomaly His greatness is not marked in box scores or at slam-dunk contests, but on the court Shane Battier makes his team better, often much better, and his opponents worse, often much worse.

New York Times, Feb 15, 2009

Team of Rivals



Joe Klein: Obama's Team of Rivals

By Joe Klein | Wednesday, June 18, 2008

Barack Obama has never been shy about comparing himself to Abraham Lincoln. He did so when he announced his candidacy at the Illinois state capitol, where both he and Lincoln served in the legislature. "The life of a tall, gangly, self-made Springfield lawyer tells us that a different future is possible," Obama said. "He tells us that there is power in words ... He tells us that there is power in hope." That was, well, audacious, to say the least — and the comparisons have continued, on issues large and small. But the most important similarity, in Obama's mind, is how he plans to govern if elected.

ILLUSTRATION FOR TIME BY STEPHEN KRONINGER; OBAMA: JOSHUA ROBERTS / BLOOMBERG

Time, June 18, 2008

But “Teams of Rivals” are not always succesful ….

Harvard Business Review OnPoint ARTICLE

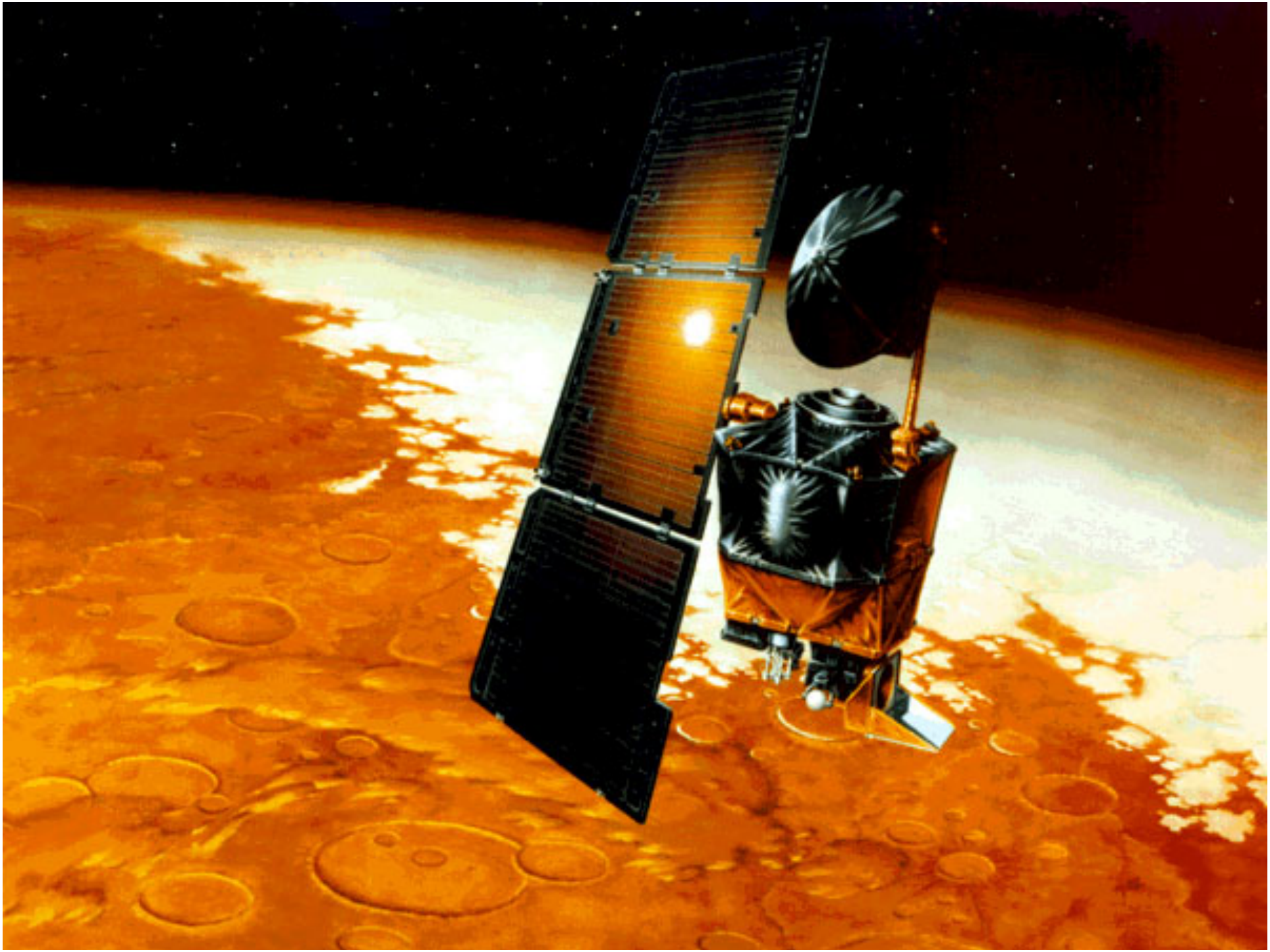
www.hbr.org

Competent Jerks, Lovable Fools, and the Formation of Social Networks

by Tiziana Casciaro and Miguel Sousa Lobo

NORTHWESTERN UNIVERSITY

SONIC advancing the science of networks in communities

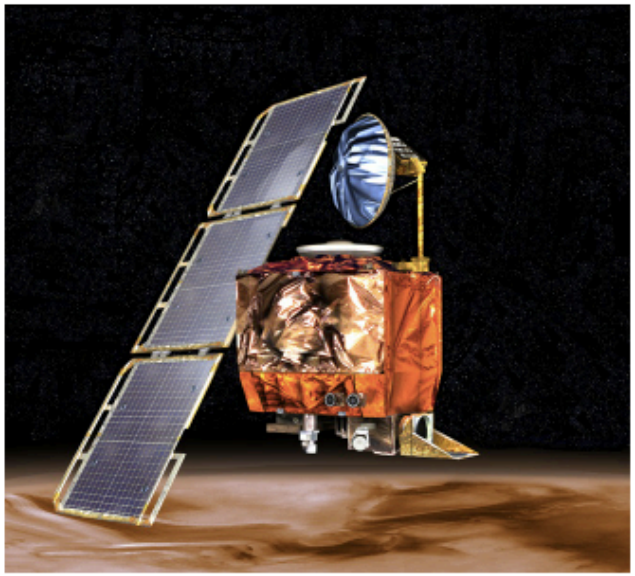


The Multi-Team System Behind the Disaster

Nov. 10, 1999: Metric Math Mistake Muffed Mars Meteorology Mission

WIRED

By Lisa Grossman ✉ November 10, 2010 | 7:00 am |
Categories: 20th century, Disasters, Space Exploration



1999: A disaster investigation board reports that NASA's Mars Climate Orbiter burned up in the Martian atmosphere because engineers failed to convert units from English to metric.

MARS PHOTO GALLERIES:


Where Will Next Mars Rover Land?


Exotic New Mars Images From Orbiting Telephoto Studio


Strange Places on Mars: What Do You Want to See Next?

The $125 million satellite was supposed to be the first weather observer on another world. But as it approached the red planet to slip into a stable orbit Sept. 23, the orbiter vanished. Scientists realized quickly it was gone for good.

"It was pretty clear that morning, within half-an-hour, that the spacecraft had more or less hit the top of the atmosphere and burned up," recalled NASA engineer Richard Cook, who was project manager for Mars exploration projects at the time.

A NASA review board found that the problem was in the software controlling the orbiter's thrusters. The software calculated the force the thrusters needed to exert in *pounds* of force. A separate piece of software took in the data assuming it was in the metric unit: *newtons*.

"The units thing has become the lore, the example in every kid's textbook from that point on," Cook said. "Everyone was amazed we didn't catch it."

NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

Tasks don't always come before Teams

Journal of Applied Statistics
Vol. 32, No. 5, 461–474, July 2005

Routledge

The Most-Cited Statistical Papers

THOMAS P. RYAN* & WILLIAM H. WOODALL**

*National Institute of Standards and Technology, Gaithersburg, Maryland, USA, **Department of Statistics, Virginia Tech, Blacksburg, Virginia, USA

(19) With 2,529 citations (120 per year),

Box, G. E. P. & Cox, D. R. (1964) An analysis of transformations, *Journal of the Royal Statistical Society, Series B*, 26, pp. 211–243 (discussion pp. 244–252).

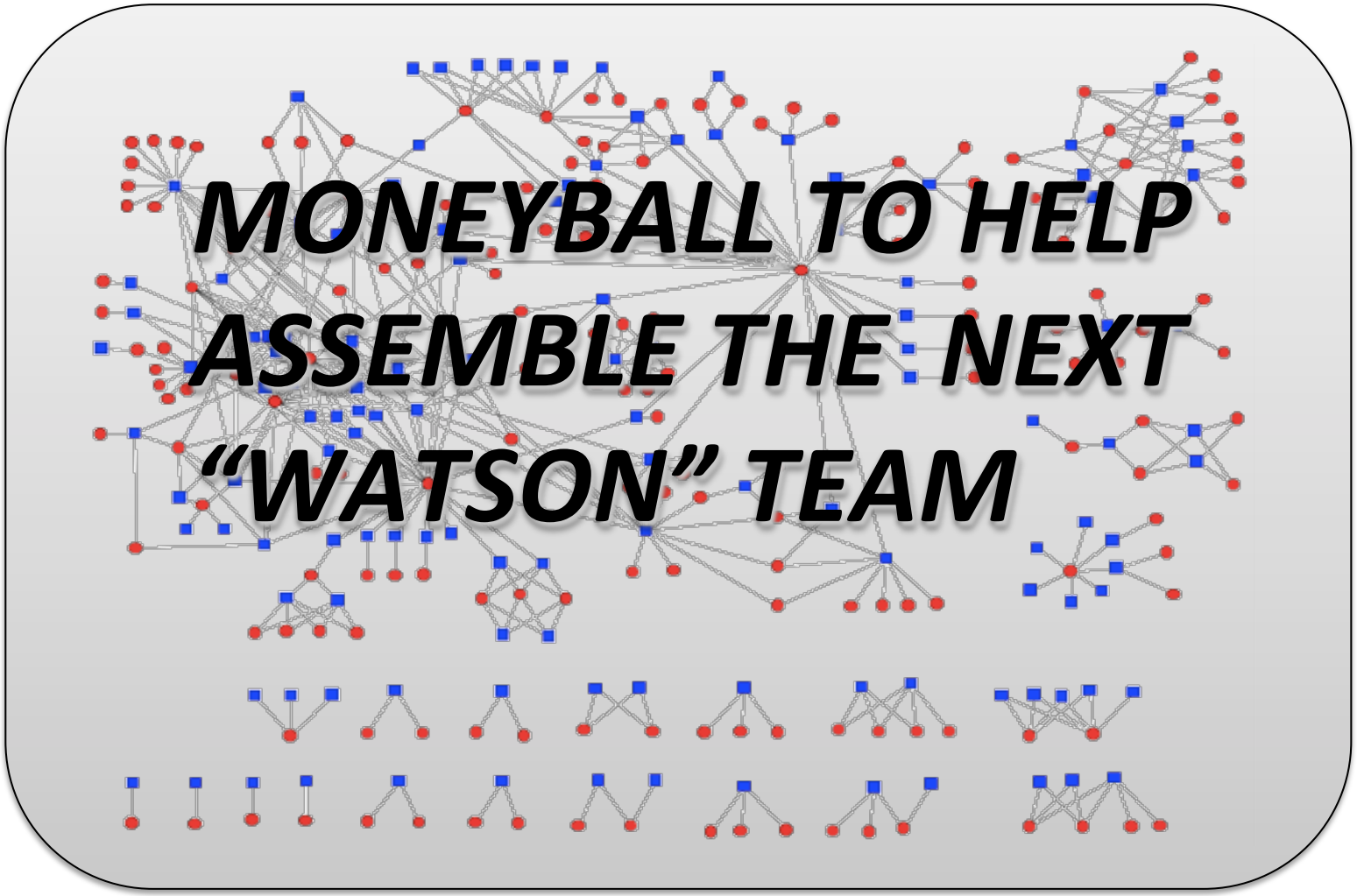
DeGroot (1987) provided some interesting background on this paper from an interview with Professor Box. Box recounted, for example, that he and Cox were on a committee of the Royal Statistical Society and several people suggested that they collaborate. Their motivation and the idea of the paper sprung, to some extent, from the similarities of their family names.

Box & Cox (1964) presented a very useful family of power transformations that have typically been used to transform the dependent variable in a regression model so as to try to meet the assumptions of homoscedasticity and normality of the error terms. The right side of the model can then be transformed in the same manner so as to retrieve the quality of the fit before the dependent variable was transformed.

NORTHWESTERN
UNIVERSITY

DeGroot, M. H. (1987) A conversation with George Box, *Statistical Science, 2*, pp. 239 – 258

advancing the
science of networks in communities



"Your goal shouldn't be to buy players. Your goal should be to buy wins. In order to buy wins, you need to buy runs." (Bakshi & Miller, 2011).

Perfect Storm ….

- We are well poised to enable team assembly by leveraging recent advances in:
  - Theories: Theories about the socio-technical motivations for creating, maintaining, dissolving and re-creating links to engage in teams
  - Methods: An ensemble of qualitative and quantitative methods techniques (such as exponential random graph modeling or *p**) to understand and enable theoretically grounded network recommendations for teams
  - Data: The development of **Virtual Web Observatories** provide the technological capability to capture, store , merge, and query relational metadata about data and tools needed to more effectively understand and enable teams.
  - Computational infrastructure: Cloud computing and petascale applications are critical to face the computational challenges in understanding and enabling teams.

Multi-theoretical Multilevel (MTML) Motivations for Team Assembly

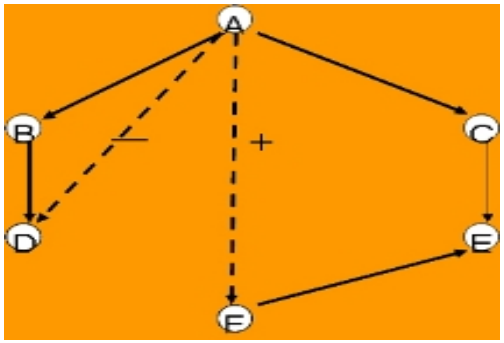
- Theories of self-interest
- Theories of social and resource exchange
- Theories of mutual interest and collective action
- Theories of contagion
- Theories of balance
- Theories of homophily
- Theories of proximity

Sources:

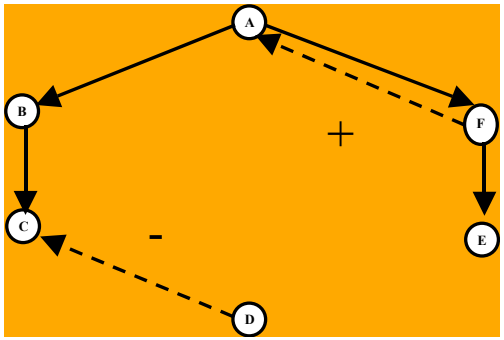
Contractor, N. S., Wasserman, S. & Faust, K. (2006). Testing multi-theoretical multilevel hypotheses about organizational networks: An analytic framework and empirical example. *Academy of Management Review*.

Monge, P. R. & Contractor, N. S. (2003). *Theories of Communication Networks*. New York: Oxford University Press.

