

# A SOCIAL NETWORK ANALYSIS OF THE GENI COMMUNITY

*GENI ENGINEERING CONFERENCE 14*

---

**Laurie J. Kirsch**

Professor

Information Systems

Katz Graduate School of Business

University of Pittsburgh

[lkirsch@katz.pitt.edu](mailto:lkirsch@katz.pitt.edu)

**Sandra A. Slaughter**

Alton M. Costley Chair & Professor

Information Technology Management

Scheller College of Business

Georgia Institute of Technology

[sandra.slaughter@mgt.gatech.edu](mailto:sandra.slaughter@mgt.gatech.edu)

# Who are we and why are we here?

- Management scholars who conduct research on the management of information technology and systems projects
- Prior and current NSF-funded studies to examine the management of large scientific research projects with a “cyber” component.
- Examples of “*cyber-infrastructure projects*” are:
  - **NEES** – George E. Brown, Jr. Network for Earthquake Engineering Simulation
  - **Teragrid / XD** – High-performance network of super computers providing cyber-infrastructure for open scientific research
  - **iPLANT** – cyber-infrastructure collaborative for plant sciences
  - **GENI** – Global Environment for Network Innovations

# Motivation

- Challenges we observe in managing cyber-infrastructure projects:
  - Large and complex
  - High uncertainty and risk
  - Volatile and emergent requirements
  - Constrained by budget and schedule
  - Distributed knowledge and collaborators across institutions
  - Diverse collaborators with different motives and incentives
  - Difficulty of communication, control and coordination

# Research Objectives for our Study of GENI

- Identify mechanisms that are needed to effectively structure, govern, and manage projects like GENI
- Understand interaction patterns (who is interacting with whom)
- Analyze the evolution of the GENI community
- Suggest control / coordination / communication techniques
- Get feedback from the GENI community

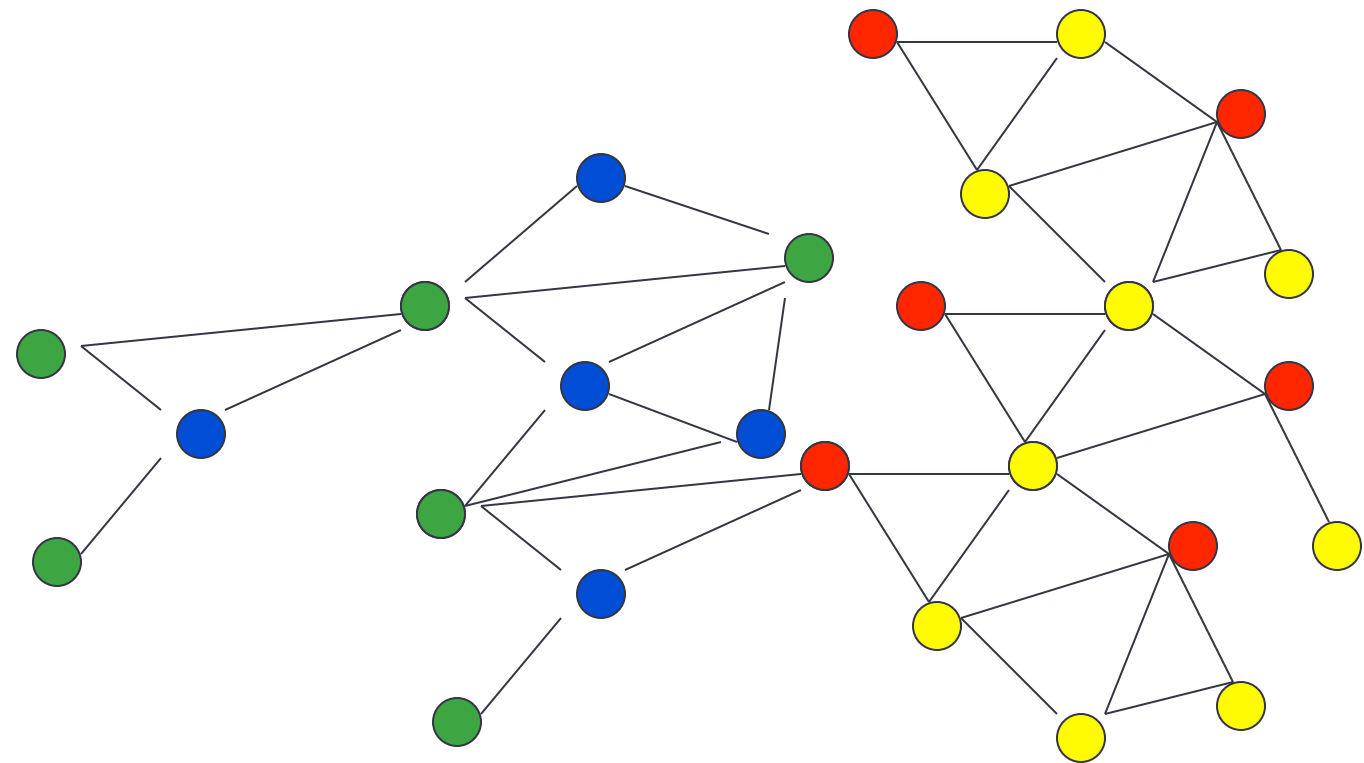
# Research Approach

- Qualitative
  - Attend GECs and conduct targeted interviews of GENI stakeholders to understand their experiences in the project
  - Synthesize recommendations for communication, coordination and commitment
- Quantitative
  - Social network analysis of patterns of interactions between GENI stakeholders

# What we have done

- Attended GECs 5-14
- Conducted interviews with ~ 50 GENI stakeholders across different roles (PI, Experimenter, Student, NSF, Company, GPO, etc.)
- Collected publicly available data capturing attendance of GENI stakeholders at GECs and obtained additional data via web searches
- Conducted a social network analysis of the GEC data

# Research Analysis & Findings: A Social Network Analysis of GECs



# A Social Network Analysis of GENI GECs

- **Social network analysis identifies**
  - communication and interaction patterns of individuals in a community
  - the most “central” (e.g. influential) individuals and those who are the information brokers (e.g., boundary spanners)
  - sub-groups and cliques, which can signal the potential for conflict
- **Patterns of communication and interaction**
  - reflect how information, knowledge and ideas are exchanged in a community
  - relate to social identity, performance, innovation and other project outcomes



# A Social Network Analysis of GENI GECs (cont'd)

- **Goal:** Ascertain the impact of GEC meetings on GEC attendees, the GENI community and GENI as a whole
  - GECs provide opportunities for attendees to interact face to face
  - GECs serve as a “window” on the GENI community

# A Social Network Analysis of GENI GECs (cont'd)

- **Data:**

- Attendee lists for GECs 1-13
  - Contain name and affiliation for each attendee
- GENI website (wiki)
  - Contains project and cluster information used to identify the “role” and “area” of attendees

# Analyses Performed by:

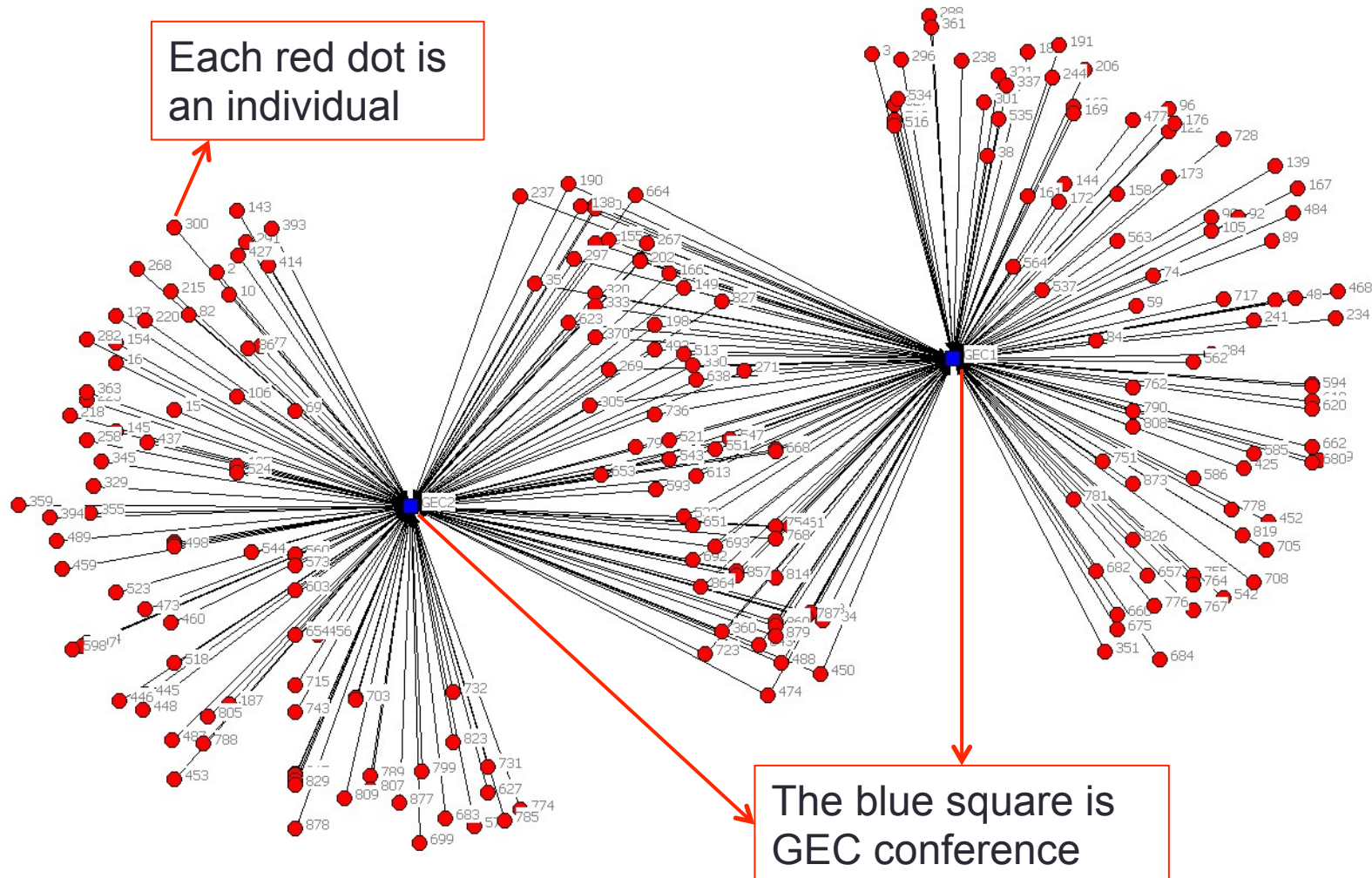
- Person-GEC\*\*
- Area-GEC
- Role-GEC\*\*
- Person-Institution
- Institution-GEC\*\*

