Measuring the Results of Science Investments

Julia Lane

Overview

- Why good measurement matters
- Why measurement is hard to do
- Designing a better system: the nascent approach in the US
  - The vision
  - What has been done
  - The pitfalls
- What might be done internationally

Why good measurement matters

Good measurement matters because it's the right thing to do

Good measurement matters because

Stakeholders demand it

Good measurement matters because

The alternative is worse

Why good measurement is hard to do

The conceptual framework is complex

- Production function framework great for aggregate impacts
  - source of result that more than 3/4 of post-1995 increase in productivity growth can be traced to science investments
- At micro level not so clear
  - Discovery – innovation highly nonlinear
  - Unit of analysis
  - Input measures
  - Dependent on organizational systems
- Outcome measures
  - Scientific; Economic; Social
- Fundamental challenge: Establishing counterfactuals
  - Selection bias
  - Random assignment not an option

And the data don't exist

The ITG undertook a literature review to determine the state of the science to date. A questionnaire was also circulated to Federal agencies to ascertain what methods are currently in use for programmatic investment decision making, as well as to ask what tools and resources are needed by Federal agencies that are currently unavailable. The ITG found that:

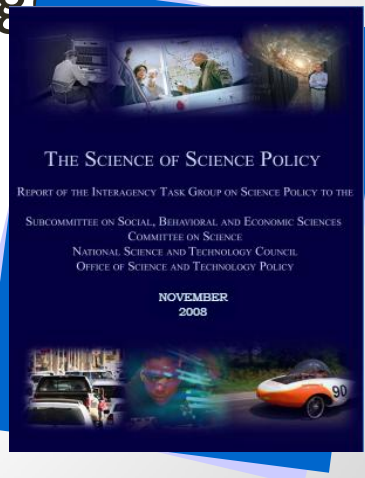
- There is a well developed body of social science knowledge that could be readily applied to the study of science and innovation.
- Although many Federal agencies have their own communities of practice, the collection and analysis of data about the science and scientific communities they support is heterogeneous and unsystematic.
- Agencies are using very different models, data and tools to understand their investments in science and technology.
- The data infrastructure is inadequate for decision-making.

THE SCIENCE OF SCIENCE POLICY: A FEDERAL RESEARCH R

The data don't exist for good reason

- No systematic documentation of inputs (who is supported by science funding)
- No systematic links between inputs and outputs
- Heavy reliance on manual reporting
- No systematic ability to capture outputs and outcomes beyond the period of an award
- Balkanized agency systems => impossible to g overview of science investments

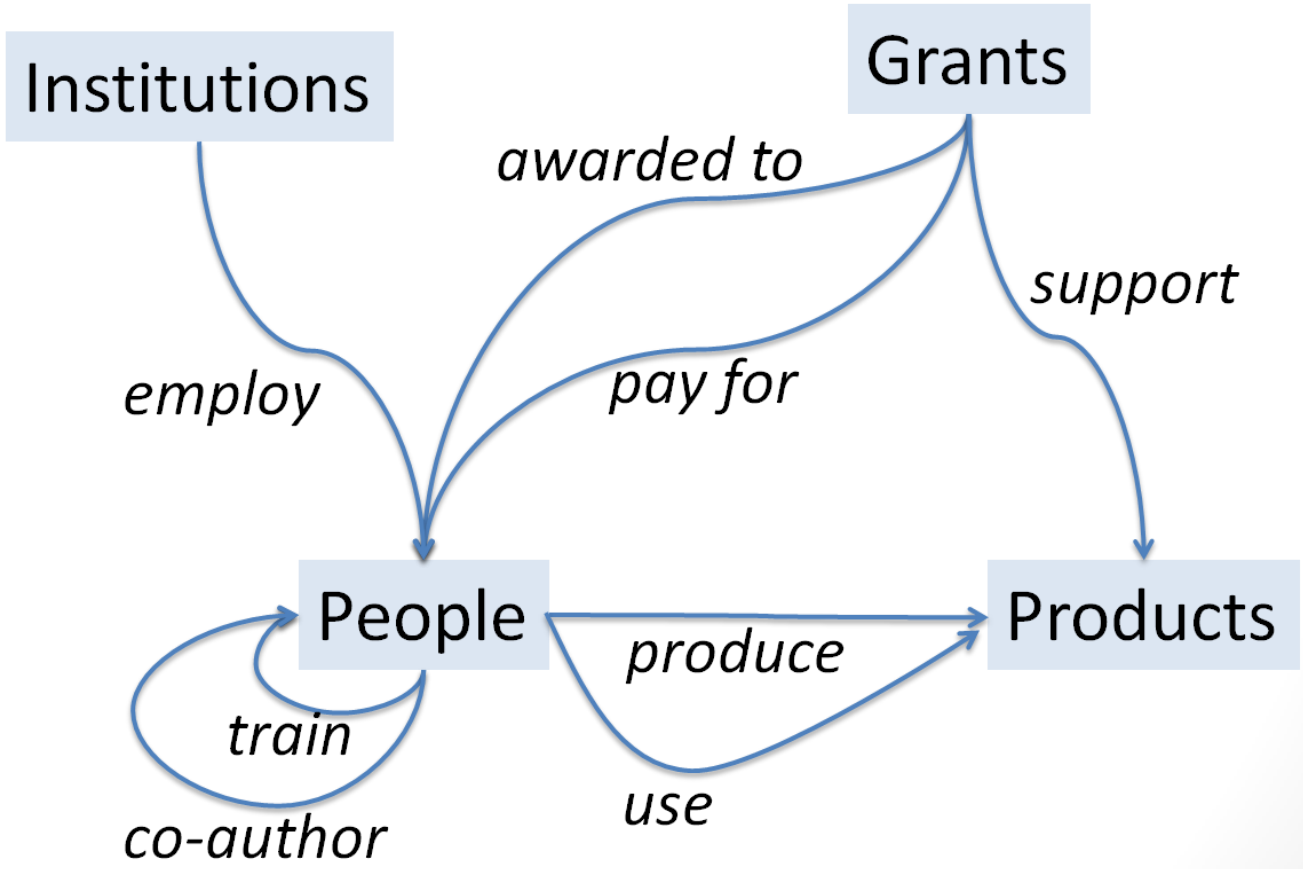
=> The data infrastructure is inadequate for decision-making