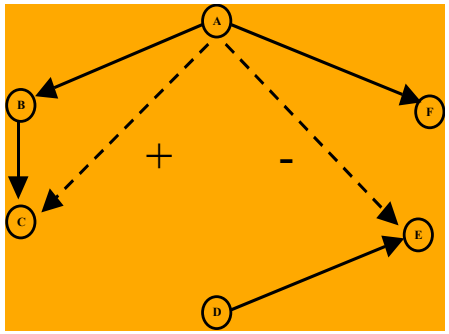
"Structural signatures" of MTML Motivations for Team Assembly



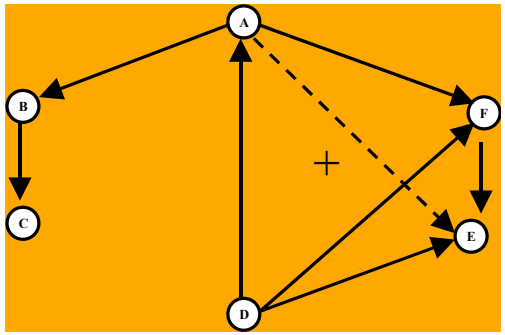
Theories of Self interest



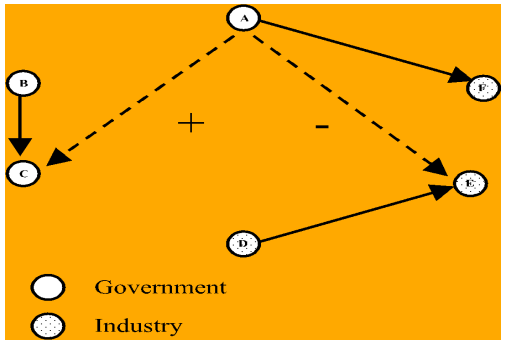
Theories of Exchange



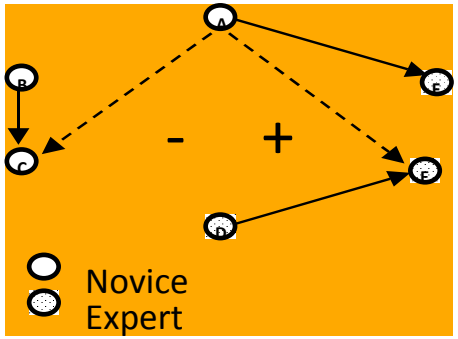
Theories of Balance



Theories of Collective Action



Theories of Homophily



Theories of Cognition

Statistical "MRI" for Structural Signatures

- p*/ERGM: Exponential Random Graph Models
- Statistical "Macro-scope" to detect structural motifs in observed networks
- Move from exploratory to confirmatory network analysis to understand multi-theoretical multilevel motivations for why we create social and information networks

Challenges of empirically testing, extending, and exploring theories about team assembly …

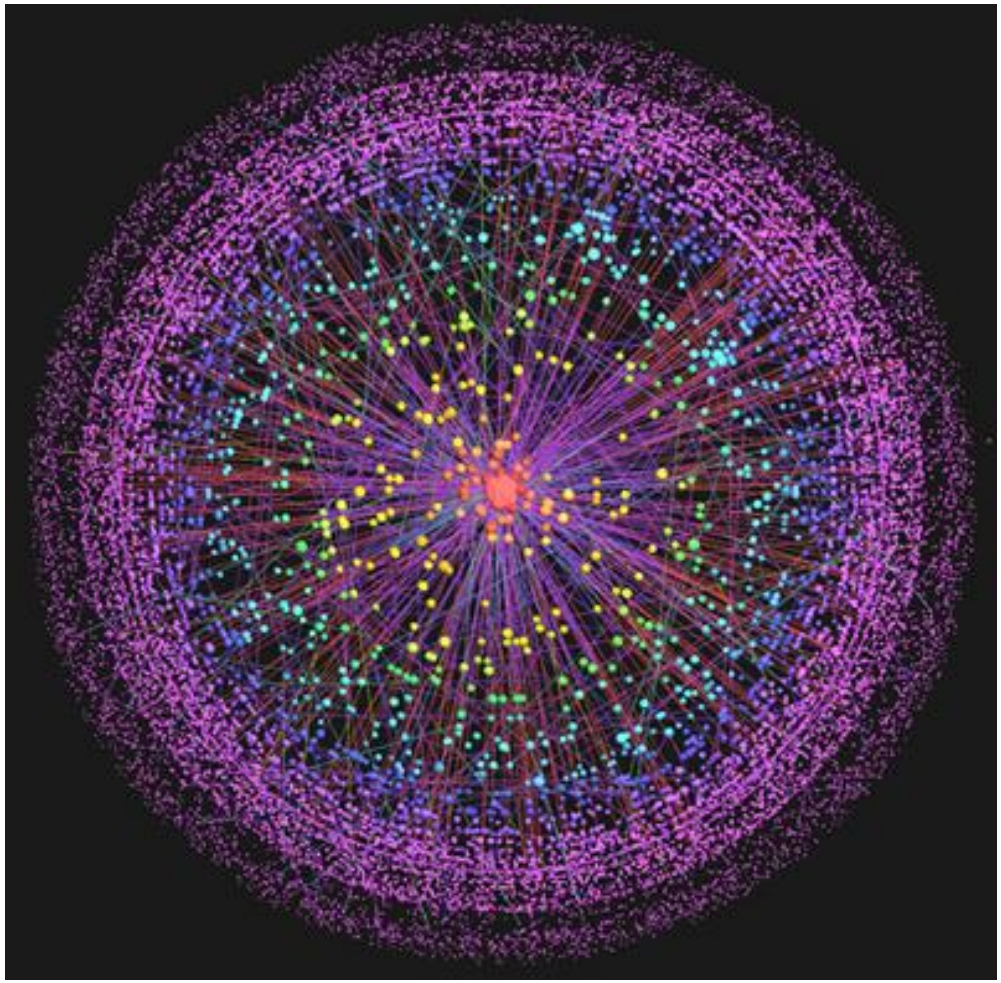
The Hubble telescope: $2.5 billion

Source: David Lazer

CERN particle accelerator: $1 billion/year

Source: David Lazer

The Web: priceless*



Source: David Lazer

** Apologies to MasterCard*

SOCIAL SCIENCE

Computational Social Science

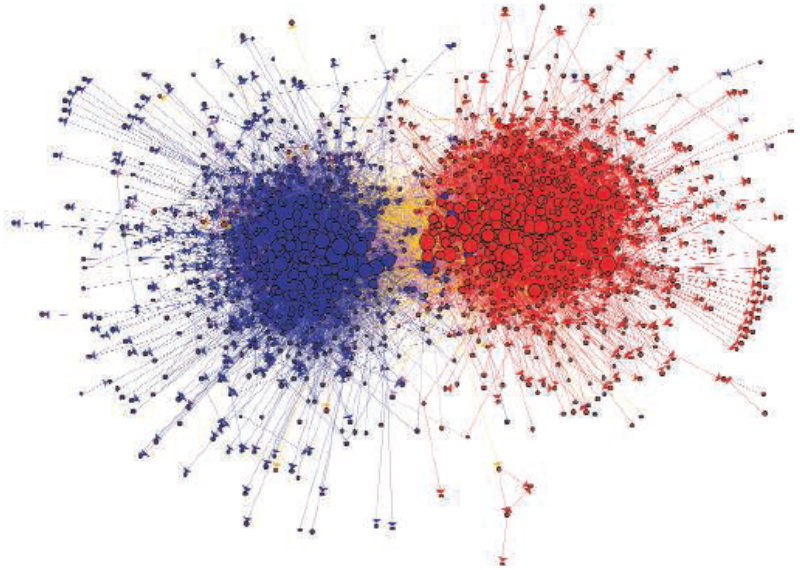
A field is emerging that leverages the capacity to collect and analyze data at a scale that may reveal patterns of individual and group behaviors.