**\*\* We will share selected results from these analyses now, but the full results are available on the GENI wiki**

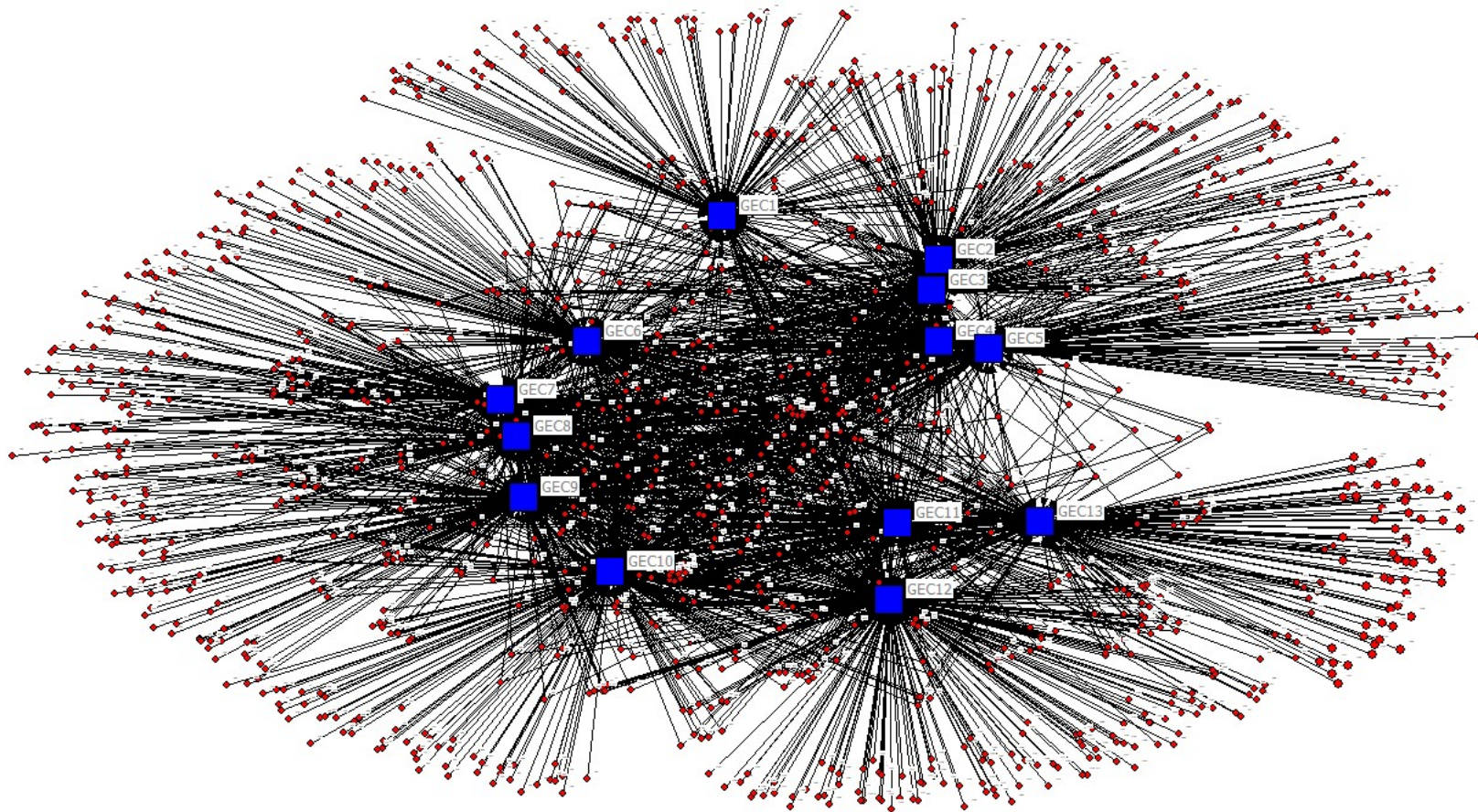
# Person-GEC Analysis

- The charts show people's GEC attendance.
- They represent cumulative attendance since GEC1.
- Red dots represent people. Blue squares represent GECs.
- An individual is 'related' to a GEC conference if he/she attended the conference.
- People who locate at the center of the chart tend to be the ones who are active in attending the GEC(s).

# GEC1~2

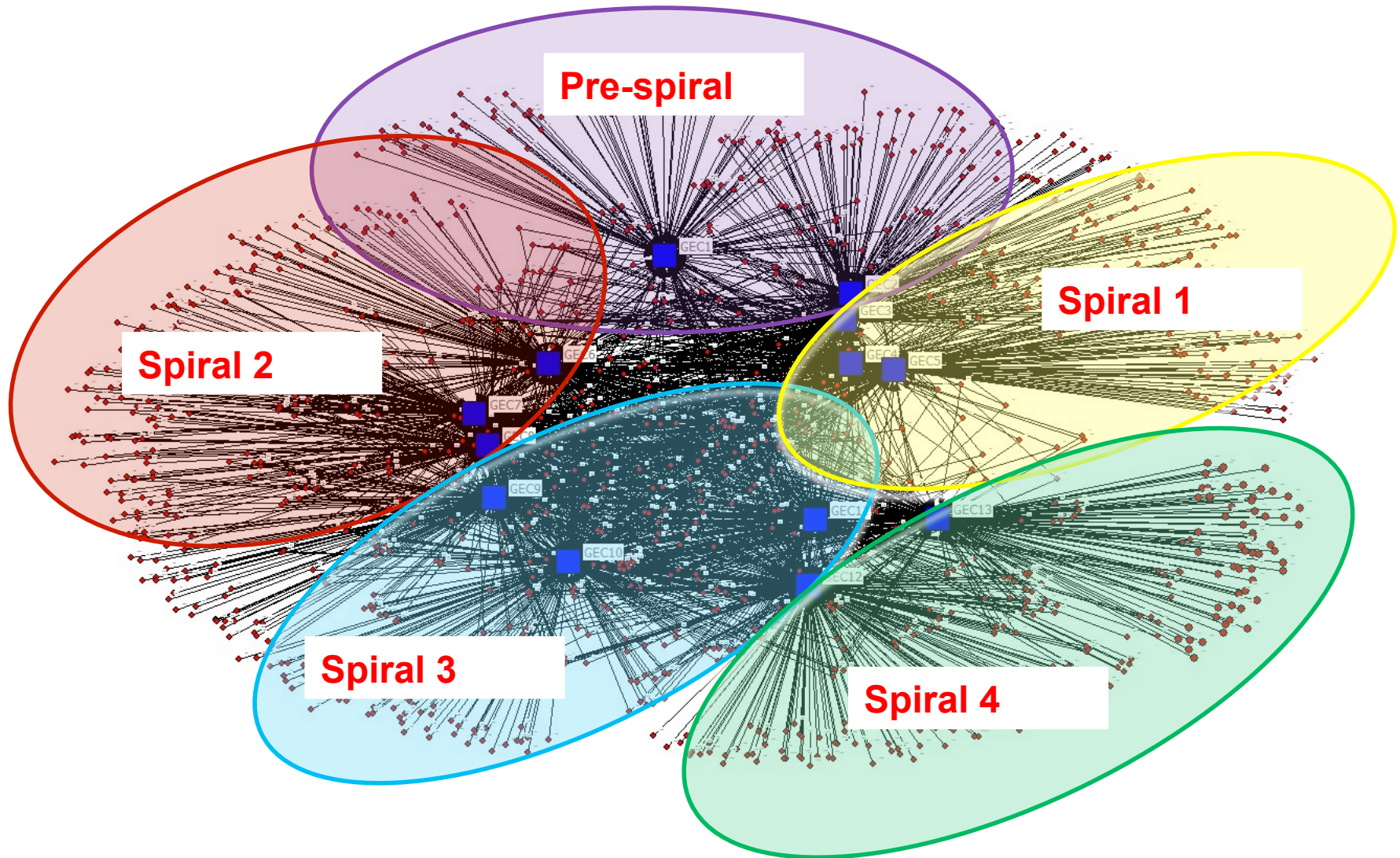


# GEC1~13





# GEC1~13



# GEC Attendance

GEC #	location	time	total # of attendees	attend future	% attend future	attend multiple	% attend multiple	attend next	% attend next	new comers	% new comers	previous comers	% previous comer
1	Minneapolis, MN	2007.1	148	92	62.16%	92	62.16%	63	42.57%	148	100.00%	0	0.00%
2	Arlington, VA	2008.3	150	94	62.67%	108	72.00%	67	44.67%	87	58.00%	63	42.00%
3	Palo Alto, CA	2008.1	193	123	63.73%	144	74.61%	92	47.67%	110	56.99%	83	43.01%
4	Miami, FL	2009.3	168	122	72.62%	143	85.12%	74	44.05%	61	36.31%	107	63.69%
5	Seattle, WA	2009.7	170	113	66.47%	126	74.12%	82	48.24%	76	44.71%	94	55.29%
6	Salt Lake, UT	2009.1	223	157	70.40%	174	78.03%	96	43.05%	100	44.84%	123	55.16%
7	Durham, NC	2010.3	216	158	73.15%	170	78.70%	115	53.24%	92	42.59%	124	57.41%
8	San Diego, CA	2010.7	247	172	69.64%	206	83.40%	125	50.61%	82	33.20%	165	66.80%
9	D.C.	2010.1	274	179	65.33%	219	79.93%	129	47.08%	88	32.12%	186	67.88%
10	Puerto Rico	2011.3	260	154	59.23%	206	79.23%	107	41.15%	88	33.85%	172	66.15%
11	Denver, CO	2011.7	201	138	68.66%	177	88.06%	117	58.21%	48	23.88%	153	76.12%
12	Kansas City, MO	2011.1	317	124	39.12%	218	68.77%	124	39.12%	126	39.75%	191	60.25%
13	Los Angeles, CA	2012.3	232			169	72.84%			62	26.72%	170	73.28%

## How to read the numbers – using GEC1 as an example:

92 (or 62.16% of total) attendees attended future GECs;  
 92 (or 62.16% of total) attendees attended multiple GECs;  
 63 (or 42.57% of total) attendees attended the next GEC (i.e. GEC2);  
 148 (or 100% of total) attendees are new to this GEC;  
 0 (or 0.00% of total) attendees attended a prior GEC.



# Observations from the Person-GEC Analysis

- Well developed core has emerged of around 150 individuals who have attended 5 or more meetings.
- Several GECs seem to attract the same types of attendees.
- Each GEC draws in a new and distinct group of attendees, that who have typically attended a future GEC.
- 1160 individuals attended at least 1 GEC; the average attendance by an individual is 2.4 GECs.