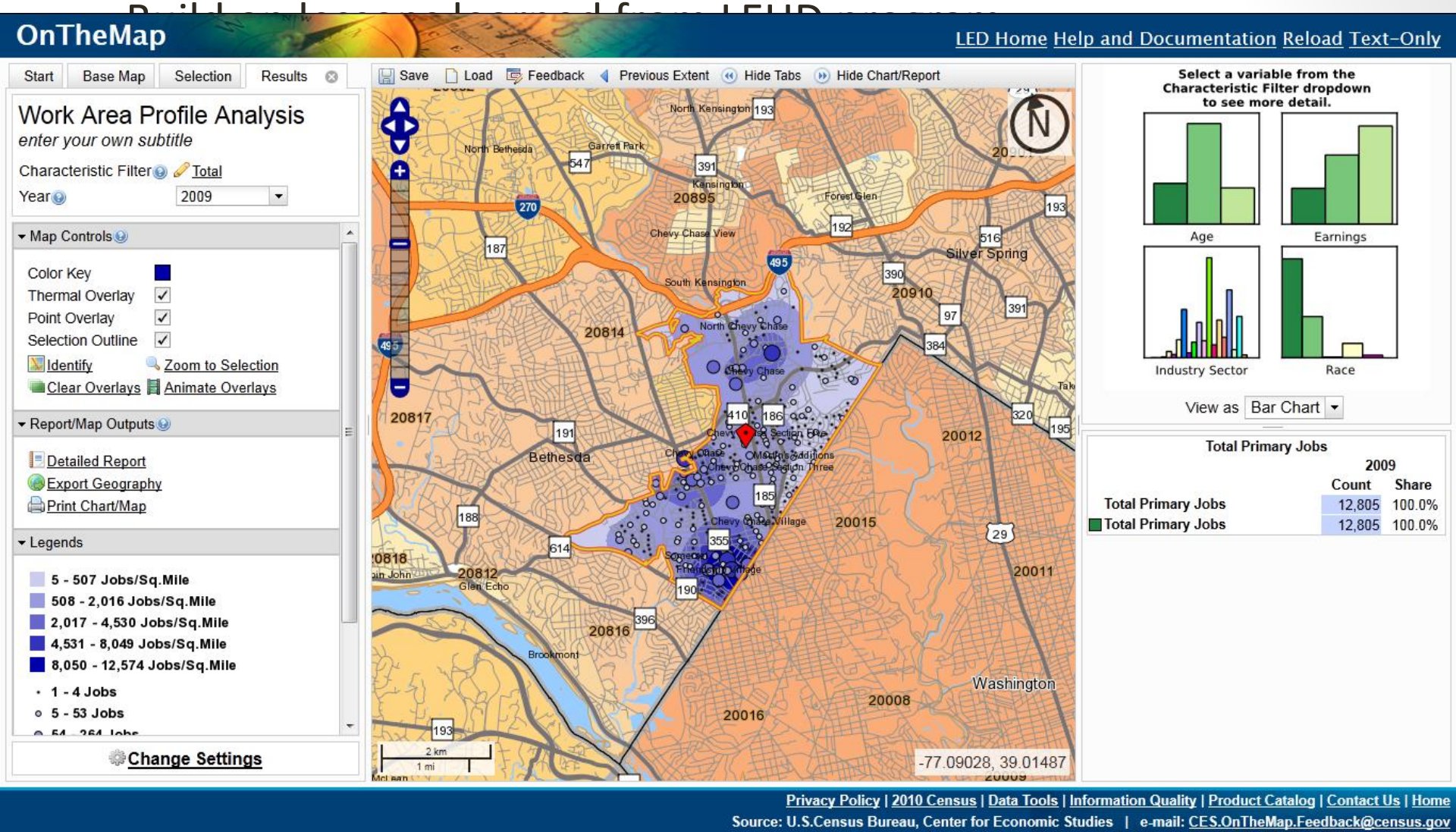
Designing a better system

What is needed

Automatically capture data about the conduct of science – inputs, outputs and the connections between the two



Designing a better system: based on collaboration

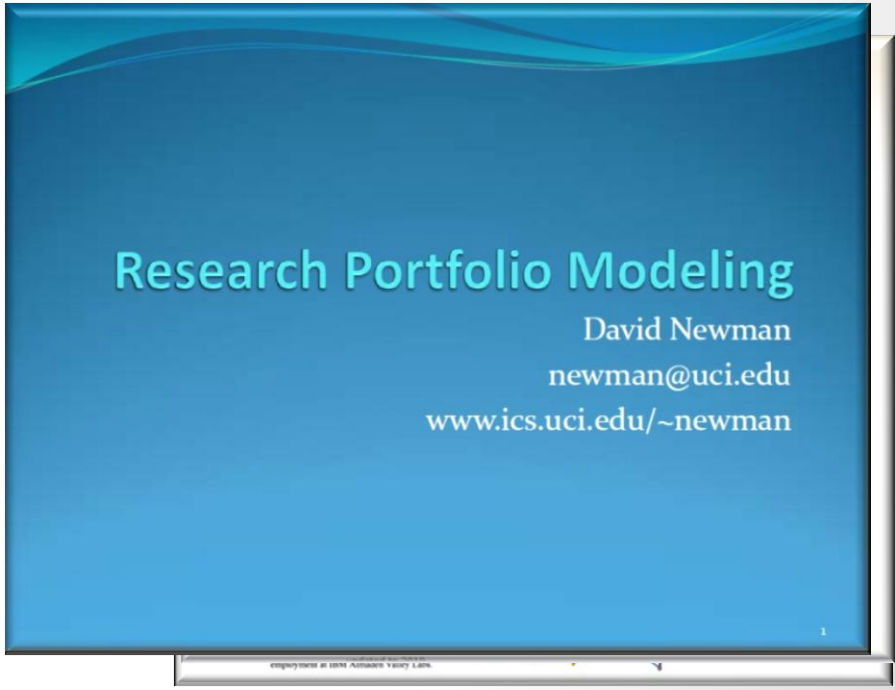


Designing a better system: based on science

Patent Network Dataverse Lee Fleming Harvard
http://dvn.iq.harvard.edu/dvn/dv/patent

COMETS (Connecting Outcome Measures in Entrepreneurship Technology and Science): Lynne Zucker and Michael Darby UCLA
http://www.kauffman.org/comets

Topic Modeling: David Newman UCI



Designing a better system: based on eScience

Use new cybertools e.g. use topic modelling and natural language processing to process massive amounts of text

Use new ways of structuring data e.g. graph databases

Use new approaches to link individuals with output e.g. webscraping

The Nascent System in the U.S.

- The vision
- What is being done
- The pitfalls

The vision

- **Level I:** Who is employed by science funding?
  - Current jobs
  - Future workforce
- **Level II:** What are the results of science?
  - scientific knowledge (such as publications and citations..)
  - economic growth (through patents, firm start ups and other measures…)
  - workforce outcomes (through student mobility and employment..)
  - social outcomes (such as health and environment…)

What has been done: STAR METRICS

Voluntary partnership between science agencies and research institutions to document the outcomes of science investments to the public

- OSTP initiative partnering with NIH, NSF, DOE and EPA; USDA has agreed to join.
- 85 research institutions participating - ~ 50 % of NSF and NIH portfolio