David Lazer,[1] Alex Pentland,[2] Lada Adamic,[3] Sinan Aral,[2,4] Albert-László Barabási,[5] Devon Brewer,[6] Nicholas Christakis,[1] Noshir Contractor,[7] James Fowler,[8] Myron Gutmann,[3] Tony Jebara,[9] Gary King,[1] Michael Macy,[10] Deb Roy,[2] Marshall Van Alstyne[2,11]

We live life in the network. We check our e-mails regularly, make mobile phone calls from almost any location, swipe transit cards to use public transportation, and make purchases with credit cards. Our movements in public places may be captured by video cameras, and our medical records stored as digital files. We may post blog entries accessible to anyone, or maintain friendships through online social networks. Each of these transactions leaves digital traces that can be compiled into comprehensive pictures of both individual and group behavior, with the potential to transform our understanding of our lives, organizations, and societies.

The capacity to collect and analyze massive amounts of data has transformed such fields as biology and physics. But the emergence of a data-driven "computational social science" has been much slower. Leading journals in economics, sociology, and political science show little evidence of this field. But computational social science is occurring—in Internet companies such as Google and Yahoo, and in government agencies such as the U.S. National Security Agency. Computational social science could become the exclusive domain of private companies and government agencies. Alternatively, there might emerge a privileged set of academic researchers presiding over private data from which they produce papers that cannot be critiqued or replicated. Neither scenario will serve the long-term public interest of accumulating, verifying, and disseminating knowledge.

What value might a computational social science—based in an open academic environment—offer society, by enhancing understanding of individuals and collectives? What are the



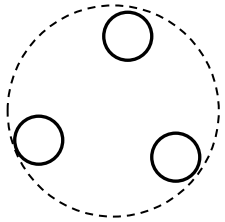
Data from the blogosphere. Shown is a link structure within a community of political blogs (from 2004), where red nodes indicate conservative blogs, and blue liberal. Orange links go from liberal to conservative, and purple ones from conservative to liberal. The size of each blog reflects the number of other blogs that link to it. [Reproduced from (*8*) with permission from the Association for Computing Machinery]

[1]Harvard University, Cambridge, MA, USA. [2]Massachusetts Institute of Technology, Cambridge, MA, USA. [3]University of Michigan, Ann Arbor, MI, USA. [4]New York University, New York, NY, USA. [5]Northeastern University, Boston, MA, USA. [6]Interdisciplinary Scientific Research, Seattle, WA, USA. [7]Northwestern University, Evanston, IL, USA. [8]University of California–San Diego, La Jolla, CA, USA. [9]Columbia University, New York, NY, USA. [10]Cornell University, Ithaca, NY, USA. [11]Boston University, Boston, MA, USA. E-mail: david_lazer@harvard.edu. Complete affiliations are listed in the supporting online material.

Published by AAAS

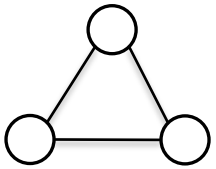
Four Levels of Influences on Team Assembly

Compositional Level



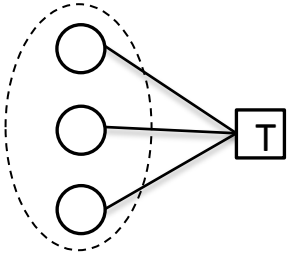
(a) Team as a collection of individuals

Relational Level



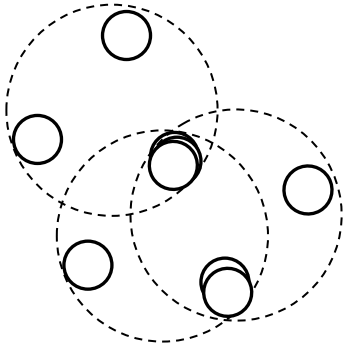
(b) Team as individuals and relations

Multimodal Network Level



(c) Team as a network of individuals and tasks

Ecosystem Level



(d) Ecosystem of teams

Individual

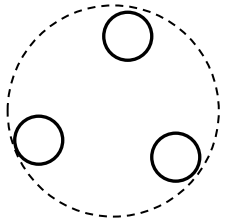
T Task

NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

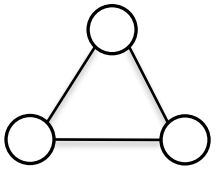
Four Levels of Influences on Team Assembly

Compositional Level



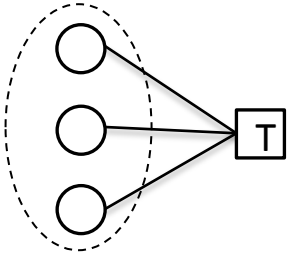
(a) Team as a collection of individuals

Relational Level



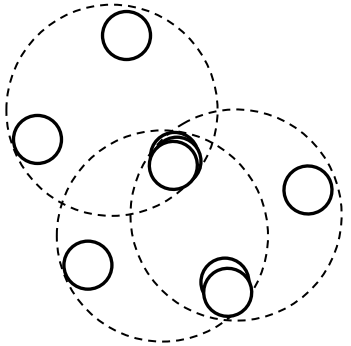
(b) Team as individuals and relations

Multimodal Network Level



(c) Team as a network of individuals and tasks

Ecosystem Level



(d) Ecosystem of teams

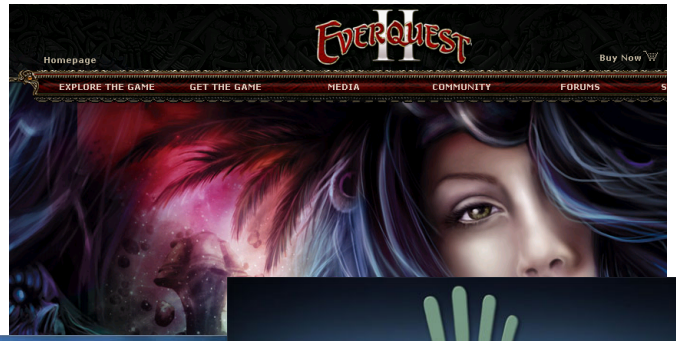
Individual

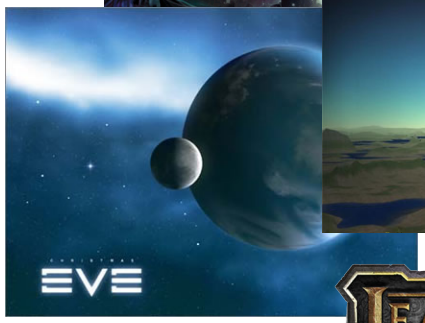
T Task

NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

Virtual World Exploratorium





vwobservatory.org

NORTHWESTERN UNIVERSITY

USC UNIVERSITY OF SOUTHERN CALIFORNIA

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

SONY ONLINE ENTERTAINMENT

Linden Lab

SHANDA GAMES

KINGSOFT

HELSINKI INSTITUTE FOR INFORMATION TECHNOLOGY

復旦大學

NORTHWESTERN UNIVERSITY

advancing the science of networks in communities

Our Dataset

- Data from a popular Massively Multiplayer Online Role Playing Game (MMORPG) EverQuest II (EQ2)
  - Fantasy based game
  - Server-side records
  - Player attributes, activities, and relations
  - Focus on Combat Teams
  - Players are “nested” within teams