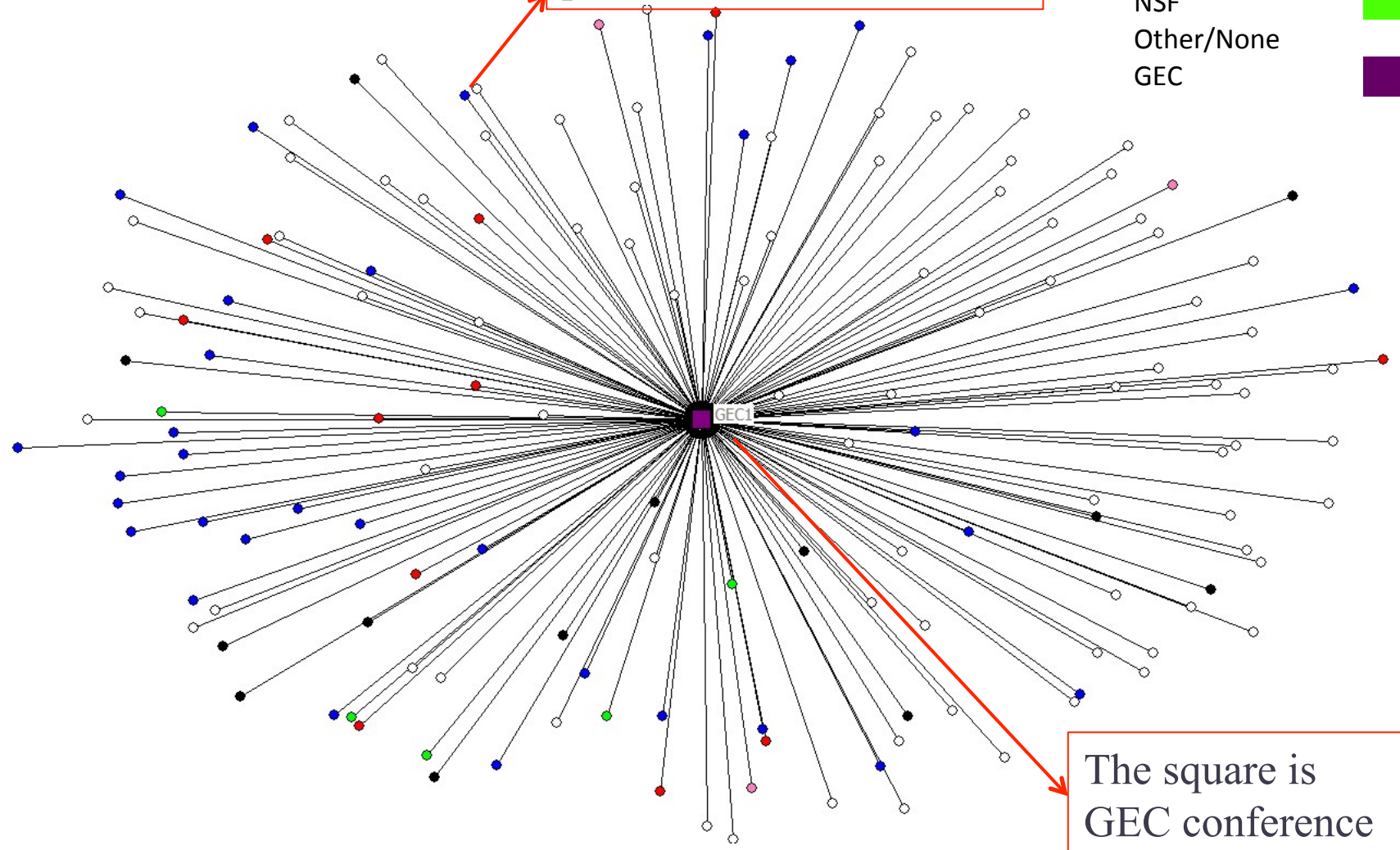
# Role-GEC Analysis

- The charts show people's GEC attendance and their roles in GENI:
  - GPO - employee of BBN
  - NSF – member of NSF
  - PI – principle investigator of a grant
  - Experimenter – any individual using GENI to run an experiment
  - Worker – any individual that works on a project, without being a PI
  - Other – attended a conference, but has no identifiable role
- The charts represent cumulative attendance since GEC1.
- Dots represent people. Blue squares represent GECs.
- People who locate at the center of the chart tend to be the ones who are active in attending GEC(s).

# GEC1

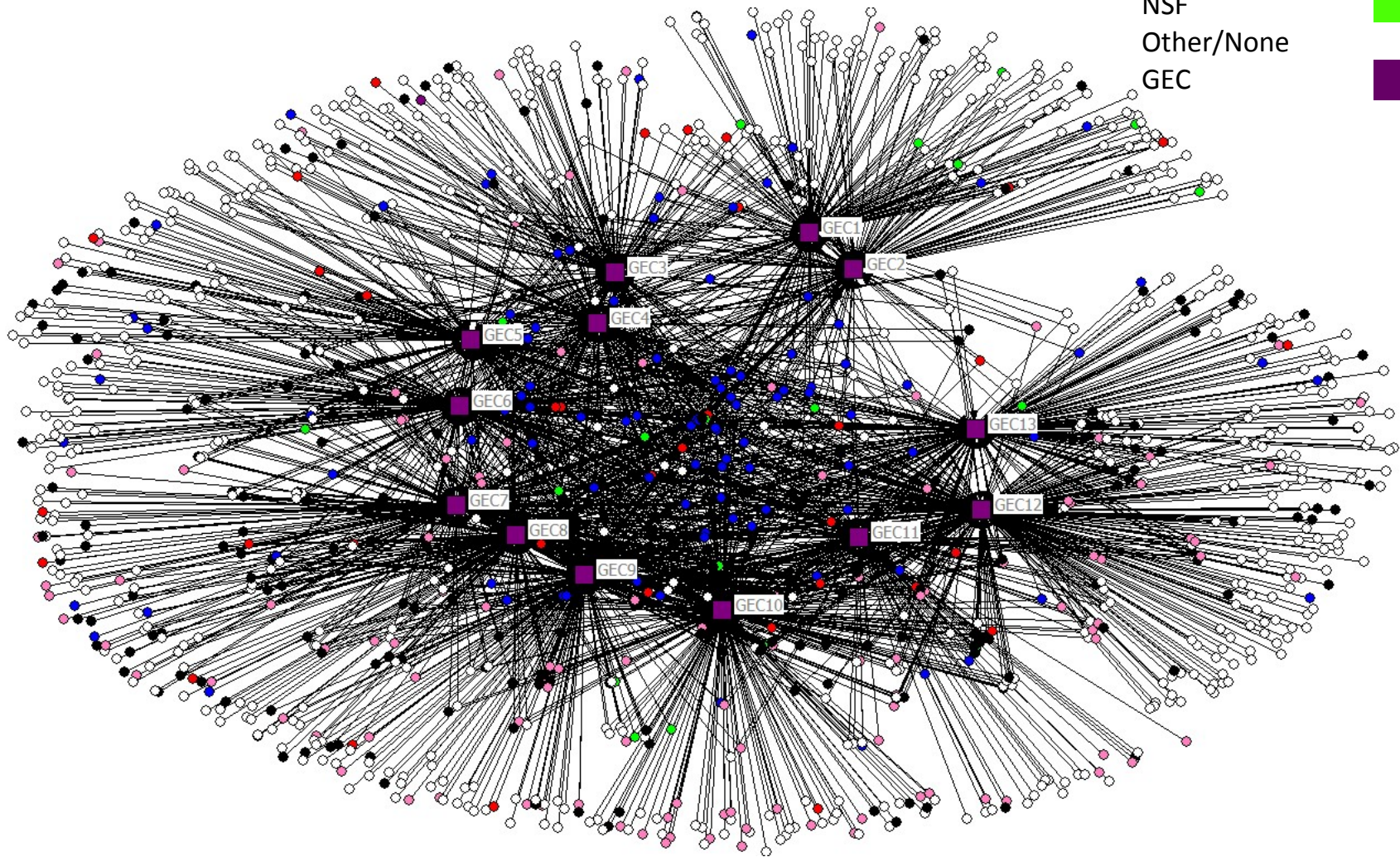
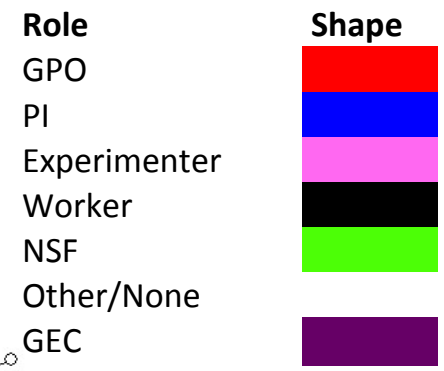
The dot is an individual;  
the color shows the  
person's role in GENI.

Role	Shape
GPO	Red
PI	Blue
Experimenter	Pink
Worker	Black
NSF	Green
Other/None	White
GEC	Purple