Star Metrics
Open API

Other Sources
TBD

Patents
Pubs

All Federal
Funding
Sources
Grant/Filter

Star Metrics

Scien
CV

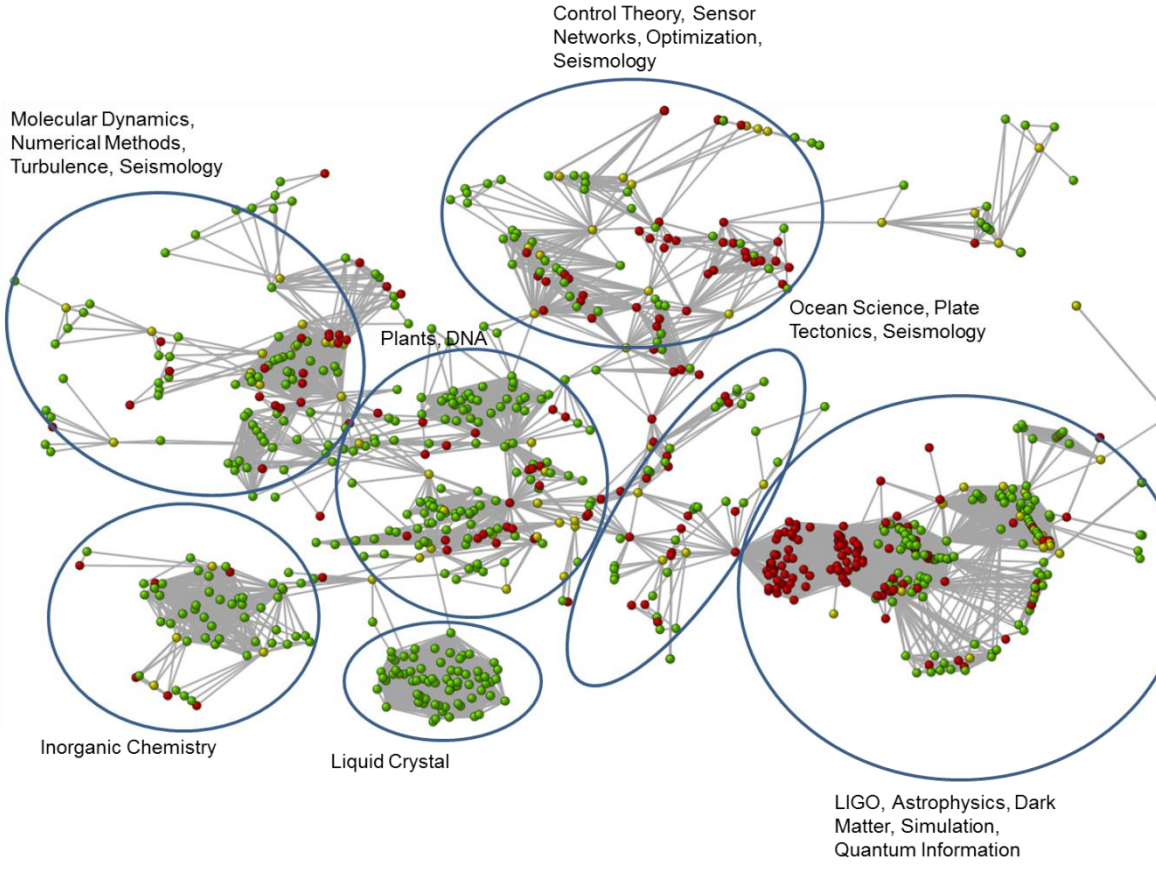
Agency, Institution and Public Portfolio Explorer

R&D Dashboard

SciENCV

Level I: Who is Supported by Science Funding

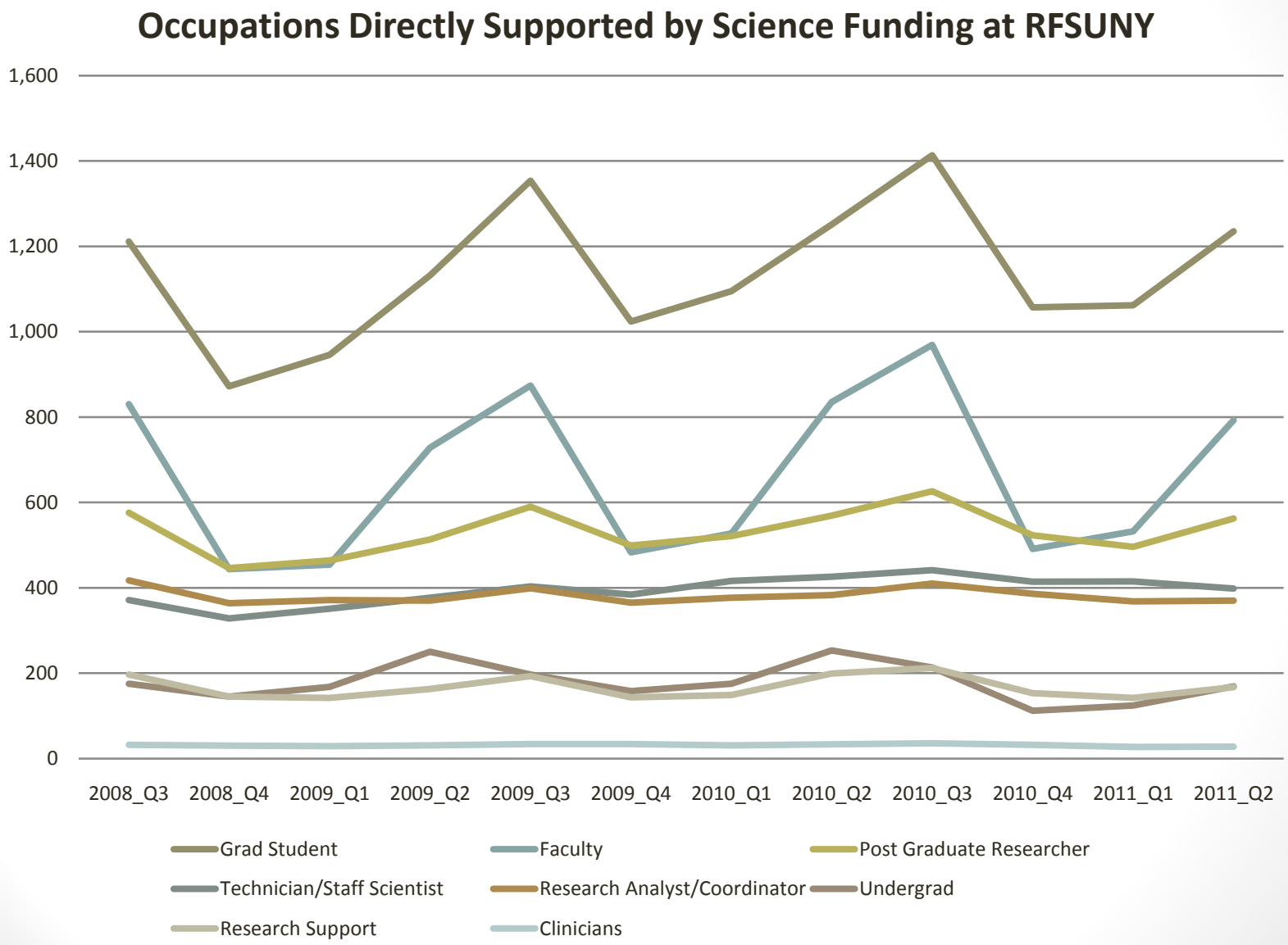
- Automated reporting from financial systems
- Documenting jobs
- And training



