# Science of Science & Innovation Policy

Julia Lane

# Overview

1. Overview of Science of Science Policy Interagency group
   1. Roadmap
   2. Workshop (particularly data collection)
   3. Next steps
2. SciSIP program
   1. Summary of investigator initiated research
   2. SRS

# Science of Science Policy Interagency Group

- Formed under Committee of Science
- 17 agencies participating
- Cochairs Bill Valdez (DOE), Julia Lane (NSF)
- ITG engaged in a number of activities
  - Questionnaire
  - Literature review
  - Roadmap

# What we have learned

**Since the Science of Science Policy (SoSP) research program was launched in FY01, we have learned the following:**

- Qualitative methods (peer review, expert judgment, logic models, strategic planning, case studies, committee of visitors, etc.) remain the gold standard for policy makers who use decision support tools when making R&D investments and policy decisions.
- The best emerging quantitative decision support tools (risk analysis, dynamic modeling, network analysis, datamining, etc.) rely heavily upon expert judgment and advice from the scientific community to be successful.
- Considerable progress has been made on process metrics for science program management, but outcome/impact measures are still in their infancy.
- The traditional tools of R&D evaluation (bibliometrics, innovation indices, patent analysis, econometric modeling, etc.) are seriously flawed and promote seriously flawed analyses.

Source: Bill Valdez

# Theme 1:
## Understanding Science and Innovation

- What are the behavioral micro-foundations of innovation?
- What explains technology adoption and diffusion?
- How and why do communities of innovation form and evolve?

# Theme 1:
## Key Findings

- Well developed body of social science knowledge: not applied to the study of science and innovation
- Study of technology adoption and diffusion largely confined to academia.  Stronger links between academic and practitioner community needed
- Although each agency has its own community of practice, the collection and analysis of data about the scientists and the communities supported by those Federal agencies is heterogeneous and unsystematic. There is little analysis of the way in which the practice of science has become distributed across space, time, and disciplines as a result of computational advances.  As a result, there is little understanding of how scientific communities respond to changes in funding within and across disciplines and countries, or to changes in program focus.

# Theme 2:
## Investing in Science and Innovation

- What is the value of publicly funded knowledge?
- Is it possible to predict discovery?
- Is it possible to describe the impact of discovery?
- What are the determinants of investment effectiveness?

# Theme 2:
## Key Findings

- Although determining the value of publicly funded knowledge is the critical outcome measure for Federal scientific agencies, the analysis is largely agency specific
- Agencies are using very different approaches and tools designed to develop scenarios that anticipate the effects of discovery and innovation
- Agencies are using a wide variety of approaches to describe the impact of discovery.
- Approaches that are used by Federal agencies to determine program effectiveness span the spectrum from mature to those in the pilot stage, but there are many open research questions.

# Theme 3:

## Using the Science of Science Policy to Address National Priorities

- What impact does science have on innovation and competitiveness?
- How competitive is the US scientific workforce?
- What is the relative importance of policy instruments in science policy?

# Theme 3:
## Key Findings

- The ITG finds that there is a real opportunity to develop new tools and data sets that could be used to quantify the impact that the scientific enterprise has had on innovation and competitiveness.
- Many critical questions about the quality and global nature of the STEM workforce cannot be answered due to a lack of data. While the models and tools exist to study flows of workers within and across disciplines and nations, lack of data means that the science policy community cannot answer important questions about the scientific enterprise.
- There has been very little investment in the U.S. and in other countries in understanding the relative importance of policy instruments. While the models and tools exist to examine the effectiveness of different approaches, there are gaps in the analytical structure, the data infrastructure, and a way of conveying information to policymakers

# Recommendations

- Create an interagency research program to address the 10 scientific challenges
  - Invest in research data infrastructure
  - Invest in models, tools and metrics using ITG Evaluation Template
- Develop a National Innovation Framework
  - Explain benefits and effectiveness of S&T investments
  - Provide scenarios and options
- Create interagency entity to develop and sustain science policy analysis efforts
  - Synthesize and provide guide to current policy analysis practice
  - Nurture the nascent community of practice consisting of researchers and practitioners

# Next steps

- Roadmap going through concurrence process – now available
- Roadmap implementation workshop in early December
- Interagency working groups to be formed around key themes

# Data questionnaire for SOSP registration

- Building an empirical platform for the science of science policy requires good data. Please provide an assessment of the current empirical basis along the following dimensions
  - Data existence
  - Data quality
  - Data documentation
  - Data accessibility
- Assign a score of 1 to 5 for each criterion. In each criterion, a low score suggests doing less of an activity, and a high score suggests doing more of an activity.