The square is  
GEC conference

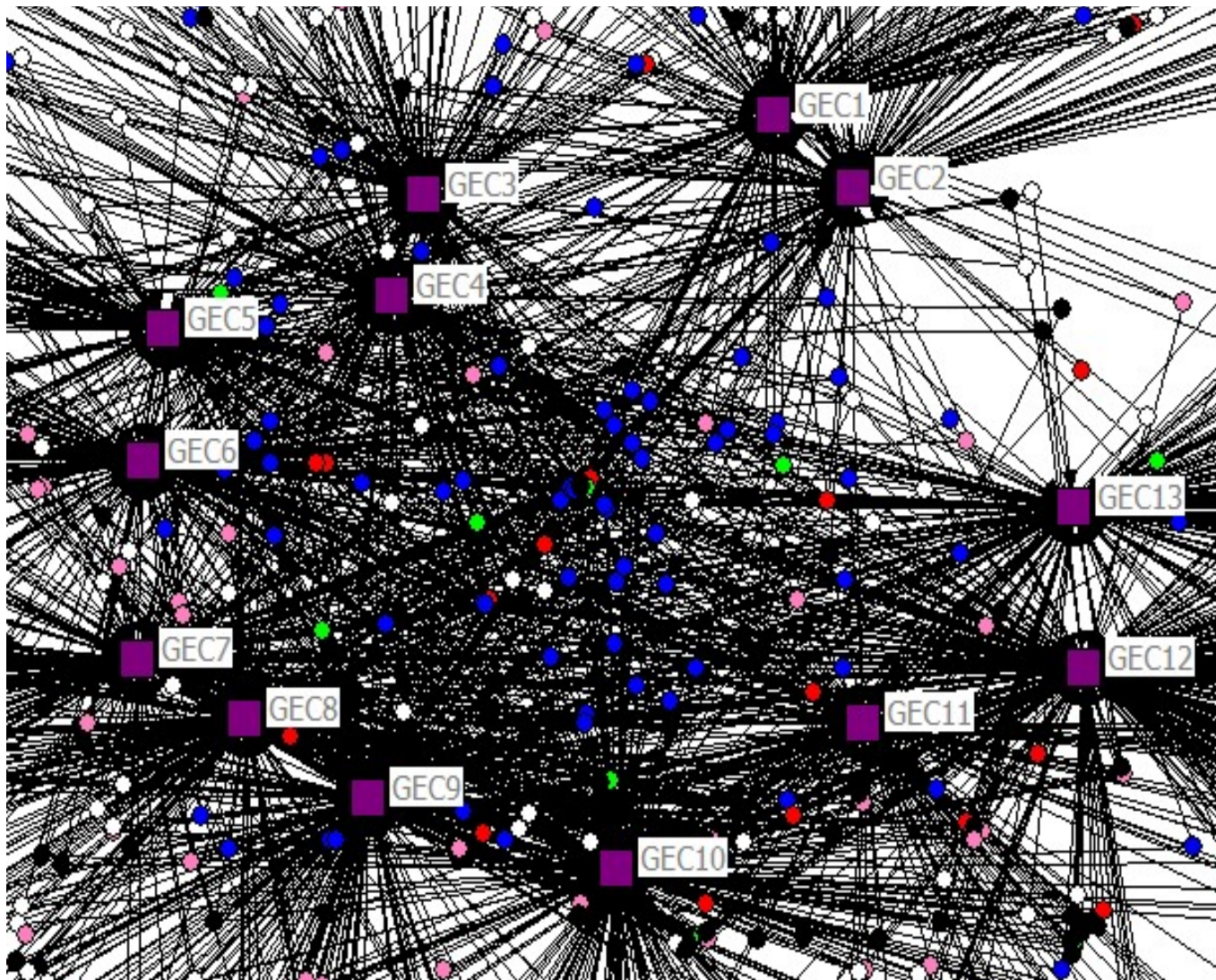
# GEC1~13





# GEC1~13 (Core)

Role	Shape
GPO	Red
PI	Blue
Experimenter	Pink
Worker	Black
NSF	Green
Other/None	White
GEC	Purple



# Some numbers

GEC#	GPO	NSF	PI	Worker	Experimenter	Other
1	11	5	32	13	3	84
2	11	11	35	14	4	75
3	12	5	46	27	14	89
4	10	5	46	37	11	59
5	10	5	37	43	15	59
6	14	6	53	55	19	75
7	17	5	39	77	21	55
8	21	7	46	82	26	64
9	19	8	44	67	45	81
10	20	6	41	73	43	77
11	20	5	42	57	31	46
12	18	6	49	79	31	133
13	17	3	40	65	22	85
1~13	200	77	550	689	285	982
%	7.19%	2.77%	19.76%	24.76%	10.24%	35.29%



From GEC1 through GEC13, 19.76% of all attendees are PIs.

# Observation from the Role-GEC Analysis

- PIs form the core of GEC attendees. These are blue dots at the center of the chart. They repeatedly attend GECs.
- GPO attendees, although in small numbers, also show in the core of GEC attendees.
- Workers are distributed mostly in the core with some in the periphery.
- Experimenters show more on the lower right corner, meaning they became more involved in later GECs.
- Overall, people taking on GENI roles tend to repeatedly attend GECs, forming the core of GEC attendees.

# Observation from the Role-GEC Analysis (cont'd)

- PIs (>19%) and workers (24%) constitute a large proportion of GEC attendees.
- We also see an increasing number of GEC attendees in GPO, PI, Experimenter, and Other categories.
- GEC 9 shows a big increase in the number of experimenters.
- The majority of attendees do not take on any role.



# Institution-GEC Analysis

- The charts show the distribution of different types of institutions among GECs.
  - Government – Institution is directly overseen by a government
  - NSF – National Science Foundation
  - BBN – BBN / GPO
  - Private – Institution is privately owned or managed
  - University – Institution of higher learning
  - Unknown – Unidentifiable institution
- Dots are institutions. Squares are GECs.
- The color shows the type of institution, and the shape depicts whether the institution is domestic or international.

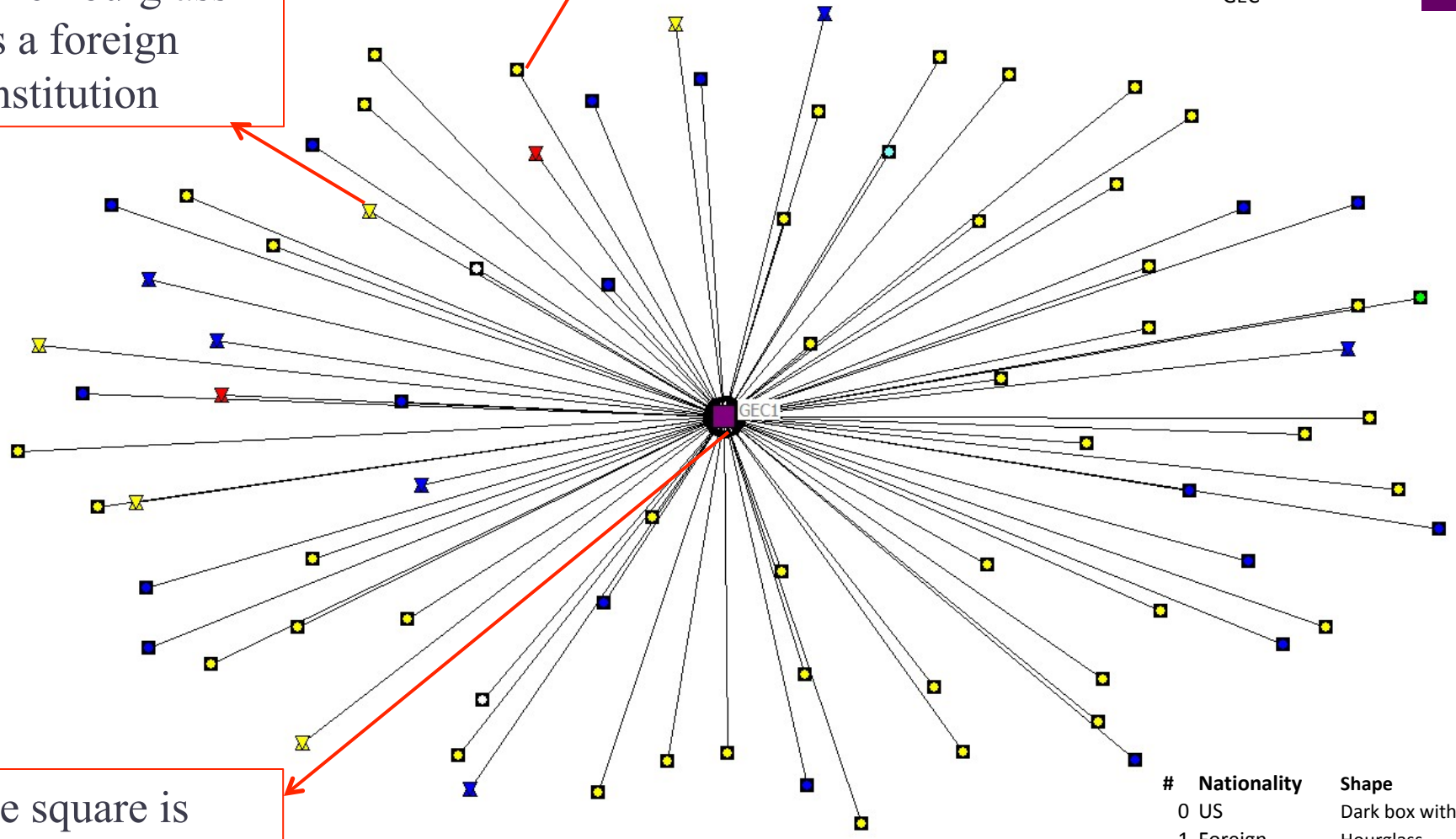
# GEC1

The hourglass is a foreign institution

The shaded box with a circle is a US institution

The square is GEC conference

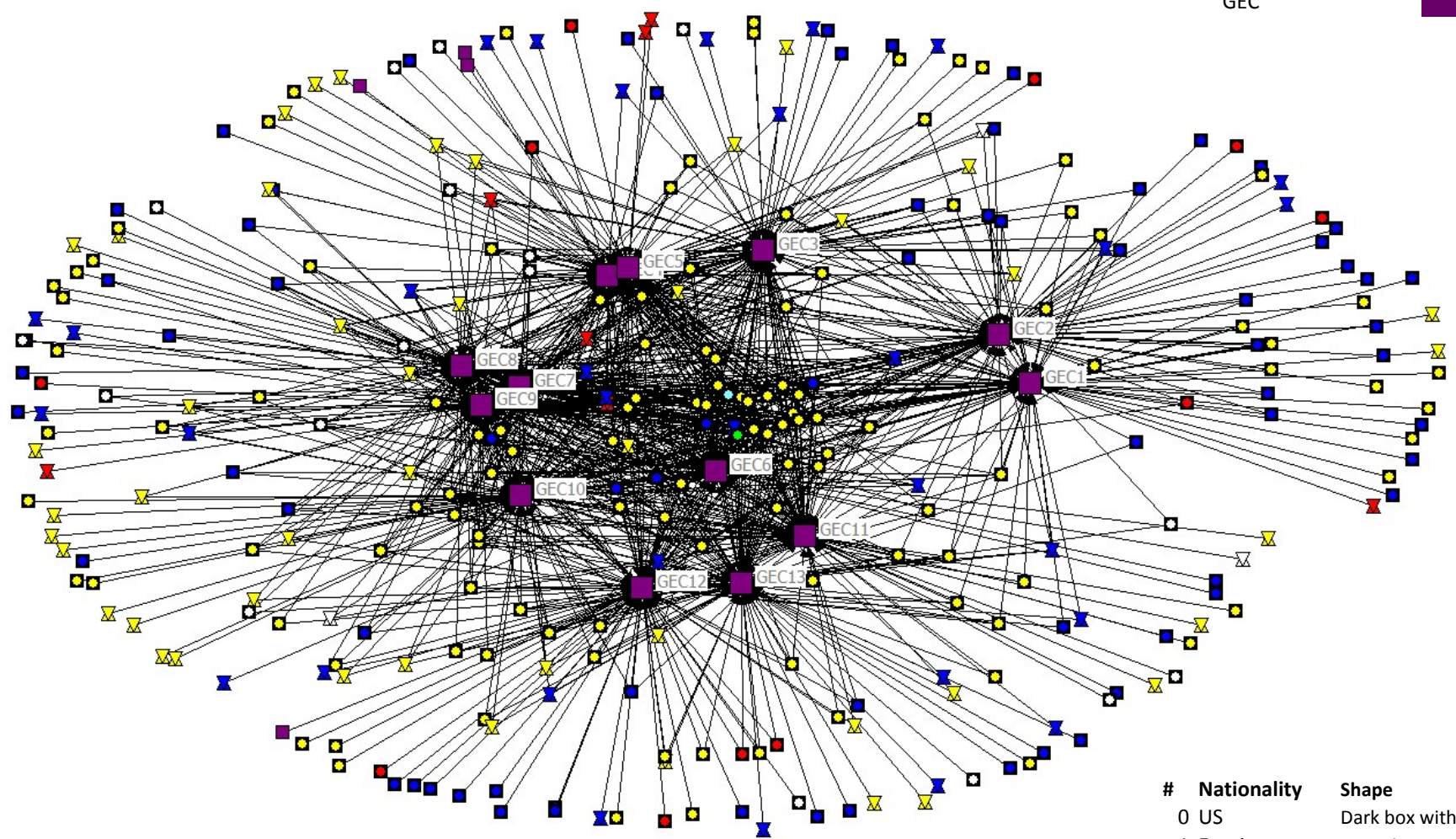
Institution Type	Color
Government	Red
NSF	Green
BBN	Cyan
Private	Blue
University	Yellow
Unknown	White
GEC	Purple