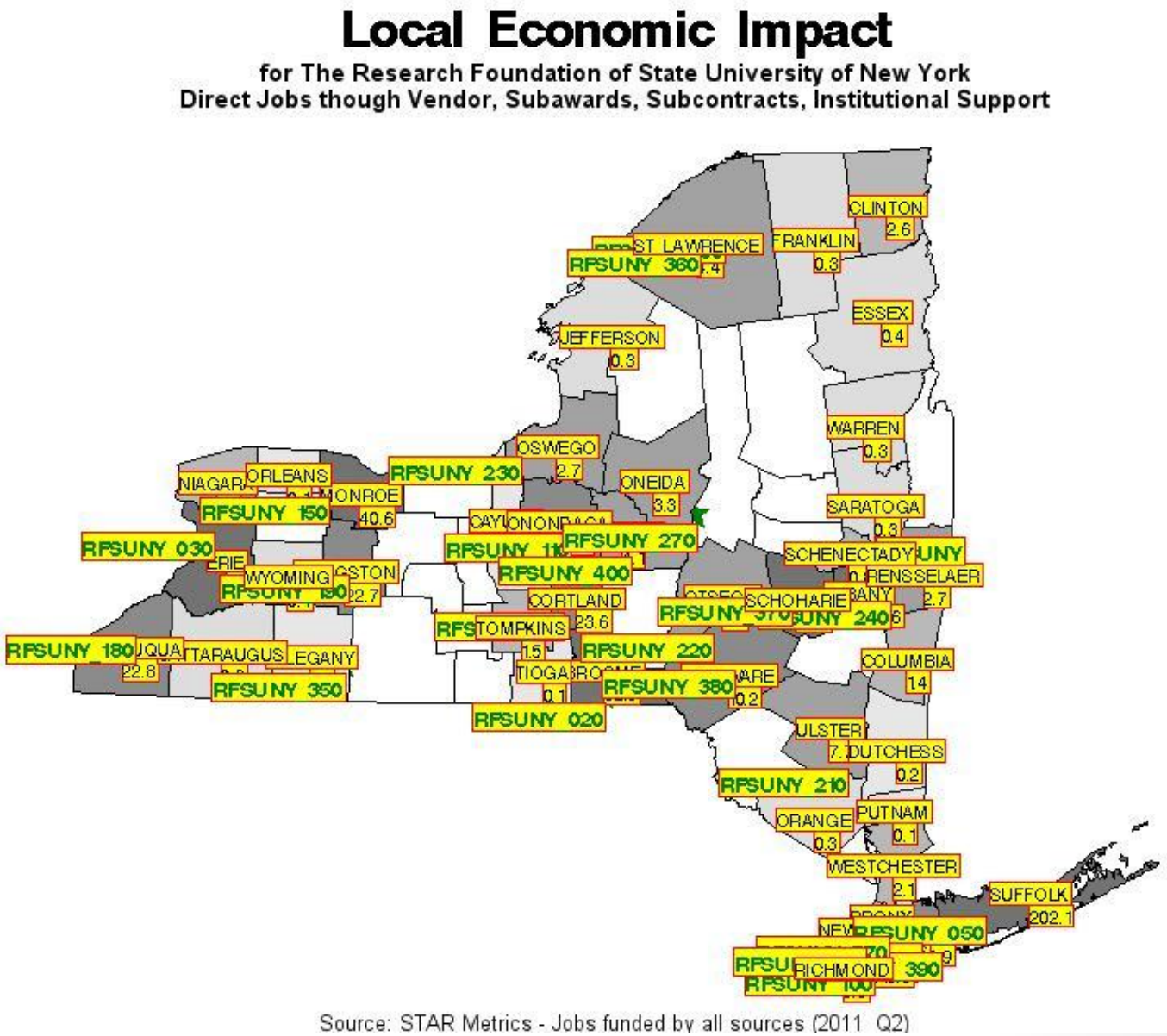
Science as an export industry

Number of active Federal science awards	1,447
Jobs	
Direct Payroll FTEs	
Counts	2,143.7
Direct Payroll Individuals	
Counts	7,444
Direct Jobs through Vendors, Sub-Award Recipients, Institutional Support	
Counts	2,029.2