# Data Input prior to SOSP workshop

| | Data covering the universe exist [1 is strongly disagree/ impact; 5 is strongly agree] | Data are high quality (e.g. have all key measures; measures reflect underlying concepts..) [1 is strongly disagree/ impact; 5 is strongly agree] | Data are well documented [1 is strongly disagree/ impact; 5 is strongly agree] | Data are available for use to the research and policy communities [1 is strongly disagree/ impact; 5 is strongly agree] |
|---|---|---|---|---|
| **"Input" Measures** | | | | |
| Generation of ideas (creativity) | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 |
| Transmission of ideas (Scientific communication) | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 |
| Source of ideas (STEM Workforce) | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 |
| Funds for ideas (Federal funding) | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 |
| Incentives for ideas (e.g. R&D tax credit) | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 |
| Discovery to innovation infrastructure (institutions) | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 | o o o o o 1 2 3 4 5 |
| **"Output Measures: Generation of Scientific Knowledge"** | | | | |

Tool Question 1: Measuring and tracking federal funding Narrative: This tool is intended to get your input on how to improve ways in which we measure federal funding. These are initial suggestions only: we also strongly encourage alternative suggestions. We have allocated 10 minutes to collect your input on this topic, to be followed by 10 minutes of group discussion. *Answer only those questions in areas which are within your expertise*

| Agree | Disagree | No Opinion | DATA ELEMENTS | Priority Rank | Comments |
|---|---|---|---|---|---|
| ☐ | ☐ | ☐ | **Option 1:** Encourage agencies to collect information on subawards and subprojects | select | |
| ☐ | ☐ | ☐ | **Option 2:** Encourage agencies to collect sufficient characteristics about those graduate students and postdoctoral researchers who receive federal funding so that their future scientific contributions can be tracked. | select | |
| ☐ | ☐ | ☐ | **Option 4:** Encourage agencies to use unique PI identifiers so that federal funding can be tracked across agencies | select | |

| | | | Data Infrastructure | | |
|---|---|---|---|---|---|
| Surveys | | | | | |
| ☐ | ☐ | ☐ | **Option 1:** Add questions to current federal surveys of researchers about sources and type of federal funding | select | |
| ☐ | ☐ | ☐ | **Option 2:**. Add questions to current federal surveys of businesses to collect additional information on federal funding. | select | |
| ☐ | ☐ | ☐ | **Option 3:**. Add questions to current federal surveys of federal agencies to collect additional information on federal funding. | select | |
| Administrative Data | | | | | |
| ☐ | ☐ | ☐ | **Option 4:** Encourage funding agencies to standardize their administrative records systems for initial awards as well as annual and final reports | select | |
| ☐ | ☐ | ☐ | **Option 4:** Encourage funding agencies to standardize their administrative records systems to capture funding by discipline | select | |
| ☐ | ☐ | ☐ | **Option 5**: Encourage the administrative sections of funding agencies to experiment with (and evaluate) different approaches to collecting PI information | select | |
| ☐ | ☐ | ☐ | **Option 6:** When developing data on organizations that conduct research, create the potential to link to funding agency administrative records | | |

| | | | Research | | |
|---|---|---|---|---|---|
| ☐ | ☐ | ☐ | **Option 7:** Establish a portal for all datasets (federal and non federal) that capture information about federal funding | select | |
| ☐ | ☐ | ☐ | **Option 8:** Establish a shared research environment so that datasets (federal and non federal) that capture information about federal funding can be integrated and analysed by researchers | | |
| ☐ | ☐ | ☐ | **Option 9:** Establish a shared research environment with award data for the research community to develop appropriate ontologies to track research | select | |

# Advancing the Science of Science & Innovation Policy

- Investigator Initiated Research
  - Three Solicitations – two sets of awards
  - Awards of interest to this group
- Statistical component
  - Many SRS surveys being redesigned
  - BRDIS

# Investigator Initiated Component Solicitations

# Awards from Solicitation I

- **Human capital development and the collaborative enterprise:**
- **Returns to international knowledge flows**
- **Creativity and innovation**
- **Knowledge production system**
- **Science policy implications**

# Awards from Solicitation II

- Describing the Role of Firms in Innovation
- Measuring and Tracking Innovation
- Measuring and Evaluating Scientific Progress
- Advancing Understanding of Collaboration and Creativity
- Knowledge sharing and creativity
- Implementing Science Policy