#	Nationality	Shape
0	US	Dark box with circle
1	Foreign	Hourglass
	GEC	Square

# GEC1-13

Institution Type	Color
Government	Red
NSF	Green
BBN	Cyan
Private	Blue
University	Yellow
Unknown	Black
GEC	Purple

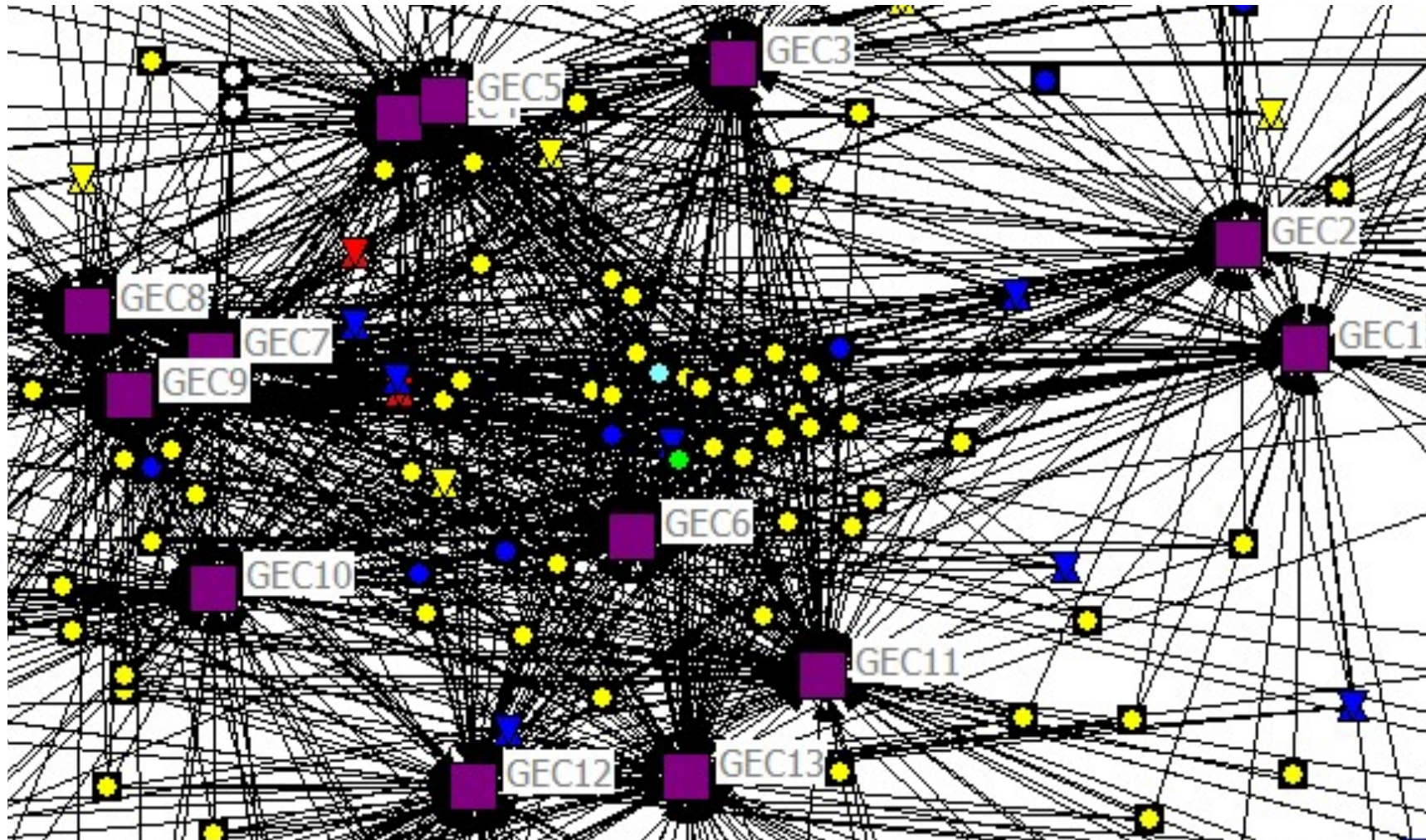


#	Nationality	Shape
0	US	Dark box with circle
1	Foreign	Hourglass
	GEC	Square



# GEC1-13 (Core)

#	Nationality	Shape	Institution Type	Color
0	US	Dark box with circle	Government	Red
1	Foreign	Hourglass	NSF	Green
	GEC	Square	BBN	Cyan
			Private	Blue
			University	Yellow
			Unknown	Black
			GEC	Purple



# Some numbers

GEC#	Government	NSF	Private	University	Unknown
1	2	1	25	48	2
2	4	1	32	51	2
3	5	1	27	59	4
4	4	1	22	56	2
5	5	1	16	56	5
6	3	1	19	58	5
7	4	1	17	60	5
8	3	1	22	66	5
9	4	1	25	67	3
10	2	1	18	80	4
11	1	1	14	72	4
12	4	1	27	72	3
13	3	1	27	62	2
1~13	44	13	291	807	46
%	3.66%	1.08%	24.23%	67.19%	3.83%

# Observations from the Institution-GEC Analysis

- Universities form the majority of the institutions in all GECs. They also form the core of the GEC institutions.
  - Universities constitute the biggest proportion (>67%) of all institutions attending GECs.
- Private institutions are the second most frequent type of institutions in all GECs. They mostly distribute in the periphery, indicating they attended a small number of GECs.
  - Private institutions constitute the second biggest proportion (>24%) of all institutions attending GECs.
- There are fewer government institutions among all GECs, and they also distribute in the periphery.
  - Government and NSF constitute the smallest proportion (about 5%) of all institutions attending GECs.

# Themes: Community

- A GENI community appears to have taken shape.
  - Early interviews suggested that the broader academic community had not coalesced around GENI and that there was no single community with a single vision.
  - Over time, we see an involved core group of participants in GENI that is growing in size.
    - About 150 participants, mostly PIs
    - Tend to assume a specific role
    - Tend to belong to a specific cluster
    - Are members of institutions that tend to send more than 1 person to GECs

# Themes: Growth

- The GENI community is organic and growing.
  - Relatively stable core (in terms of participants) that has evolved over time
  - Overall the number of attendees at GECs is generally increasing over time
  - Significant number of new participants at each GEC
  - Significant number of students at each GEC
  - More than 39% of attendees at one GEC attend the next GEC, with a maximum of 58% of GEC11 attendees attended GEC12



# Themes: Participation & Identity

- Those who “belong” participate more.
  - Most people unassociated with a particular cluster only attend 1 GEC
  - Majority of attendees don't take on a specific role
    - Those with specific role attend multiple GECs and often become part of the core
- Participants who identify with GENI through a GENI cluster tend to get more involved in GENI and attend more GEC conferences.

# Themes: Institutional Involvement

- Strong and sustained involvement in GENI by many institutions suggests a commitment to GENI.
  - About 89% attend more than one GEC
- Interesting partnerships and “mentorships” evolving between:
  - Academia and industry
  - Large and small universities
  - Research and teaching schools
- The type of institution attending GECs is largely universities (53%), with a smaller proportion of both private (34%) and government (6%) institutions attending.
- Roughly 28% of the attending institutions are international

# Themes: Social Networks

- GENI has provided tremendous networking opportunities for individuals who can establish and nurture collaborations.
- The high profile of GENI provides a form of “safety net” for junior faculty.
- GENI provides a mechanism for students and for researchers from smaller schools to become more integrated into the network of scholars.

# Thank you!

Our contact information:

**Professor Laurie Kirsch**

University of Pittsburgh

[lkirsch@katz.pitt.edu](mailto:lkirsch@katz.pitt.edu)

**Professor Sandra Slaughter**

Georgia Institute of Technology

[sandra.slaughter@mgt.gatech.edu](mailto:sandra.slaughter@mgt.gatech.edu)