Jobs at all levels

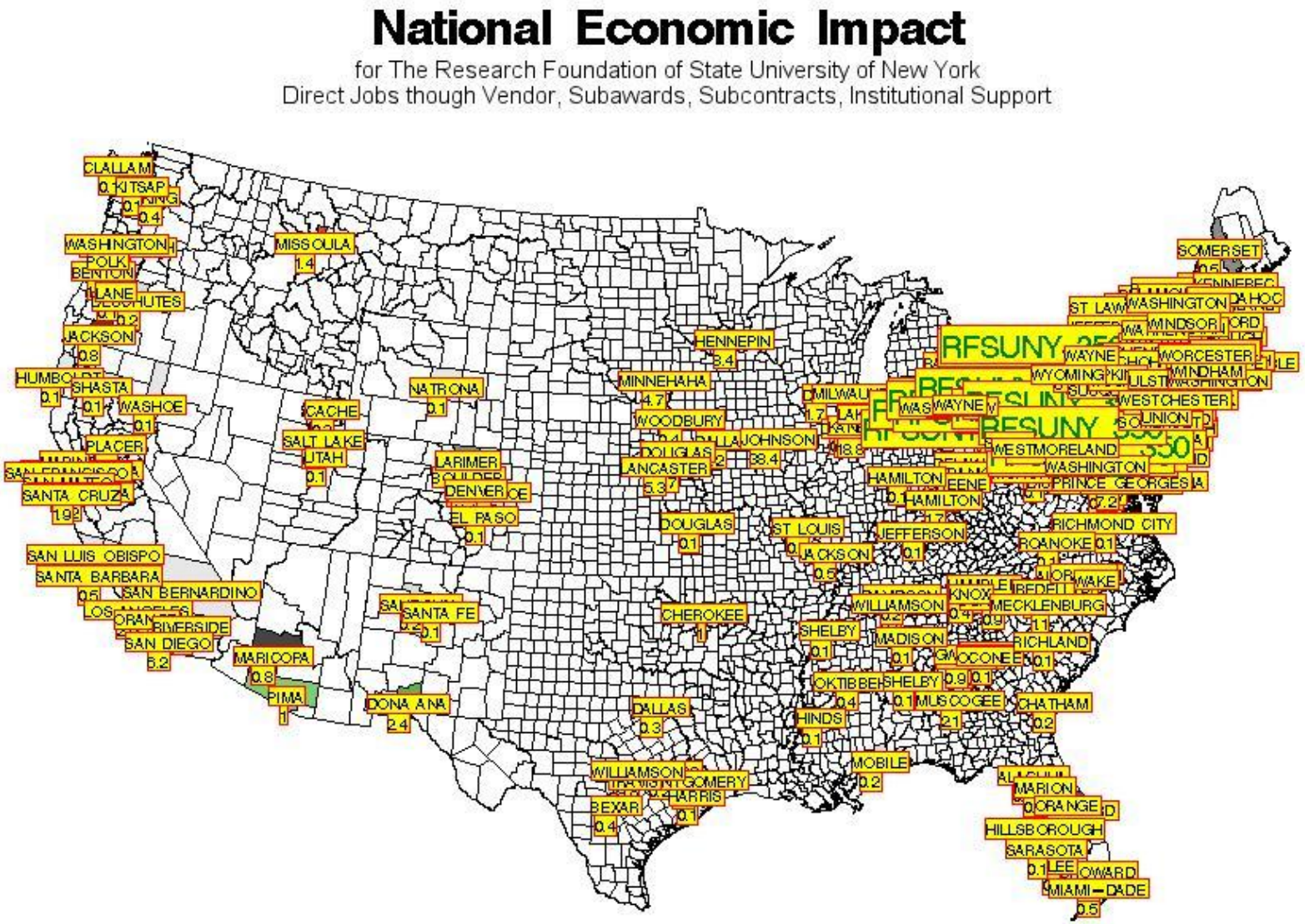


Jobs across the state



Source: STAR Metrics - Jobs funded by all sources (2011 Q2)

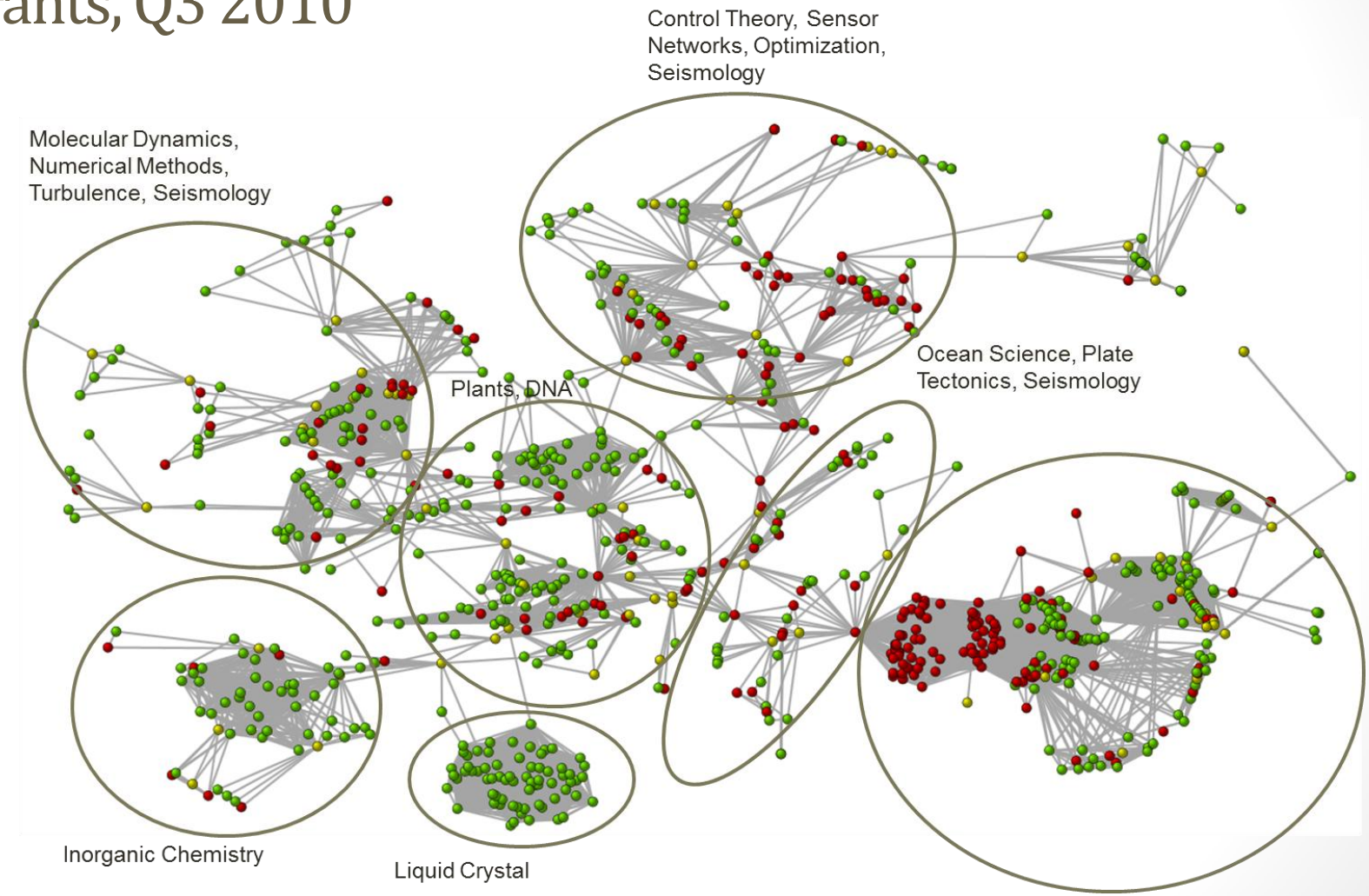
And jobs across the nation



Source: STAR Metrics - Jobs funded by all sources (2011_Q2)
Note: Map excludes Alaska, Hawaii, and Puerto Rico

Training:

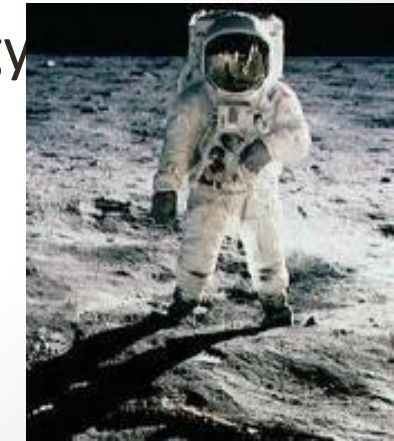
A private University's science staff paid by science agency grants, Q3 2010



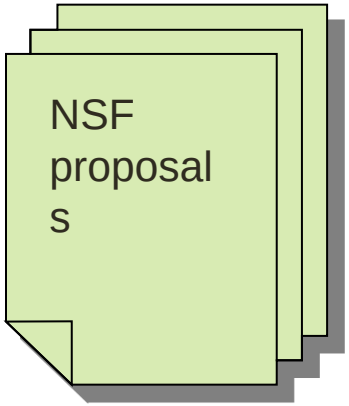
Source: Jason Owen-Smith

Level II: Approach

- A platform that can link inputs and outputs/outcomes using automated approaches
- Leverage existing data and knowledge (results of $60 million in investments)
- Collaborative development of data infrastructure on broad categories of impact:
  - knowledge (e.g. publication, citations…)
  - economic (patents, spin off companies…)
  - workforce (employment, student mobility…)
  - social (e.g. health, environment, energy



Automated capture of scientific topics



Topic Model:
- Use words from (all) text
- Learn T topics

Automatically learned topics (e.g.):