http://everquest2.station.sony.com/screenshots.vm

Combat Groups in EverQuest II

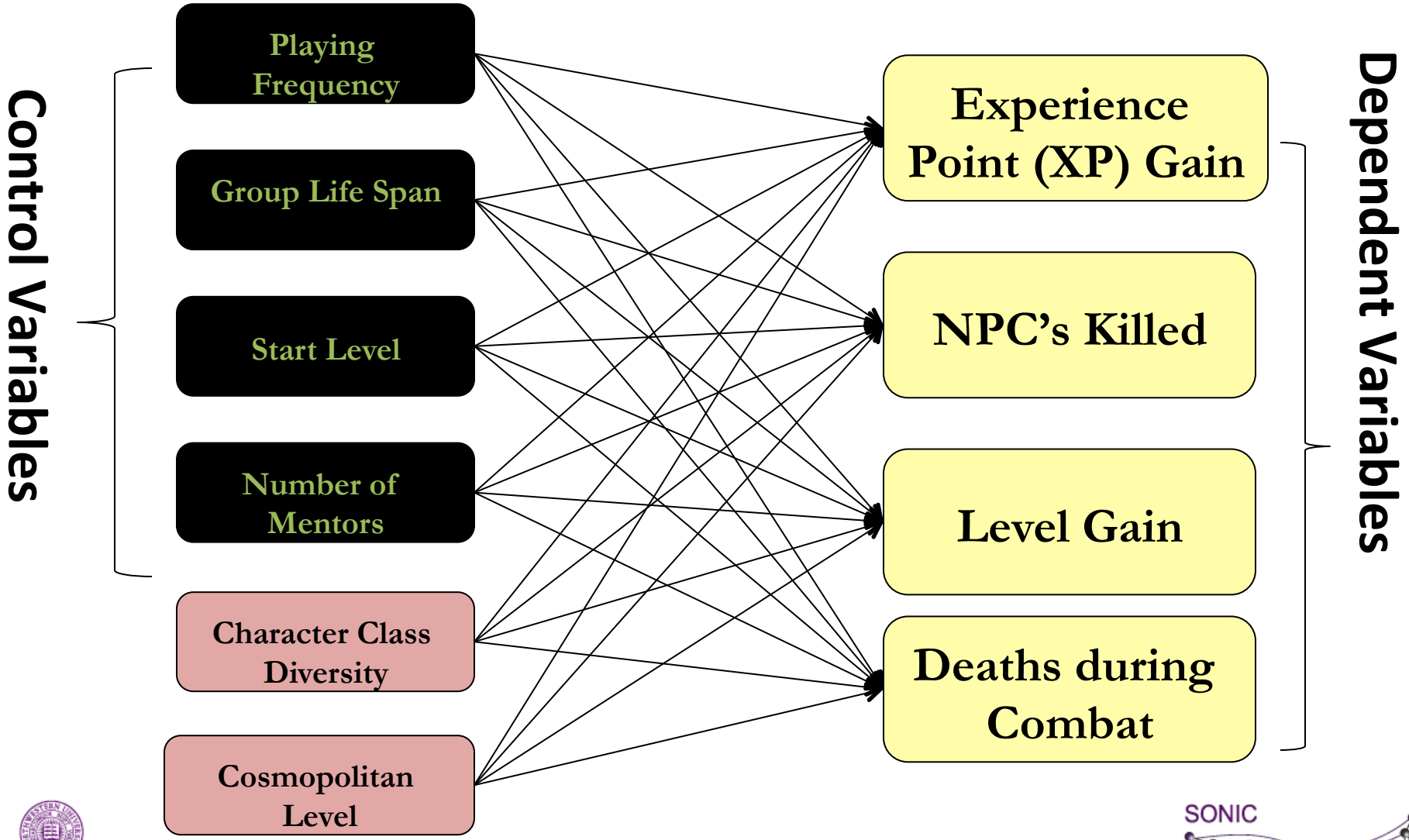
- Difficult combat tasks require collaboration of multiple players and assembly of combat groups
- From 2006-08-27 to 2006-09-11 on Antonia Bayle Server
  - 8,423 players
  - 46,393 groups
  - 9,436,741 combat related records

What makes a group successful?

- **Group Diversity**
  - Four character classes in the game: Fighter, Mage, Scout and Priest, each having a different role in a group
  - Measure Group Diversity: Blau's Index

- **Group member's cosmopolitan level**
  - Group members being involved in multiple different groups

Effects of Group Attributes on Performance Measures



Regression Analysis Results on Combat Groups of Four Players

	XP	NPC's	Level Gain	Deaths
Constant	-20939.926**	-3.376 (.361)	.717**	4.011**
Frequency	-1553.494 (.105)	4.127**	-.010 (.816)	.601**
Life Span	736.797**	1.174**	.015**	.063**
			3**	-.050**
Num of Mentor			2**	-.647**
Diversity	20819.998**	14.342**	.726**	-1.873 (.095)
Member Cosmo.	30.254 (.612)	-.025 (.698)	-.010**	-.032**
R^2	.666	0.821	0.371	0.244
F	595.213 (p=.000)	1368.793 (p=.000)	176.071 (.000)	96.274 (.000)

Diversity helps the groups to achieve more.

Members being cosmopolitan doesn't help with gains but helps to avoid loss.

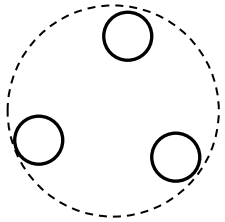
** indicates significant results at .01 level

NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

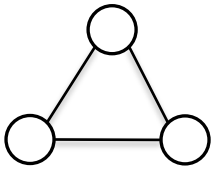
Four Levels of Influence on Team Assembly

Compositional Level



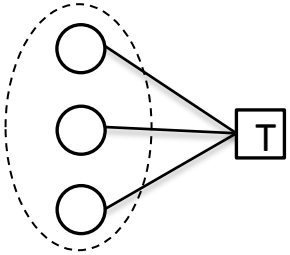
(a) Team as a collection of individuals

Relational Level



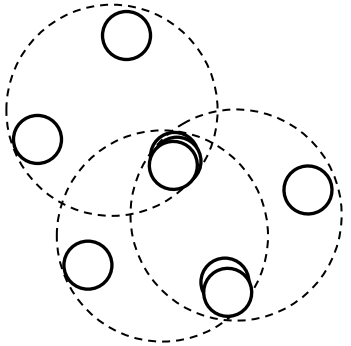
(b) Team as individuals and relations

Multimodal Network Level



(c) Team as a network of individuals and tasks

Ecosystem Level



(d) Ecosystem of teams

Individual

T Task

NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

Relational Influences on nanoHUB Team Assembly

- Outcome variables
  - Co-contribution network(s)
- Explanatory variables
  - Contributor attributes
  - Network structures
  - Covariate networks (co-authorship and citation)
  - Positions in co-authorship and citation networks
- Methods: p*/Exponential random graph model

Relational Influences on nanoHUB Team Assembly

Co-contribution in …	Teams (>250 users)	Teams (<250 users)
Female	0.16 (.20)	0.17 (.21)
Same country origin	-0.01 (.21)	0.17 (.17)
Same university	**0.86*** (.10)	**1.59*** (.14)
H-index	**-0.04*** (.01)	**-0.05**** (.02)
H-index difference	**0.04*** (.02)	**0.10*** (.03)
Publication difference	-0.002 (.002)	**-0.009*** (.003)
Co-author relation (Ln)	**1.69*** (.39)	**1.39*** (.53)
Citation relation (Ln)	0.36 (.29)	**1.46*** (.37)
Control:		
Purdue	**-0.39*** (.09)	**-0.26*** (.10)
NCN	**0.57*** (.14)	**1.16*** (.20)
Edge	**-3.69*** (.50)	**-2.05*** (.53)
Alternating stars	**-1.51*** (.12)	**-2.14*** (.18)
Alternating triangles	**3.62*** (.21)	**3.13*** (.18)
N	87	118