# Awards of interest to this group

- Linking Government R&D Investment, Science, Technology, Firms and Employment: Science & Technology Agents of Revolution (Star) Database (Lynne Zucker and Michael Darby, University of California, Los Angeles)
  - Data creation with links from government investment in R&D through the path of knowledge creation, its transmission and codification; then commercialization
    - NSF, NIH, DoD and DoE grants,
    - All journal articles and citations, high-impact articles, highly-cited authors, UMI ProQuest Digital Dissertations
    - US utility patents (complete/parsed/cleaned),
    - Venture capital, IPOs, web-based firm data, and links to major public firm databases via ticker symbols and/or CUSIP numbers.
    - Concordance linking STAR IDs to the IDs in the Census Bureau’s Integrated Longitudinal Business Database (ILBD) and Longitudinal Employer-Household Dynamics (LEHD) program, Census data, for use within the Census Research Data Centers.
  - Dissemination
    - a public graphics-based site primarily oriented toward policymakers and the media,
    - a public site providing access to researchers for downloads and database queries limited to the public constituent databases or aggregates derived from the licensed commercial databases, and
    - on-site access at the National Bureau of Economic Research providing researchers access to the complete STAR Database

Figure 1 – Major Features of the U.S. National Innovation System in the STAR Database: Policy, Innovation, Institutional Processes, and Economic Growth