…

t6. conflict violence war international military …

t7. model method data estimation variables …

t8. parameter method point local estimates …

t9. optimization uncertainty optimal stochastic …

t10. surface surfaces interfaces interface …

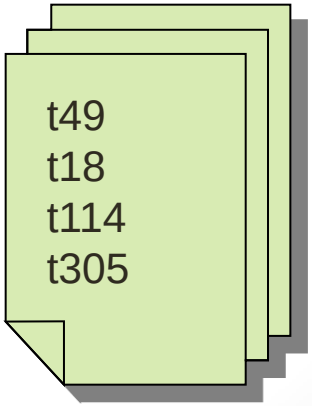
t11. speech sound acoustic recognition human …

t12. museum public exhibit center informal outreach

t13. particles particle colloidal granular material …

t14. ocean marine scientist oceanography …

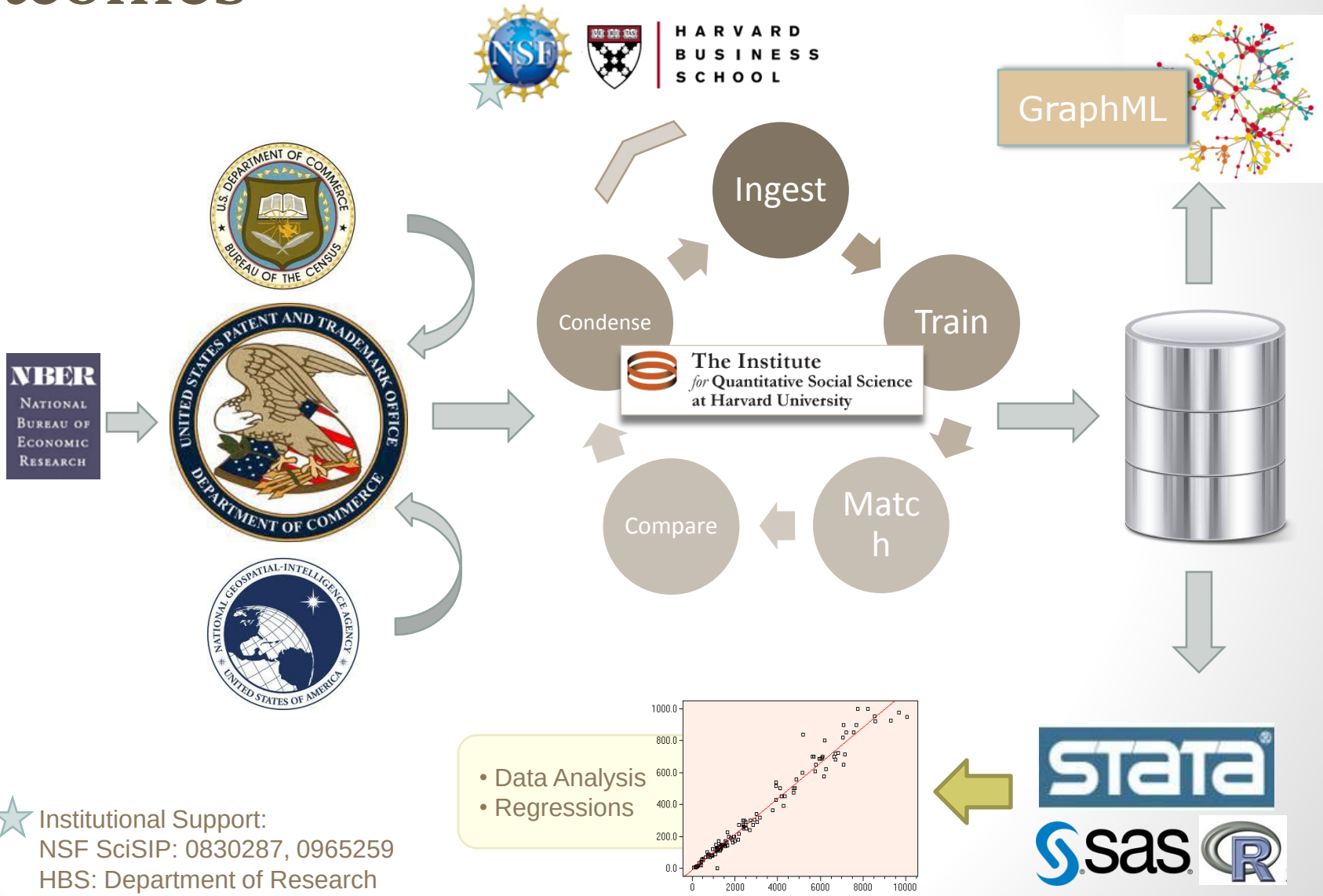
…



Topic tags for each and every proposal

David Newman - UC Irvine

Automated capture of economic outcomes



Institutional Support:
NSF SciSIP: 0830287, 0965259
HBS: Department of Research

What is being done: Specifics

- Science based Portfolio Characterization: Portfolio Explorer
- Place based Portfolio Reporting
- Reduced scientist burden: SciENCV platform