Note: * p<.10, ** p<.05, *** p<.01

NORTHWESTERN UNIVERSITY

SONIC advancing the science of networks in communities

I.2 Motivations for creating network links for Team Assembly

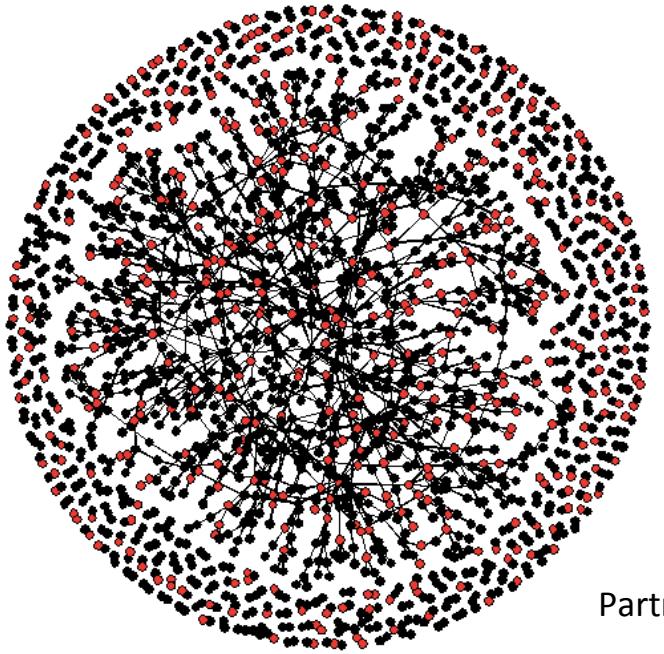
Yun Huang
Mengxiao Zhu
Brian Keegan
Jeff Treem
Noshir Contractor

Network Data on Team Assembly

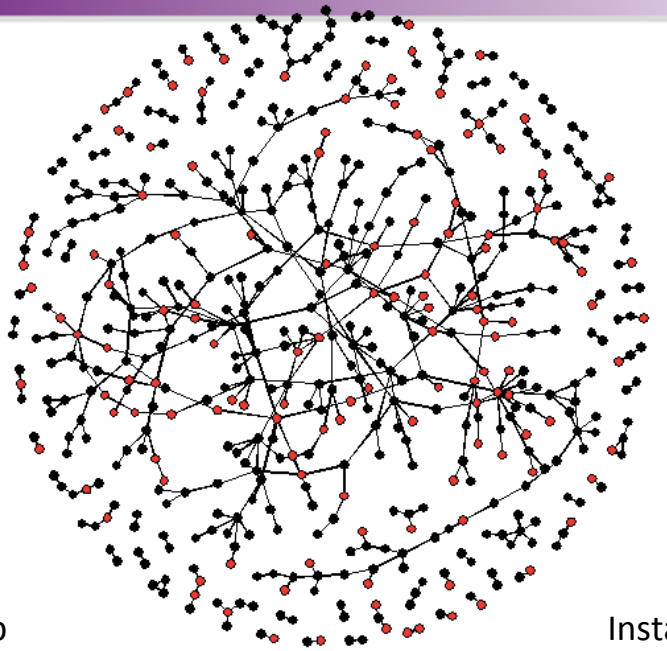
NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

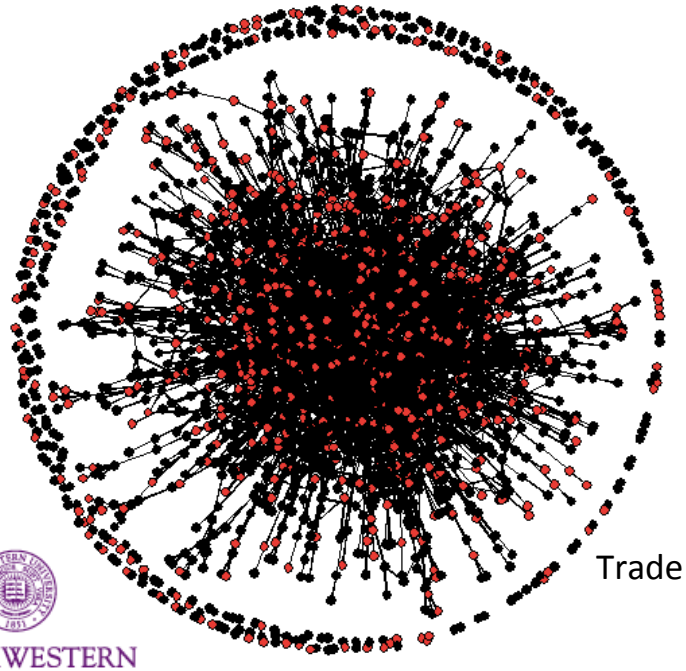
Black: male
Red: female



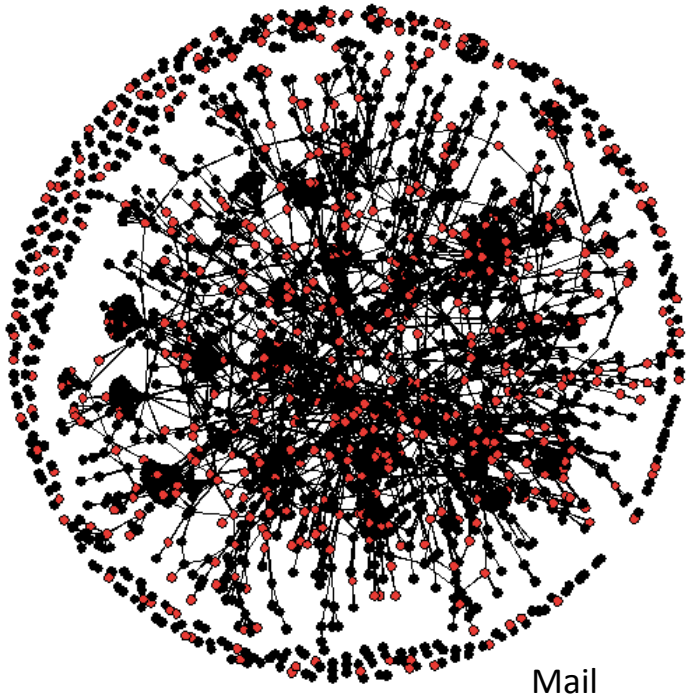
Partnership



Instant messaging



Trade



Mail

NORTHWESTERN UNIVERSITY

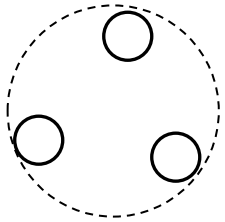
SONIC
advancing the science of networks in communities

Results

- Selectivity and transitivity (friend of a friend) exists in all online relations.
- Homophily of age and game experience is supported in all four relations.
- Distance matters but short distances are more important. Individuals living within 50 Km are 22.6 times more likely to be partners than those who live between 50 and 800 Km.
- Time zones impacts gaming and trading but not IM and mail. Individuals in the same time zone are 1.25 times more likely to be game partners than the individuals with one hour difference (but no time zone effect for
- Gender homophily is not supported for all relations and female players are more likely to interact with the male players.