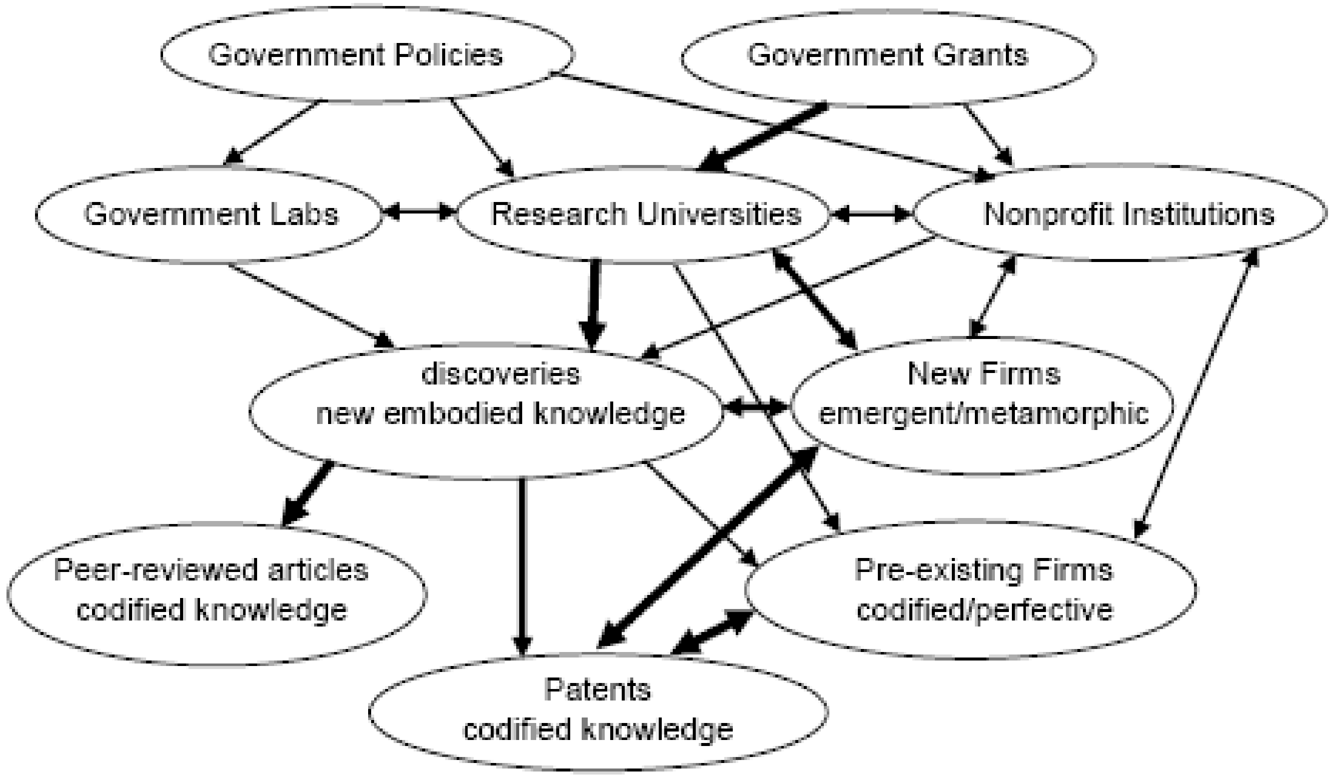
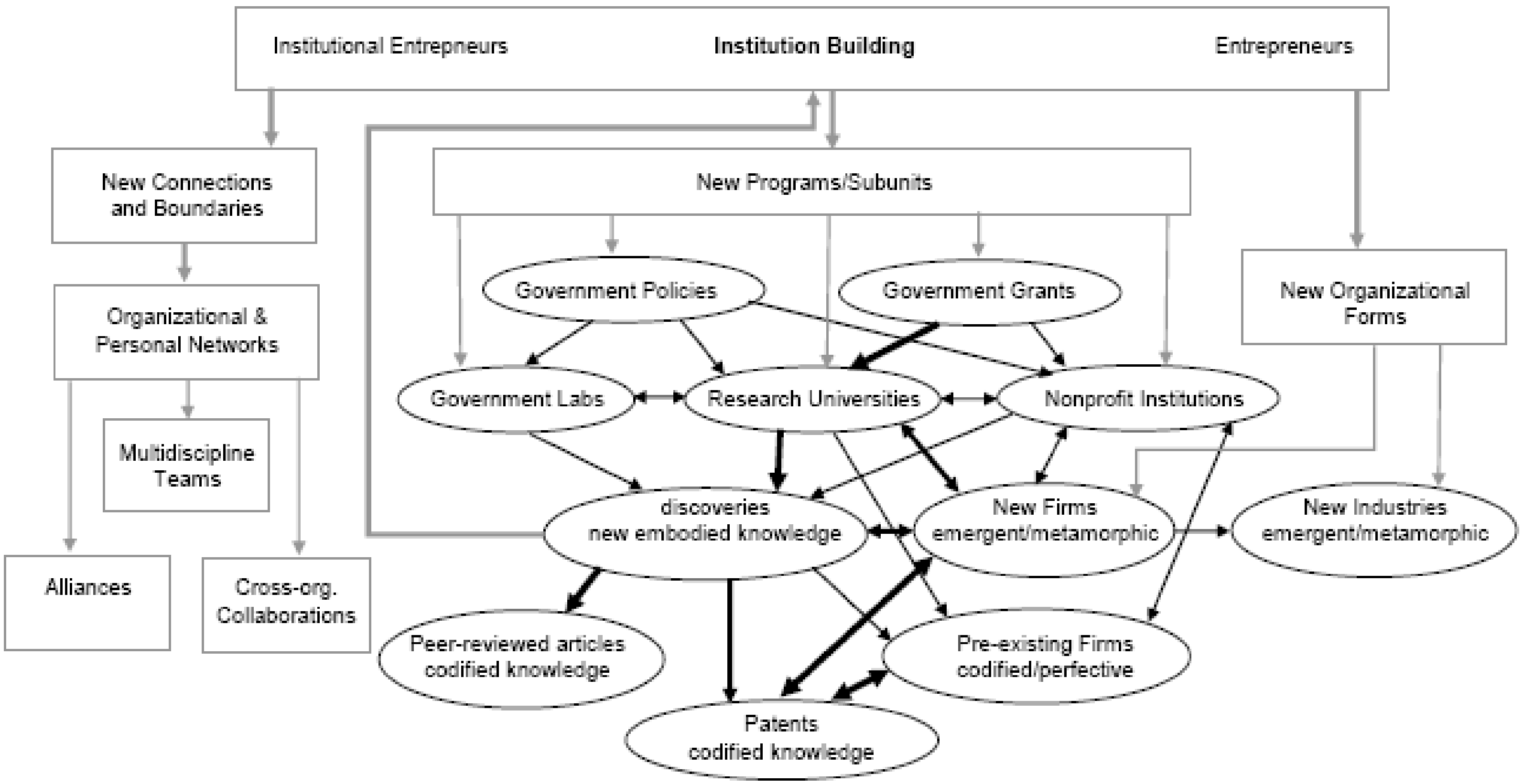


Figure 2 – Institutional Processes in Tandem with Knowledge Creation, Transmission and Use




Note: grey boxes and arrows denote institutional processes

# Awards of interest to this group

- A Social Network Database of Patent Co-authorship to Investigate Collaborative Innovation and its Economic Impact (Lee Fleming, Harvard University)
  - Develops a freely available social network database built from all U.S. patent co-authorships since 1963; Complements NBER patent database
  - Unit of analysis at the individual inventor and aggregate levels including organizational, regional, and technological
  - 1) refines inventor identification by encouraging inventors to check the identification algorithm,
  - 2) develops currently unavailable social network variables,
  - 3) makes the relational data easily available via the Harvard-MIT Dataverse infrastructure
  - 4) develops real time capability to visualize patent co-authorship networks.