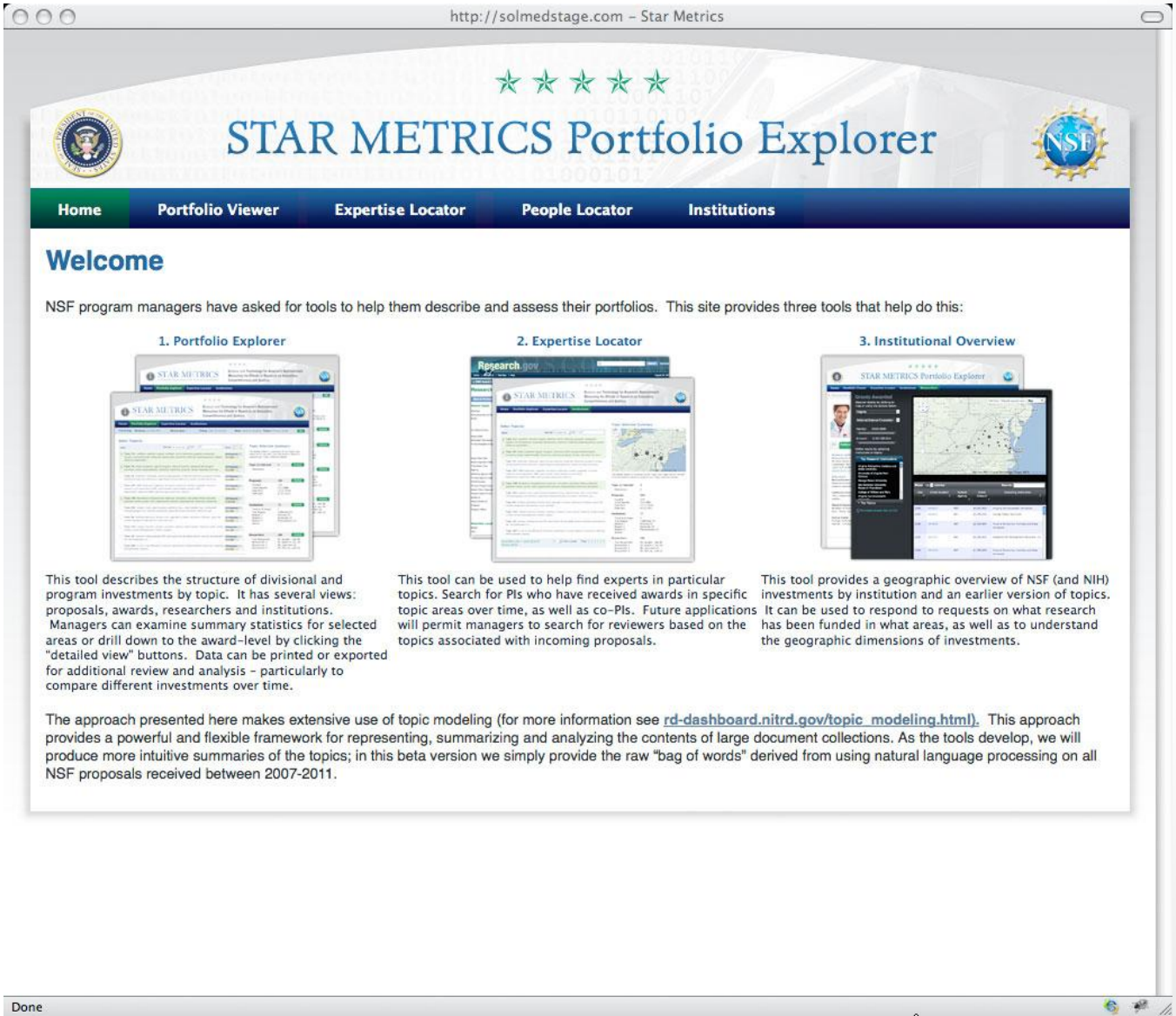
Key principles

- Build an open, transparent and automated system
- Facilitate the building of apps

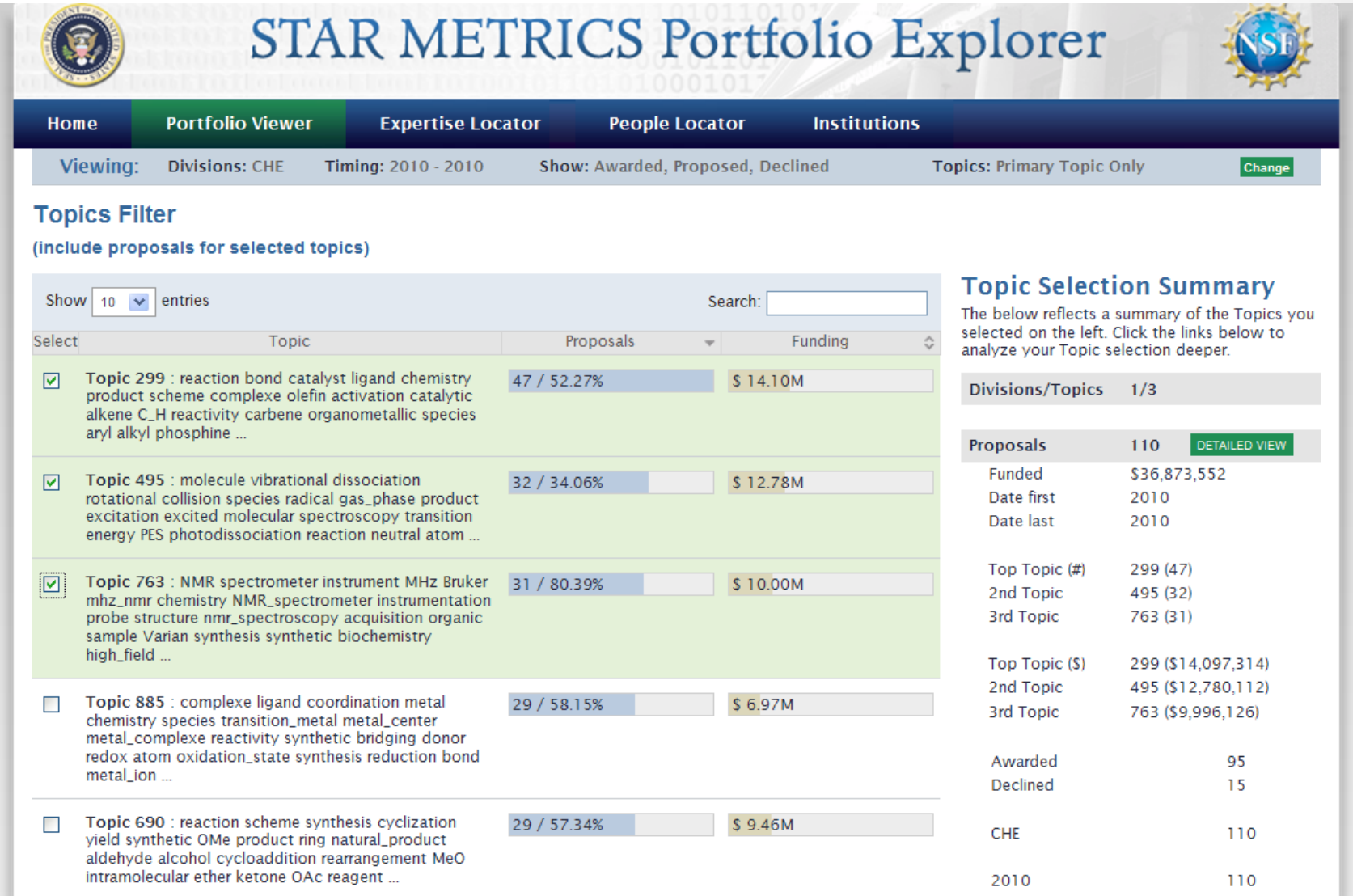
Science Based Portfolio Characterization

- For agencies
  - Gap analysis: What is being funded in which areas?
  - Expertise Locator: Who is doing research in which topics?
- For Researchers
  - Funding information: What programs are funding research like mine?
  - Expertise Locator: Who else is doing research like mine?
- For VPs for Research and their Institutions
  - Gap analysis: Where are my institutional research strengths?
  - Expertise Locator: How can I connect researchers?