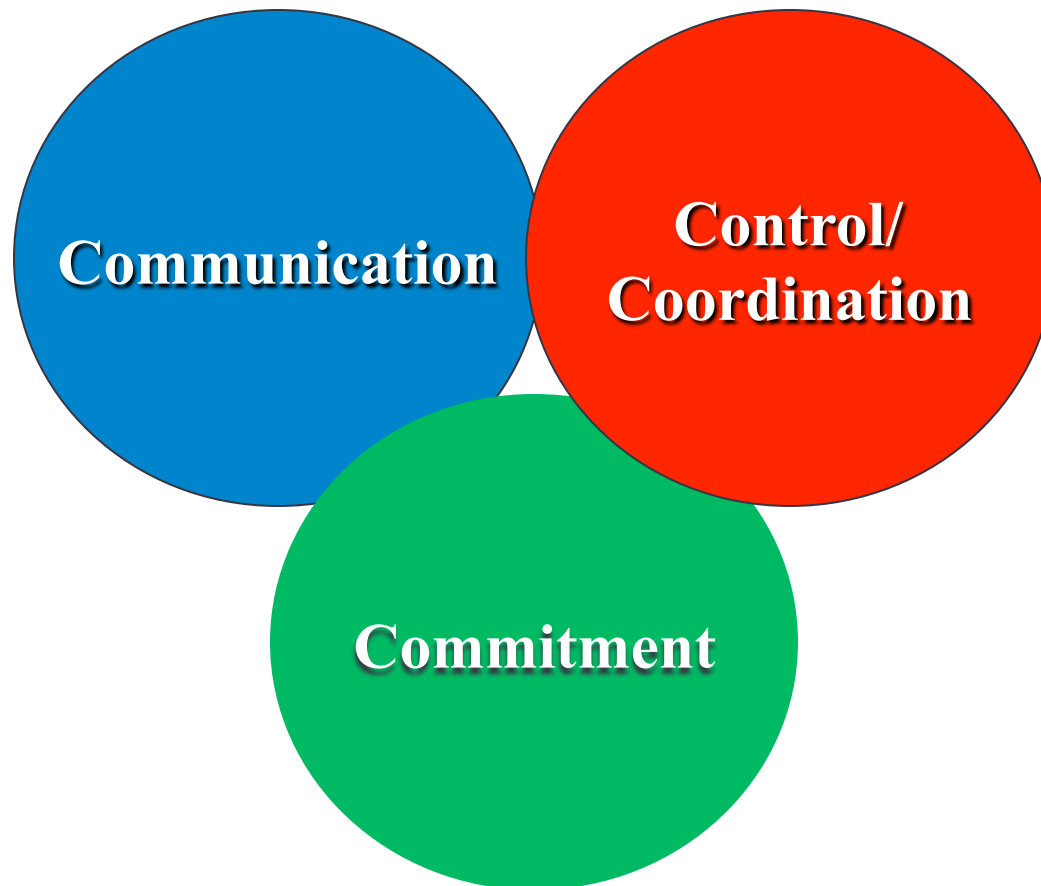
# ADDITIONAL SLIDES

---

# Why is the type of social structure important?

- Different structures have different advantages and disadvantages:
  - **Traditional hierarchy** – efficient but inflexible
  - **Fully connected team** – effective but time consuming
  - **Autonomous** – no information transfer
  - **Nearest neighbor** – convenient, but slow and information distorted as passed along
  - **Core-periphery** – brings in many perspectives, but depends on boundary spanners

# Critical Dimensions for Managing Cyber-Infrastructure Projects



# Communication Challenges

- Distributed communities of stakeholders
  - Must have effective mechanisms for distributed communication (not face-to-face) between team members
  - Common, shared infrastructure for project management and technical development is helpful
- Diverse groups of stakeholders
  - Requirements determination processes are critical
  - Need for communication and requirements discovery mechanisms to foster collaboration *across* stakeholders
  - Stakeholders who span across multiple groups are needed to facilitate communication



# Control / Coordination Challenges

- Constraints on schedule, budget, quality
  - Extensive project planning, monitoring and reporting is needed
  - Formal oversight is required
- Iterative development (innovation) process
  - Must manage “incubation” or “experimental” process
  - Need to facilitate technical integration which can be very complex
  - How to reconcile need for flexibility in development process with need for formality in project management

# Control / Coordination Challenges

- Funding
  - Funded by external agencies and must report to them
  - Complex project funding arrangements must be managed
- Different stakeholder communities
  - Control is indirect, complex and difficult to exert (i.e., different cultures, organizations)
  - Project Director plays an especially critical role in bridging

# Commitment Challenges

- Community-based instead of formal organizational roles
  - Stakeholders need to have clearly defined roles, standards and codes of conduct
  - Self-regulation mechanisms (reputation, trust, etc.) are required
  - Shared vision and consensus-based decision making processes are vital for commitment

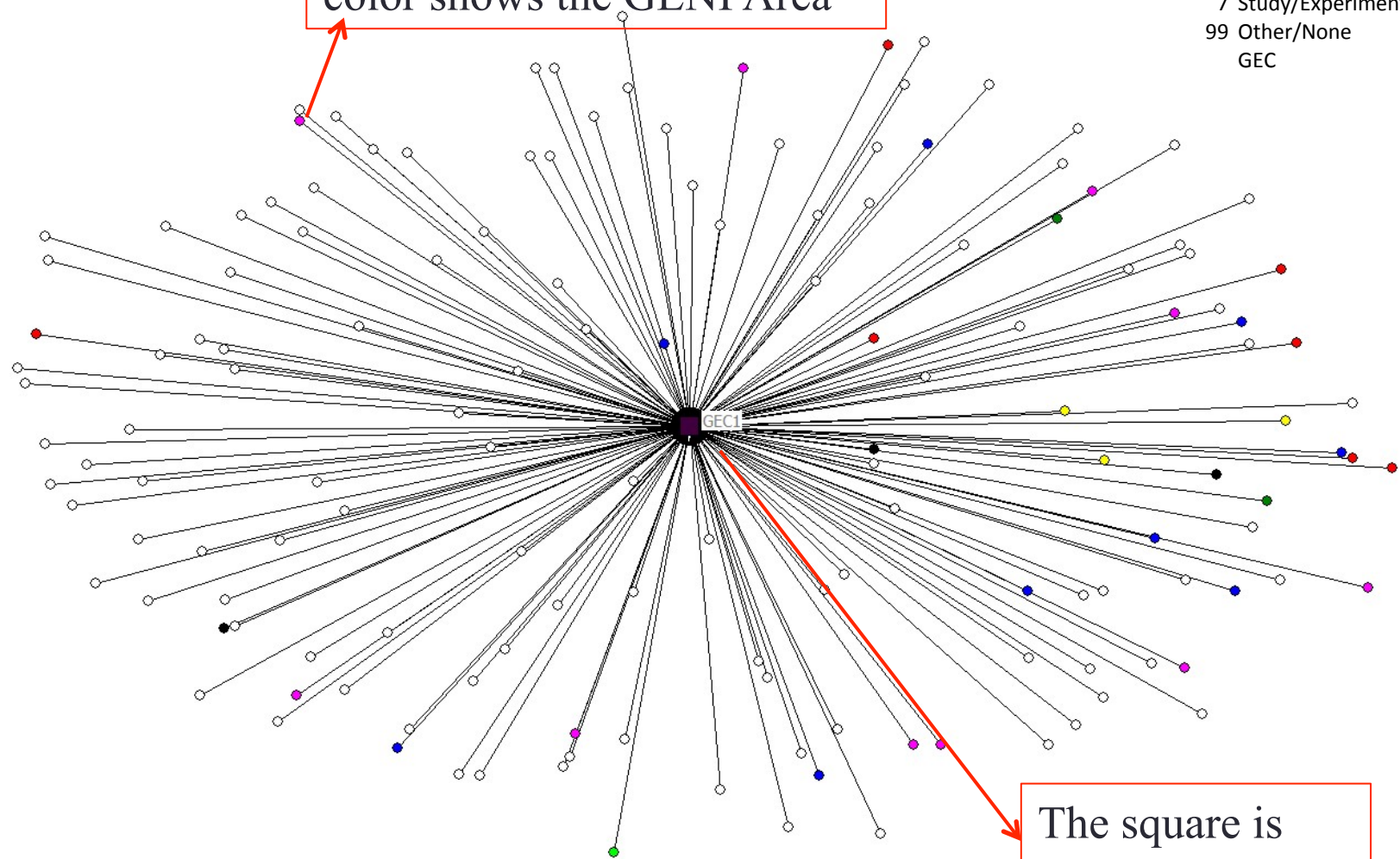
# Area-GEC Chart

- The charts show people's GEC attendance and their areas.
- They represent cumulative attendance since GEC1.
- Dots represent people. Squares represent GECs.
- People that locate at the center of the chart tend to be the ones that are active in attending GEC(s).
- The GENI areas are described in the following legend.

# GEC1

#	GENI Grouping	Color
1	Cluster A	Green
2	Cluster B	Pink
3	Cluster C	Red
4	Cluster D	Black
5	Cluster E	Light Green
6	Meso	Blue
7	Study/Experiment	Yellow
99	Other/None	Purple

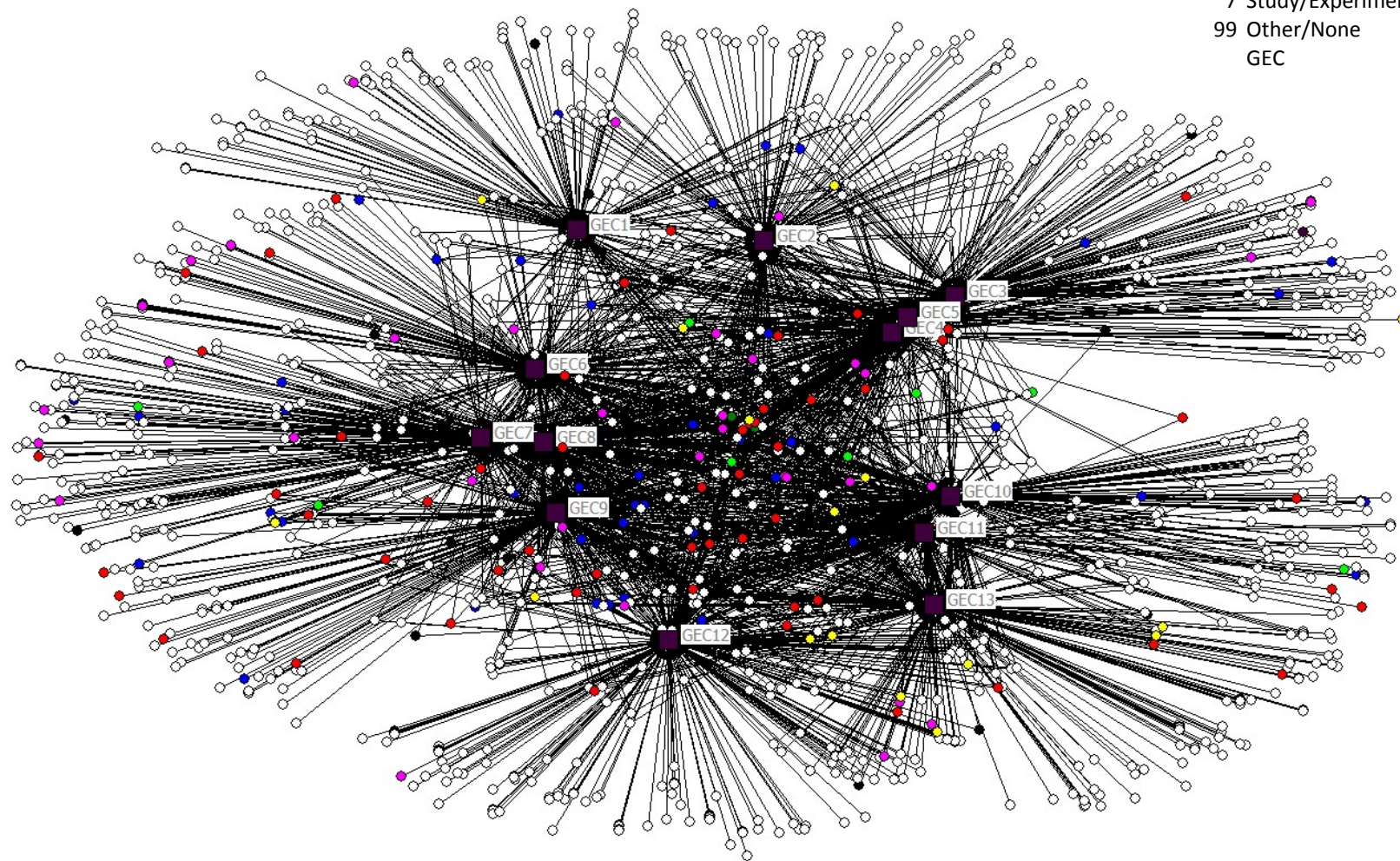
Each dot is an individual,  
color shows the GENI Area