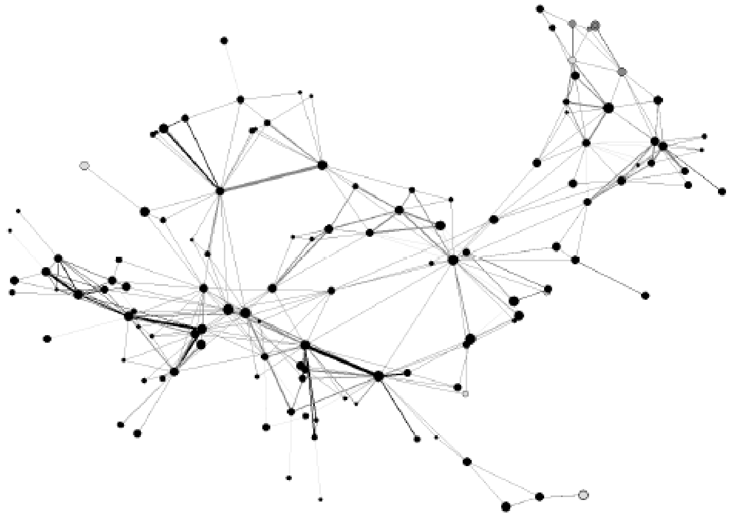


Figure 1: Bosch carburetor patents, circa 1980 (unpublished, developed with Dan Snow and Venkat Kuppuswamy). Note the difference with Figure 3, in that Bosch is much more collaborative. Nodes represent inventors and node size corresponds to the number of patents. Black nodes represent inventors who work in physical technologies, dark grey nodes represent electronic technologies, and light grey nodes represent inventors in both technologies. Tie width corresponds to the number of co-authored patents. Light grey ties represent later ties, black ties earlier ties, and dark gray ties intermediate

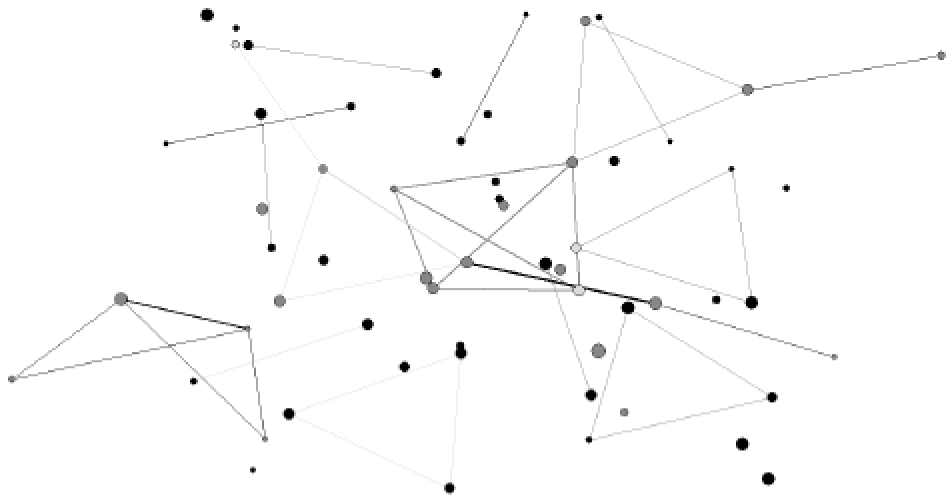


Figure 2: Ford carburetor patents, circa 1980 (unpublished, developed with Dan Snow and Venkat Kuppuswamy). Ford inventors are much more isolated and less collaborative than Bosch inventors illustrated in Figure 1.