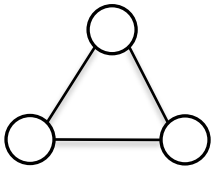
Four Levels of Influence on nanoHUB Team Assembly

Compositional Level



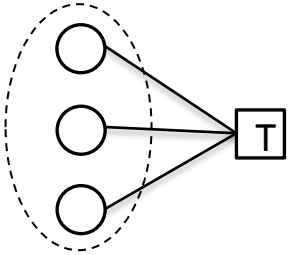
(a) Team as a collection of individuals

Relational Level



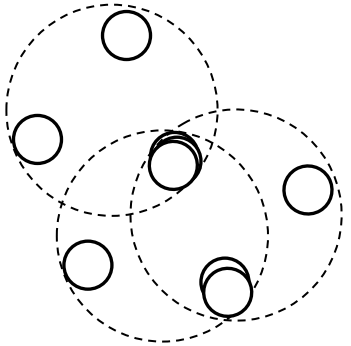
(b) Team as individuals and relations

Multimodal Network Level



(c) Team as a network of individuals and tasks

Ecosystem Level



(d) Ecosystem of teams

Individual

T Task

NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

Multimodal influences on nanoHUB Team Assembly

- Outcome variables
  - Team affiliation network(s)
- Explanatory variables
  - Contributor attributes
  - Team attributes
  - Network structures
  - Positions in co-authorship and citation networks
- Methods: p*/BPnet

Multimodal influences on nanoHUB Team assembly

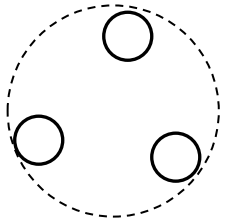
	Teams (>250 users)	Teams (<250 users)
Female	-0.24 (.48)	-0.18 (.33)
Same country origin	-0.07 (.13)	**0.20**** (.10)
Different university	**-0.53***** (.09)	**-1.57***** (.13)
H-index	-0.01 (.01)	0.006 (.02)
H-index difference	0.007 (.008)	0.01 (0.01)
Publication difference	-0.001 (.001)	-0.003 (.002)
Team:		
Tool difficulty	0.05 (.18)	**0.39**** (.16)
Open source	**-1.57***** (.53)	-0.71 (.67)
Ratings (Binary)	0.15 (.27)	0.02 (.21)
Num citations (Ln)	**0.67***** (.18)	-0.06 (.27)
Num users (Ln)	-0.27 (.23)	0.001 (.12)
Control:		
Purdue	**-1.01***** (.28)	**-1.22***** (.16)
NCN	**2.89***** (.45)	**2.51***** (.33)
Edge	0.31 (2.01)	0.17 (1.04)
Contributor stars	**-0.96***** (.30)	**-0.97***** (.22)
Team stars	-0.06 (.61)	**-1.12**** (.53)

NORTHWESTERN UNIVERSITY

SONIC advancing the science of networks in communities

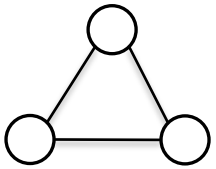
Four Levels of Influence on Team Assembly

Compositional Level



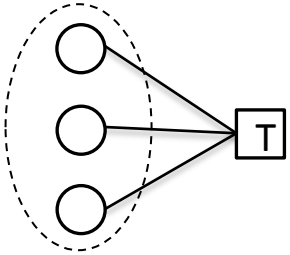
(a) Team as a collection of individuals

Relational Level



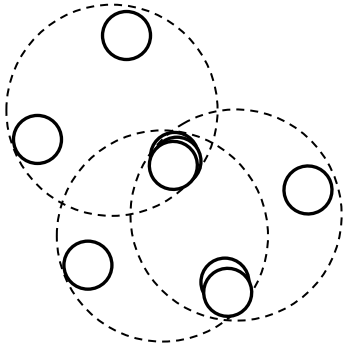
(b) Team as individuals and relations

Multimodal Network Level



(c) Team as a network of individuals and tasks

Ecosystem Level



(d) Ecosystem of teams

Individual

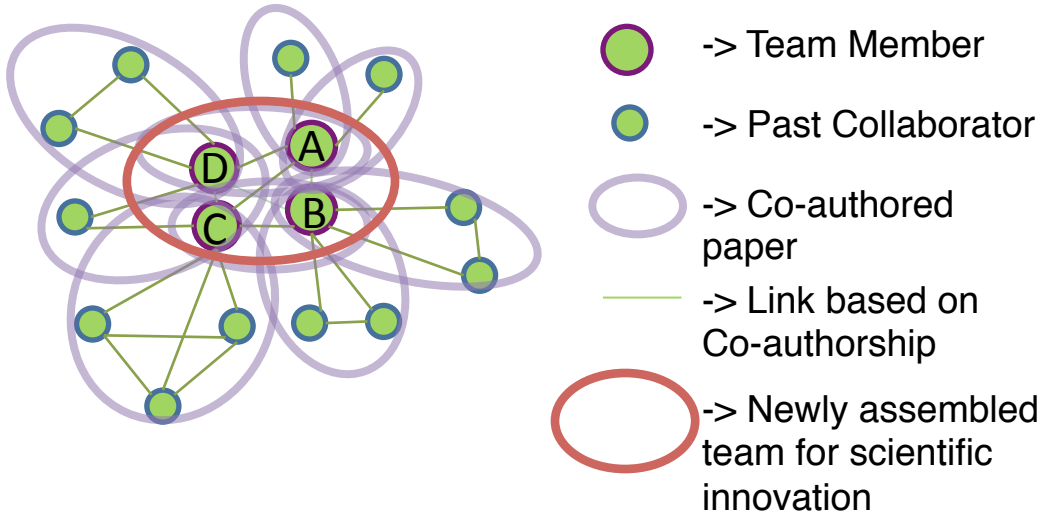
T Task

NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

Scientific Ecosystem as Antecedent of Team Assembly and Performance

- Teams do not assemble in a "vacuum"
- Teams emerge from networks of prior collaborations in a particular space
  - An "ECOSYSTEM"

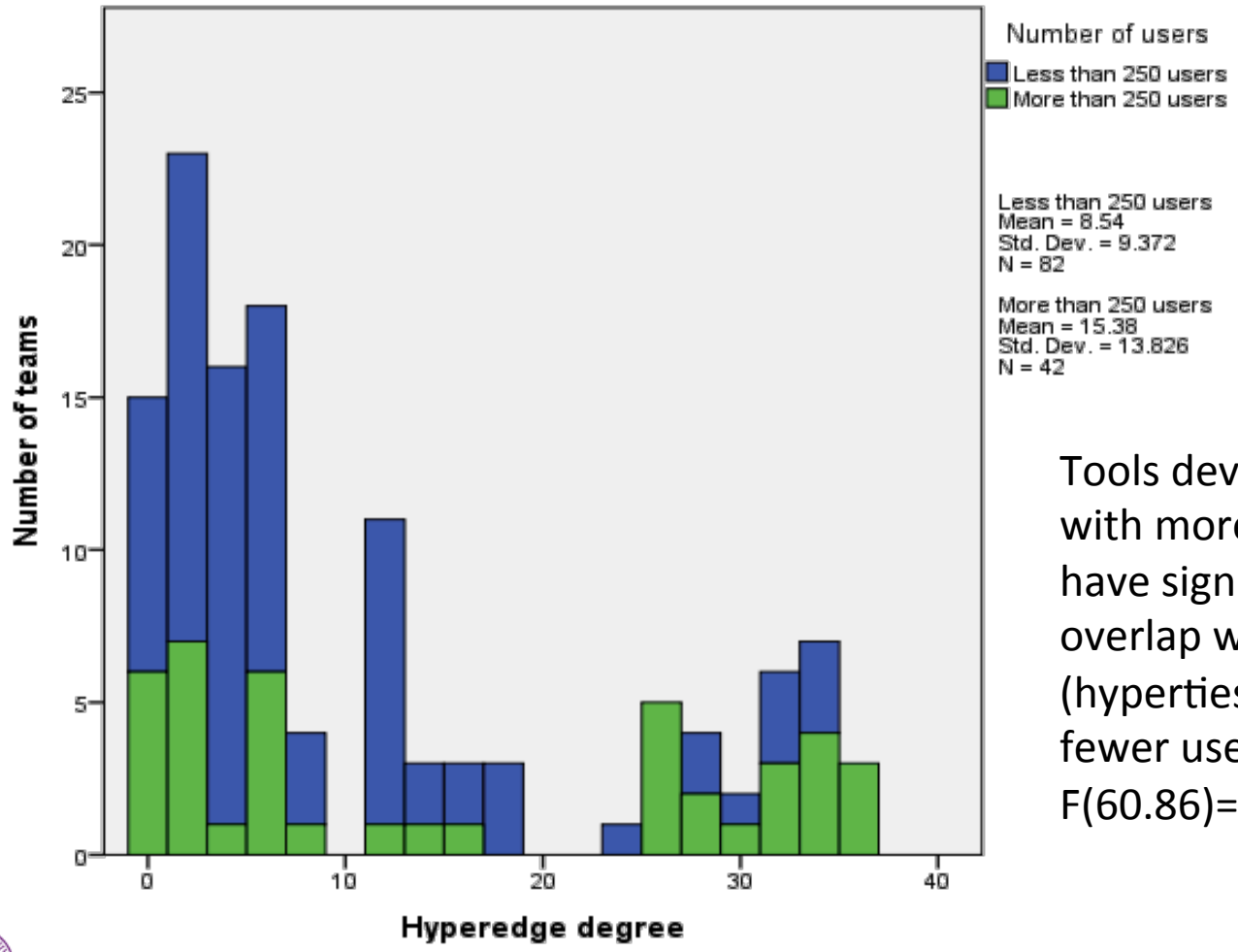


- Are there certain characteristics of the scientific ecosystem that lead to team assembly?
- Do variations in these ecosystem characteristics predict team performance?