The square is  
GEC conference

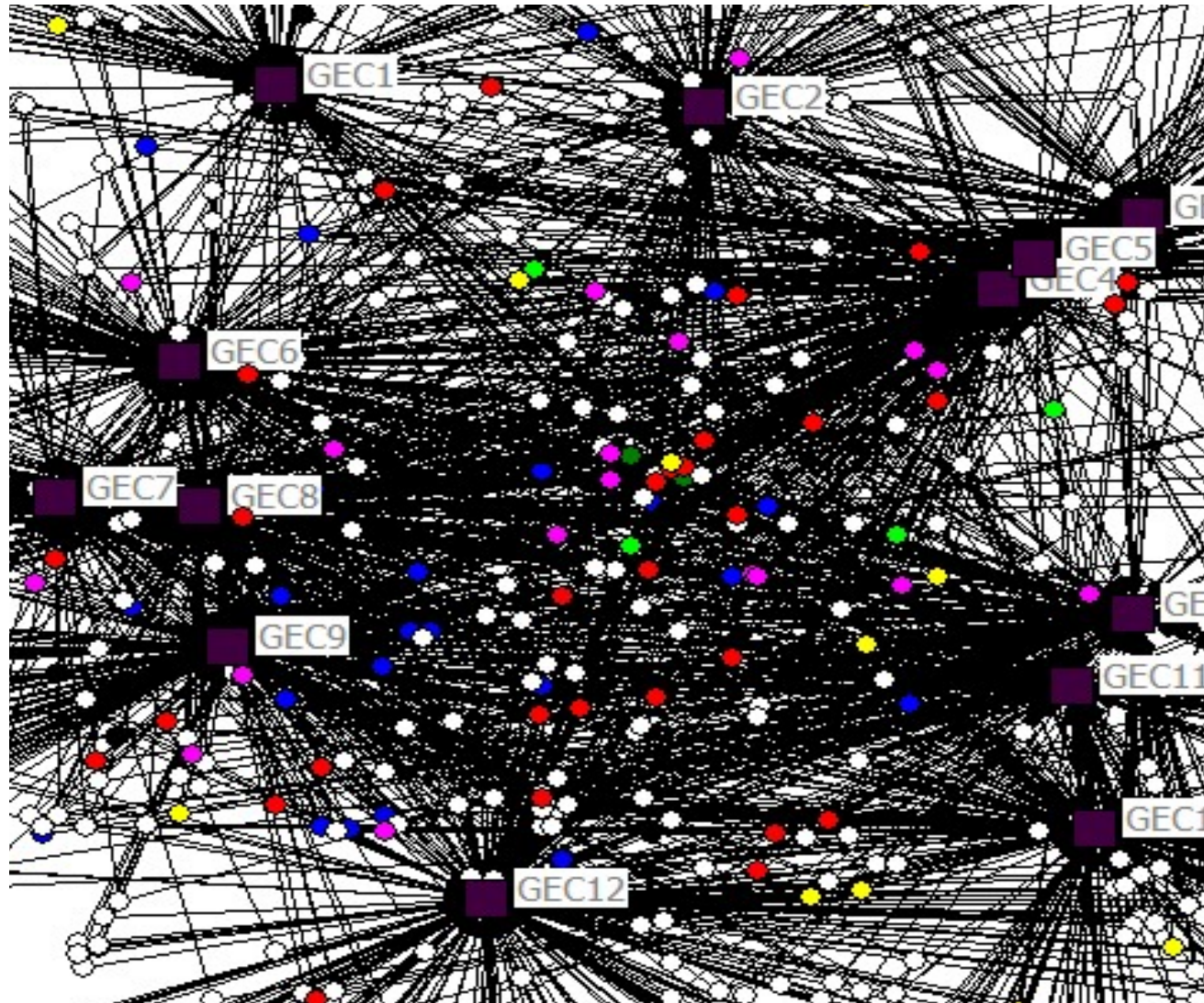
# GEC1~13

#	GENI Grouping	Color
1	Cluster A	Green
2	Cluster B	Pink
3	Cluster C	Red
4	Cluster D	Black
5	Cluster E	Light Green
6	Meso	Blue
7	Study/Experiment	Yellow
99	Other/None	Purple
	GEC	Purple





# GEC1~13 (core)



#	GENI Grouping	Color
1	Cluster A	Green
2	Cluster B	Pink
3	Cluster C	Red
4	Cluster D	Black
5	Cluster E	Light Green
6	Meso	Blue
7	Study/Experiment	Yellow
99	Other/None	Purple

# Observations from the charts

- People from different areas gradually form the core of the GEC attendees. These are people who locate at the center of the charts. They include people from all areas.
- Most of the core group are from the following areas:  
Clusters B, C, D & Meso
- Most of the people in the “None” area attend only one GEC conference. This supports our argument that when people identify with GENI through an GENI area, they tend to get more involved in GENI and attend more GEC conferences.
- Experimenters are mostly at the periphery in early GECs, indicating they only attend one or two conferences. But they started to get involved more and more in later GECs.



# Some numbers

GEC#	Cluster A	Cluster B	Cluster C	Cluster D	Cluster E	Study/ Experiment	Meso	None
1	2	10	7	3	1	3	9	113
2	2	8	7	7	3	4	8	111
3	2	12	11	7	4	4	13	140
4	2	11	18	9	3	6	10	109
5	2	9	12	6	2	5	13	120
6	2	17	25	12	2	5	17	142
7	2	11	20	12	3	4	20	142
8	2	15	19	10	6	5	26	163
9	2	15	27	12	3	9	19	177
10	2	10	21	9	2	9	14	193
11	2	12	19	11	6	11	19	121
12	2	14	22	14	2	7	17	238
13	2	7	21	12	4	8	10	167
1~13	26	151	229	124	41	80	195	1936
%	0.90%	5.40%	8.23%	4.46%	1.47%	2.88%	7.01%	69.59%

Throughout GEC1 to GEC13, 8.23% of the total attendees are from Cluster C.

# Observations from the numbers

- Among all the GENI areas studied, Meso group (>7%) and Cluster C (>8%) consistently sent the most people to the GEC conferences.
- We also see other clusters send more and more people to the GEC conferences.
- We see an increasing number of people in the group “study/experiment” attending the GEC conferences, although still in small numbers.
- Over 69% of attendees have not yet been identified to belong to a GENI area. This implies a great potential for getting more people involved in GENI development.

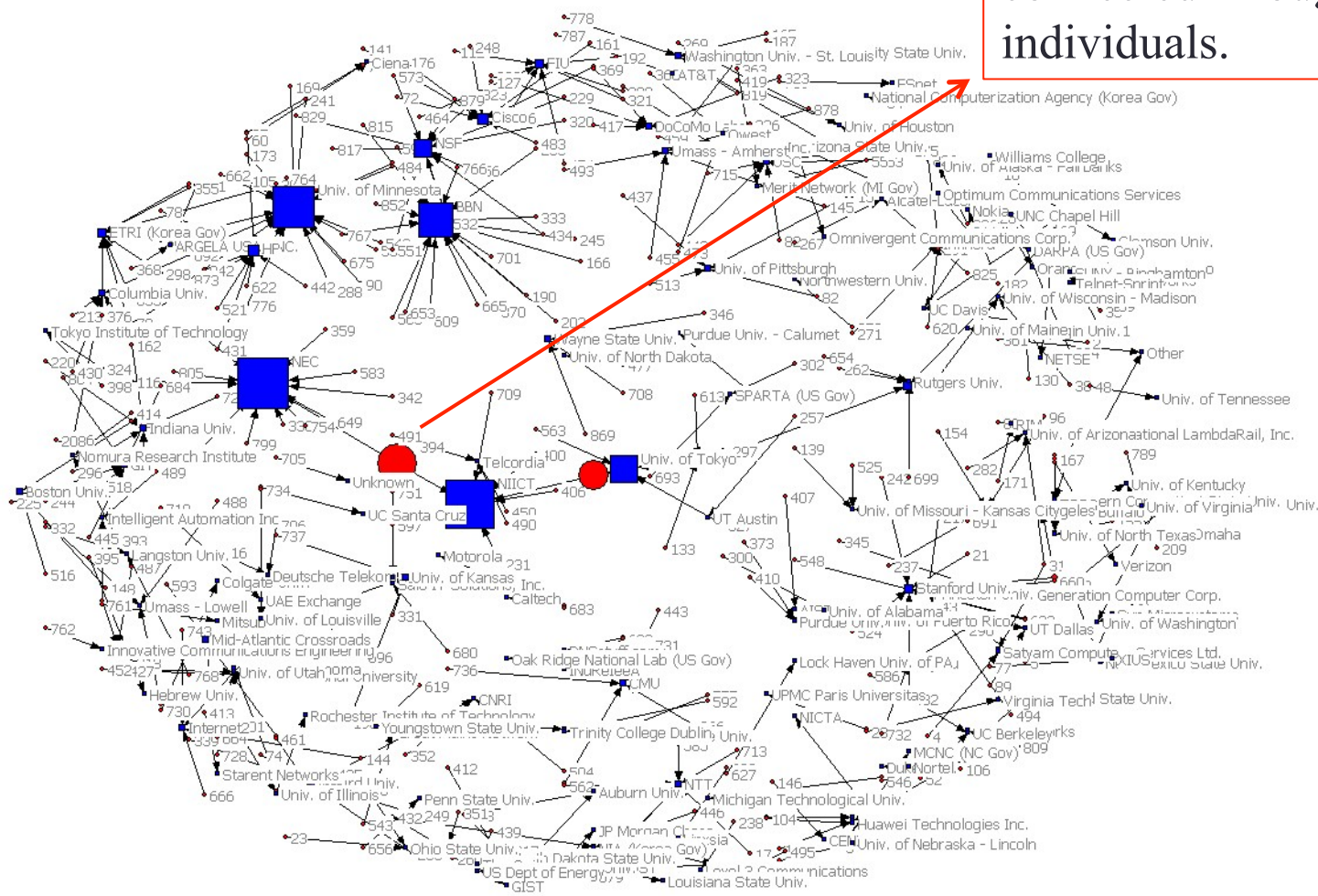
# Person-Institution Chart

- The charts show the relationships between GEC attendees and the institutions that they come from.
- They represent cumulative attendance since GEC1.
- Red dots are people. Blue squares are institutions.
- The size of the red dot represents the number of people are connected through these individuals. The size of the square represents the number of GEC attendees from the institution.