# Awards of interest to this group

- Modeling Productive Climates for Virtual Research Collaborations (Sara Kiesler, Carnegie Mellon University and Jonathon Cummings, Duke University)
  - Unit of analysis is project-based research collaboration involving researchers from different institutions
  - Studies  the institutional environments of a sample of projects that were supported by the National Science Foundation.
  - Examines importance of a productive climate for distributed research collaboration,
  - Traces the linkages among productive climate and the institutional environments of these collaborations.
  - => better metrics for measuring and predicting performance and innovation in collaborations.

| Index | Items |
| --- | --- |
| Knowledge outcomes ('ideas') | Started new field or area of research; developed new model or approach in field; came up with new grant or spin-off project; submitted patent application; presented at conference or workshop; published article(s), book(s), or proceeding(s); recognized with award(s) for contribution to field(s). Alpha = .63 (7 items) |
| Tools outcomes ('tools') | Developed new methodology; created new software; created new hardware; generated new dataset; generated new materials; created data repository; created website to share data; created collaboratory; created national survey; developed new kind of instrument; created online experiment site. Alpha = .65 (11 items) |
| Training outcomes ('people') | Grad student finished thesis or dissertation; grad student/post-doc got academic job; grad student/post-doc got industry job; undergrad/grad student(s) received training; undergrad(s) went to grad school. Alpha = .70 (5 items) |
| Outreach outcomes ('people') | Formed partnership with industry; formed community relationship through research; formed collaboration with researchers; established collaboration with high school or elementary school students; established collaboration with museum or community institution; established collaboration with healthcare institution. Alpha = .45 (6 items) |

Table 1. Project outcomes studied in Cummings & Kiesler, 2007.

# Awards of interest to this group

- Dynamics of Creativity and Innovation in Cyber-enabled Scientific Commons (Levent Yilmaz, Auburn University)
  - Agent simulation models
  - (1) considers the discourse of scientific activity, including the contribution of new knowledge in virtual scientific commons, growth of the domain knowledge, and the clustering of research into specialties,
  - (2) views science as an autonomous and self-regulating socio-cognitive system through the introduction of motivation and competitive nature of knowledge production, and
  - (3) explores the impact of alternative community cultures (e.g., exploration-oriented, service-oriented, and utility-oriented), peer evaluation styles (e.g., centralized, decentralized) on the sustainability and innovation potential of SCs.
  - Creates an integrated and customizable agent simulation framework, called SciSIM, for science policy mechanism design and decision analysis for virtual scientific communities to improve sustainable innovation.