Ecosystem influences on nanoHUB Team Assembly

- Target network statistics
  - Team hyperties
    - Number of teams with which the focal team has overlapping members (i.e. hypertie degree).
  - Closure of team hyperties
    - Ratio of overlap among teams with which focal team has overlapping members. (i.e. mean clustering coefficients)
- Methods: Estimation by comparison to distribution generated by simulating hypergraphs

Ecosystem influence on nanoHUB Team Assembly



Tools developed by teams with more than 250 users have significantly more overlap with other teams (hyperties) than tools with fewer users
F(60.86)=-2.89, p=0.005.

NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

Ecosystem influences on assembly of clinical and translational science teams

- Bibliographic data on teams of researchers who submitted research proposals to a grant competition funded by NUCATS Pilot Grants
- 100 research proposal teams in original data set
  - Total of 147 participants.
  - 7 teams awarded funding (i.e., "Successful" teams)
  - 47 single author proposals excluded from sample
  - 53 coauthored proposals
- 11 teams included in current analyses
  - 4 funded, 7 unfunded

Impact of ecosystem on assembly of successful teams

Research Question 1: How does the presence of established "*key teams*" that dominate the intellectual discourse influence assembly of new teams?

They inhibit the assembly of successful teams in the ecosystem

Impact of ecosystem on assembly of successful teams

Research Question 2: How does the coherence of the "*intellectual neighborhood*" impact assembly of successful teams?

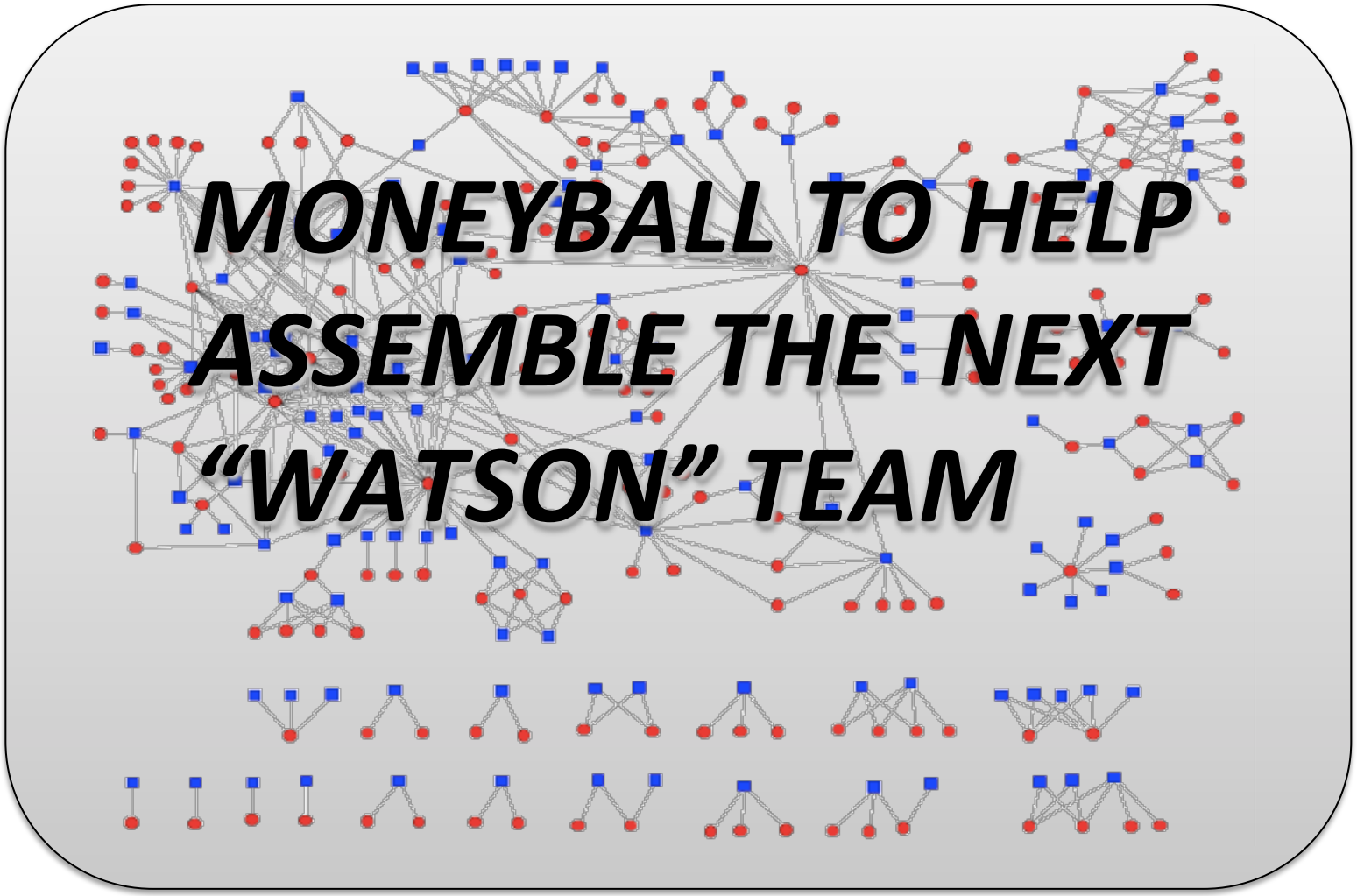
Successful teams are more likely to assemble when there is modest overlap in team membership within the ecosystem

Impact of ecosystem on assembly of successful teams

Research Question 3: How does the coherence of the "*immediate* intellectual neighborhood" impact assembly of successful teams?

*Teams are more likely to assemble when their immediate neighborhood has **lower** overlap in team membership than the entire ecosystem*

From Understanding to Enabling Team Assembly Or ….



"Your goal shouldn't be to buy players. Your goal should be to buy wins. In order to buy wins, you need to buy runs." (Bakshi & Miller, 2011).

Demo

- Intra-university Research Networking: NUCATS Semantic C-IKNOW
- Inter-university Research Networking: VIVO C-IKNOW
- Dream Team Builder

Perfect Storm ….

- We are well poised to enable team assembly by leveraging recent advances in:
  - Theories: Theories about the socio-technical motivations for creating, maintaining, dissolving and re-creating links to engage in teams
  - Methods: An ensemble of qualitative and quantitative methods techniques (such as exponential random graph modeling or *p**) to understand and enable theoretically grounded network recommendations for teams
  - Data: The development of **Virtual Web Observatories** provide the technological capability to capture, store , merge, and query relational metadata about data and tools needed to more effectively understand and enable teams.
  - Computational infrastructure: Cloud computing and petascale applications are critical to face the computational challenges in understanding and enabling teams.

Acknowledgements