# GEC1~4

The bigger the red dots, the more people are connected through these individuals.











# Observations from the charts

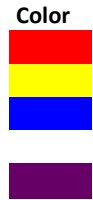
- Over the course of GEC conferences, the charts get denser, suggesting that more people and institutions attended GECs.
- Some institutions such as BBN, NSF, Umass-Lowell are shown in larger squares, indicating that more relationships were formed through them. For example, people who moved to or from these institutions help form relationships with other institutions.

# Institution-GEC by Size Chart

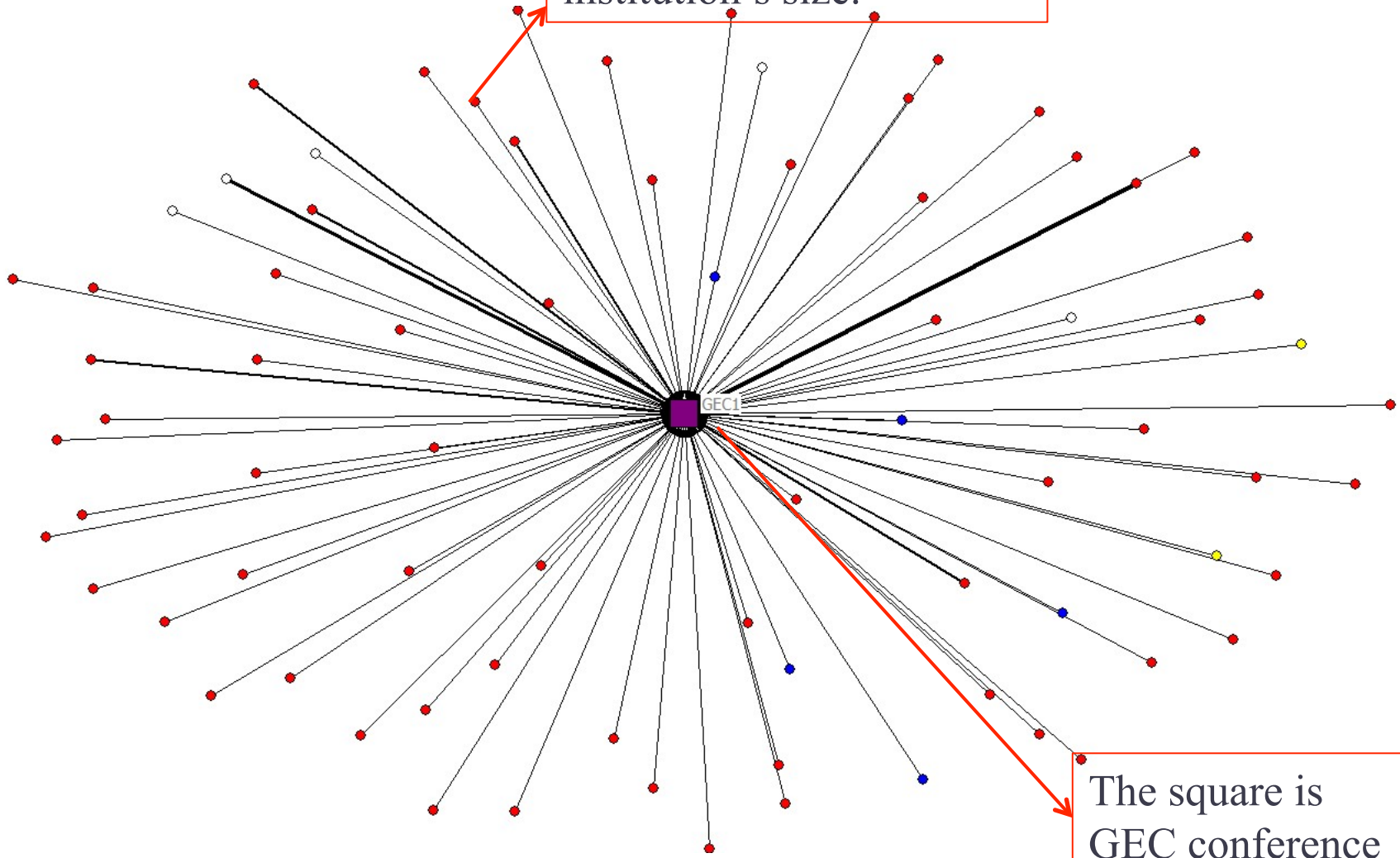
- The charts show the distribution of different sizes of institutions among GECs.
- Dots are institutions. Squares are GECs.
- The color shows the size of the institutions.

# GEC1

- # GENI Grouping
- 1 Large
  - 2 Medium
  - 3 Small
  - 4 Unknown
- GEC



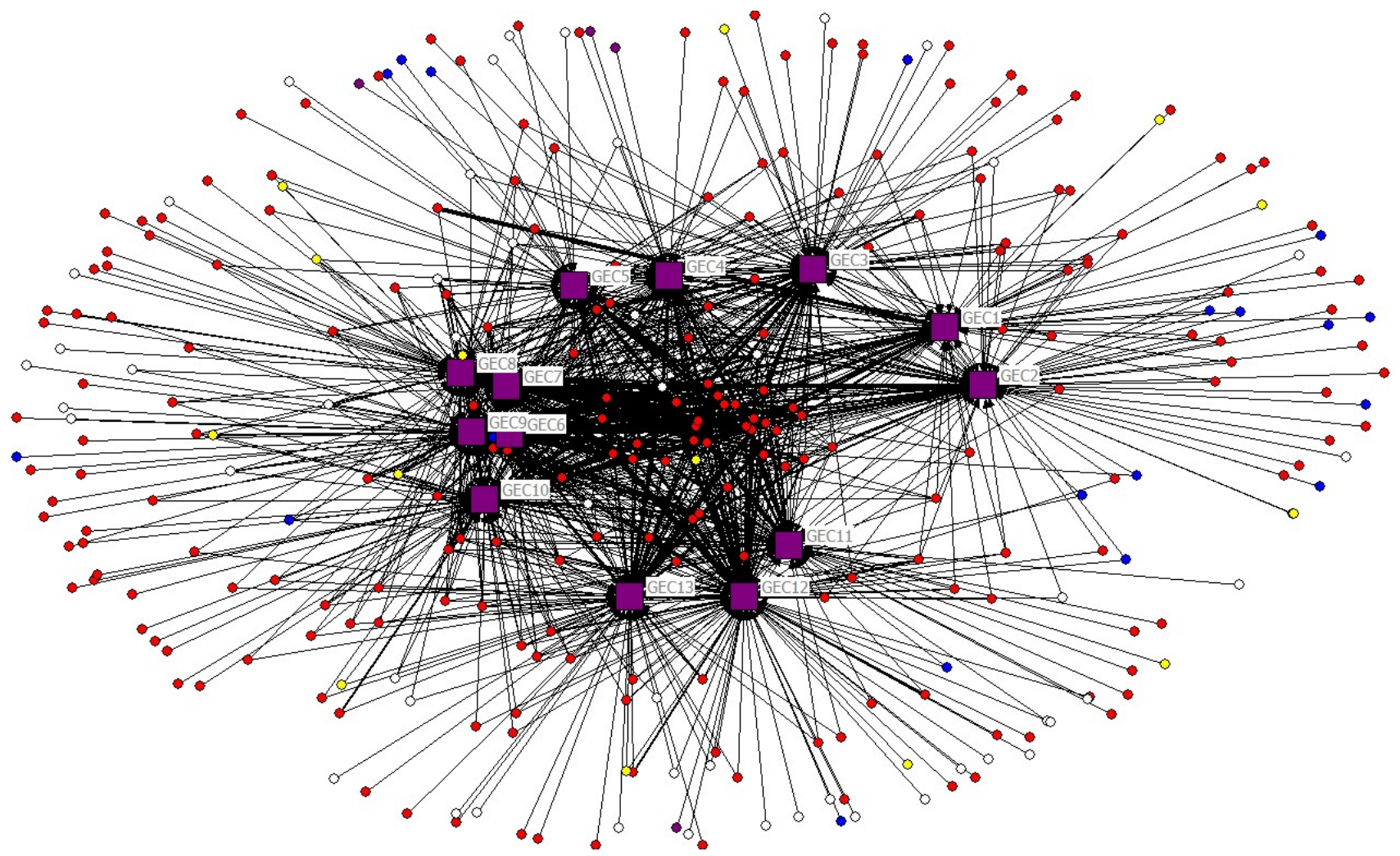
The dot is an institution;  
the color shows the  
institution's size.



The square is  
GEC conference

# GEC1~13

- # GENI Grouping
- 1 Large
  - 2 Medium
  - 3 Small
  - 4 Unknown
  - GEC



# Observations from charts

- Large institutions (represented by the red dots) form the majority of the institutions in all GECs. They also form the core of the GEC institutions.
- Most small size institutions (represented by the blue dots) distribute in the periphery, indicating they only attended a small number of GECs.
- There are even fewer medium size institutions (represented by the yellow dots) among all GECs. They also distribute in the periphery.

# Some numbers

GEC#	Large	Medium	Small	Unknown
1	66	2	25	5
2	72	2	32	9
3	80	3	27	10
4	67	4	22	11
5	66	4	16	10
6	71	2	19	10
7	72	4	17	10
8	79	6	22	10
9	82	5	25	11
10	88	5	18	10
11	77	2	14	10
12	81	4	27	18
13	80	3	27	10
1~13	981	46	291	134
%	67.56%	3.17%	20.04%	9.23%

# Observation from the numbers

- Average across all GECs, large institutions constitute the biggest proportion (>67%) of all institutions attending GECs.
- Small size institutions constitute the second biggest proportion (>20%) of all institutions attending GECs.
- Medium size institutions constitute the smallest proportion (about 3%) of all institutions attending GECs.
- Overall, there is a slight increase in the number for all sizes of institutions attending GECs from GEC1~13.