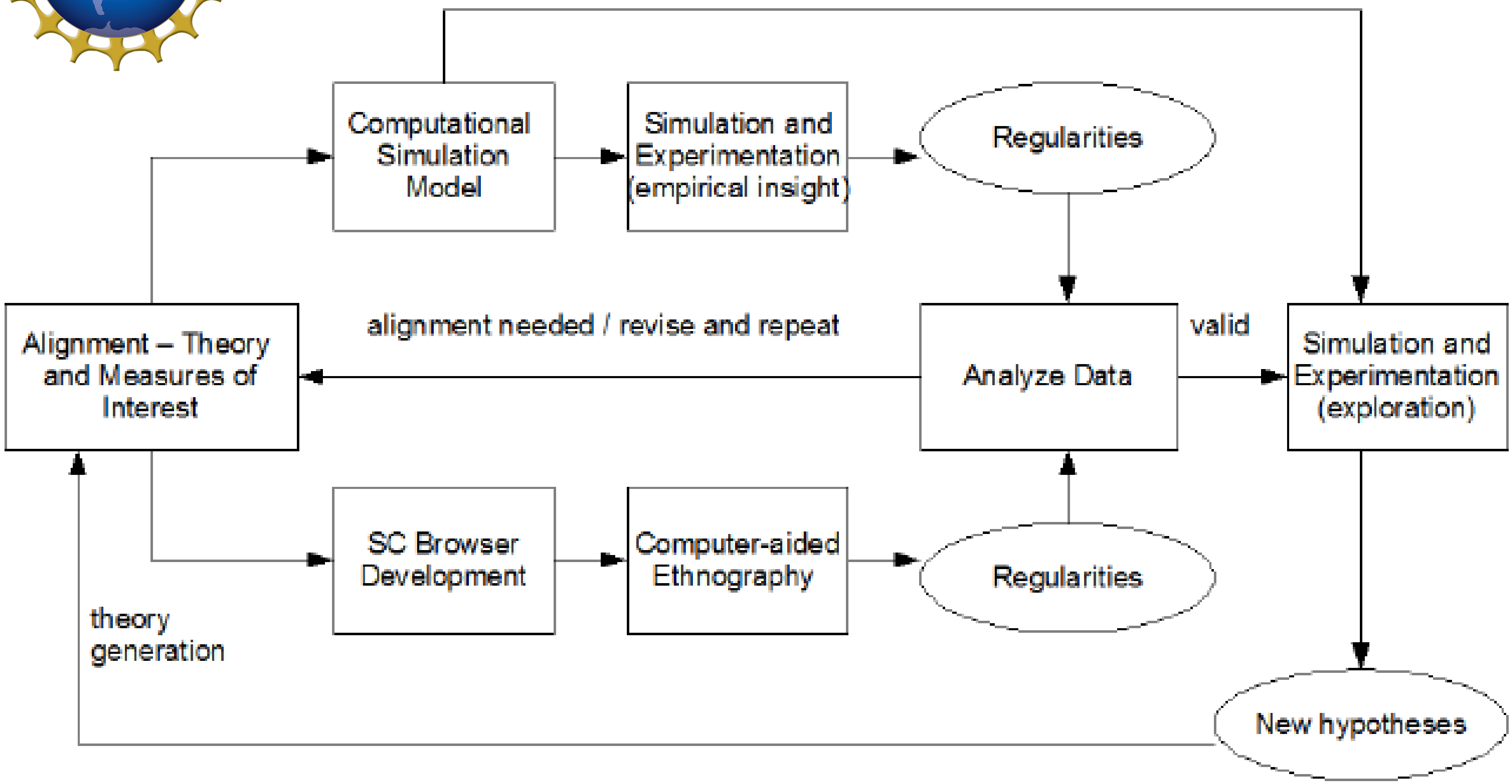



Figure 2: Research Strategy

# Awards of interest to this group

- Integrating Social and Cognitive Elements of Discover and Innovation (Chris Schunn, University of Pittsburgh)
  - Examines video data collected from a recent highly successful case of science and engineering, the Mars Exploration Rover. Traces the path from
    - the structure of different subgroups (such as having formal roles and diversity of knowledge in the subgroups)
    - to the occurrence of different social processes (such as task conflict, breadth of participation, communication norms, and shared mental models)
    - to the occurrence of different cognitive processes (such as analogy, information search, and evaluation)
    - and finally to outcomes (such as new methods for rover control and new hypotheses regarding the nature of Mars).

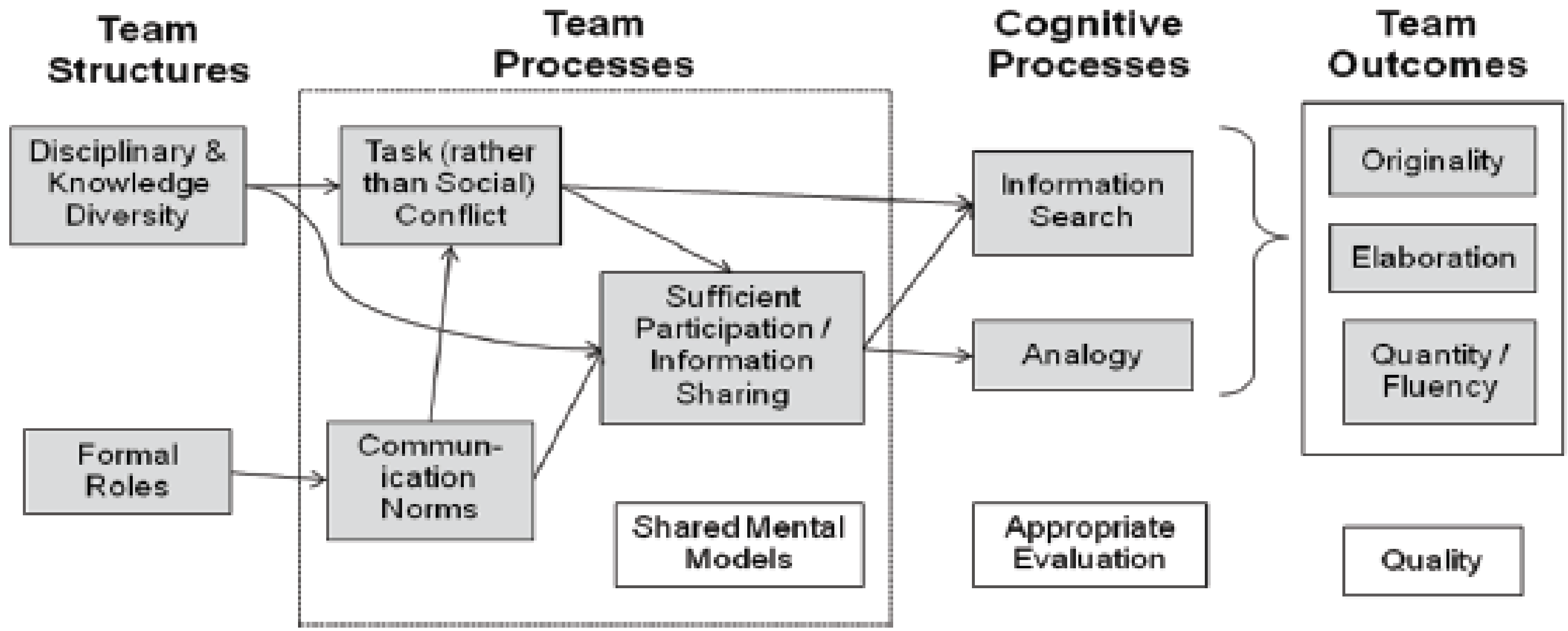


Figure 1: Hypothesized Social-Cognitive Pathways of Team Divergent Thinking

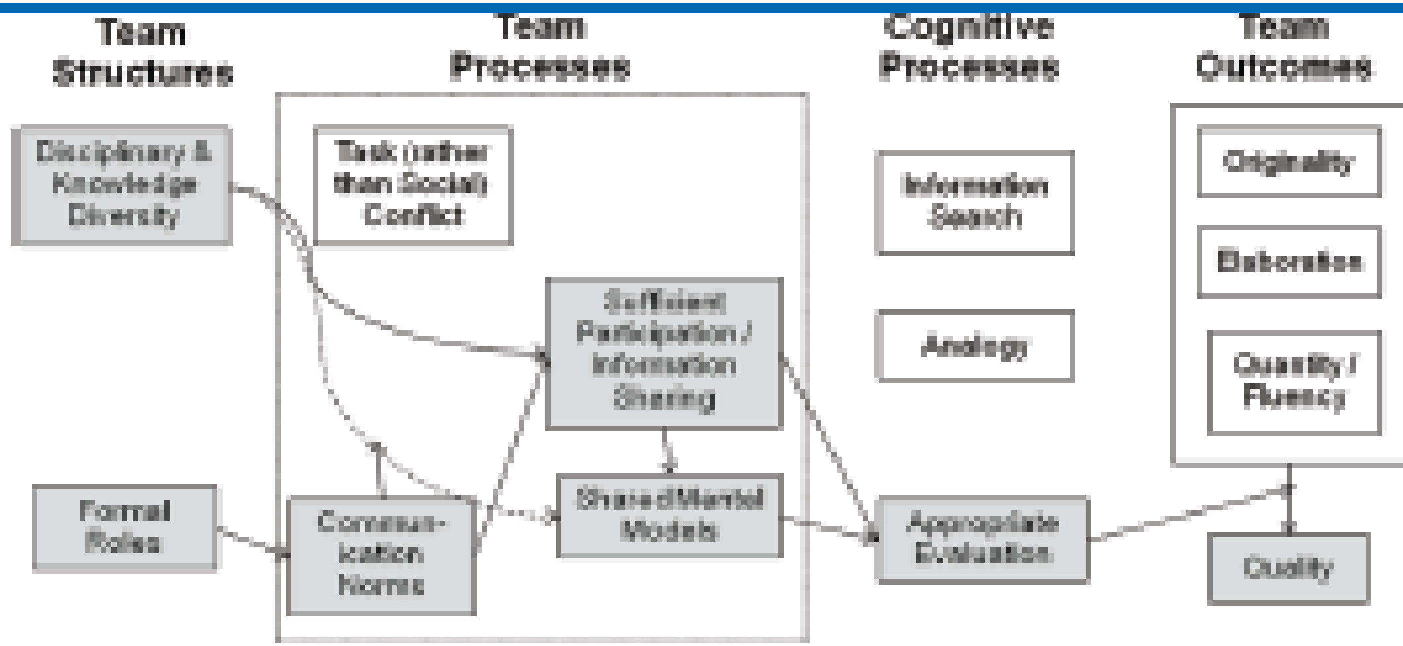


Figure 2: Hypothesized Social-Cognitive Pathways of Team Convergent Thinking

# Solicitation III

- Demonstration projects on Organizations and Innovation
- Visualization (drawing particularly on visual analytics)
- International Collaborations

SMART Board
Europe
Eurasia
N America
Middle East
Asia
Africa
S Americas
Oceania
government, security, minister
china, government, country
president, security, cooperation
government, people, minister
people, government, president
government, united, north
president, people, party
north, korea, government
security, government, people
Live News
IN-SPIRE

# SciSIP Milestones

- **Longer term:**
  - An evidence-based understanding of the impacts of the S&E enterprise
  - A capacity to better nourish and harness the capabilities of the national STEM workforce
  - The development of a Community of Practice

Thank you!

Comments and questions invited.

For more information please contact:

Julia Lane

jlane@nsf.gov