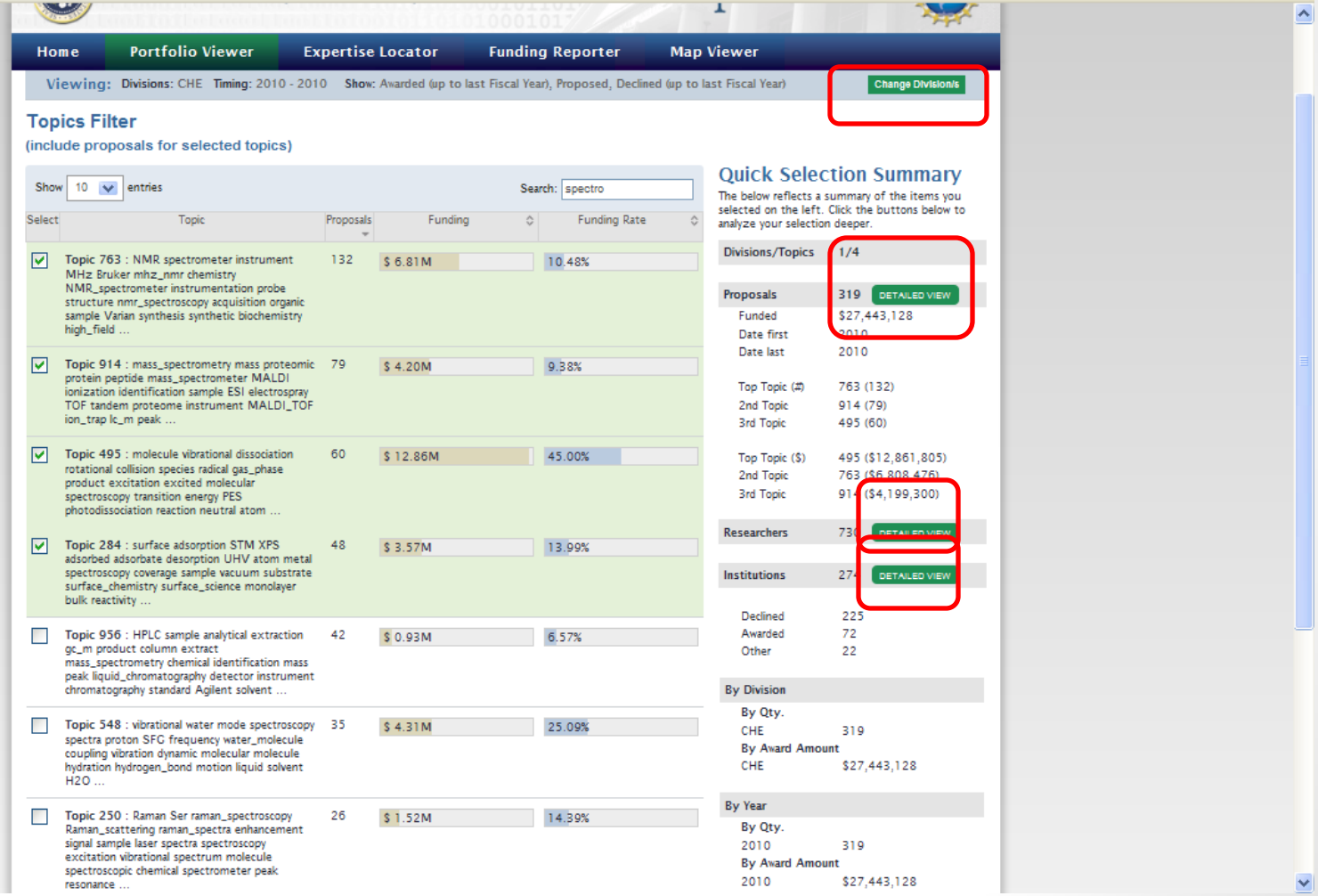
Portfolio Characterization



Describing the Chemistry Portfolio



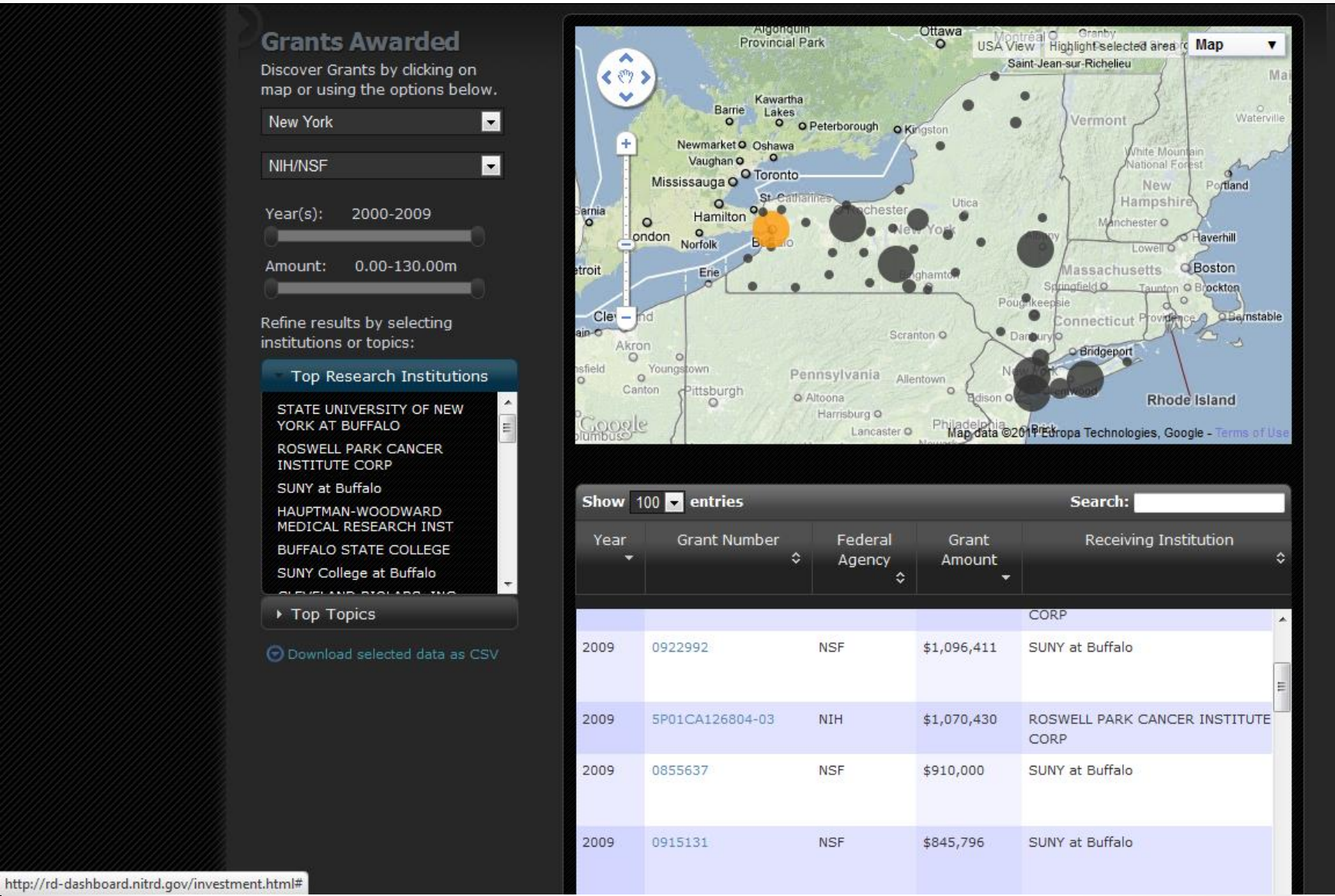
Describing the Chem Portfolio

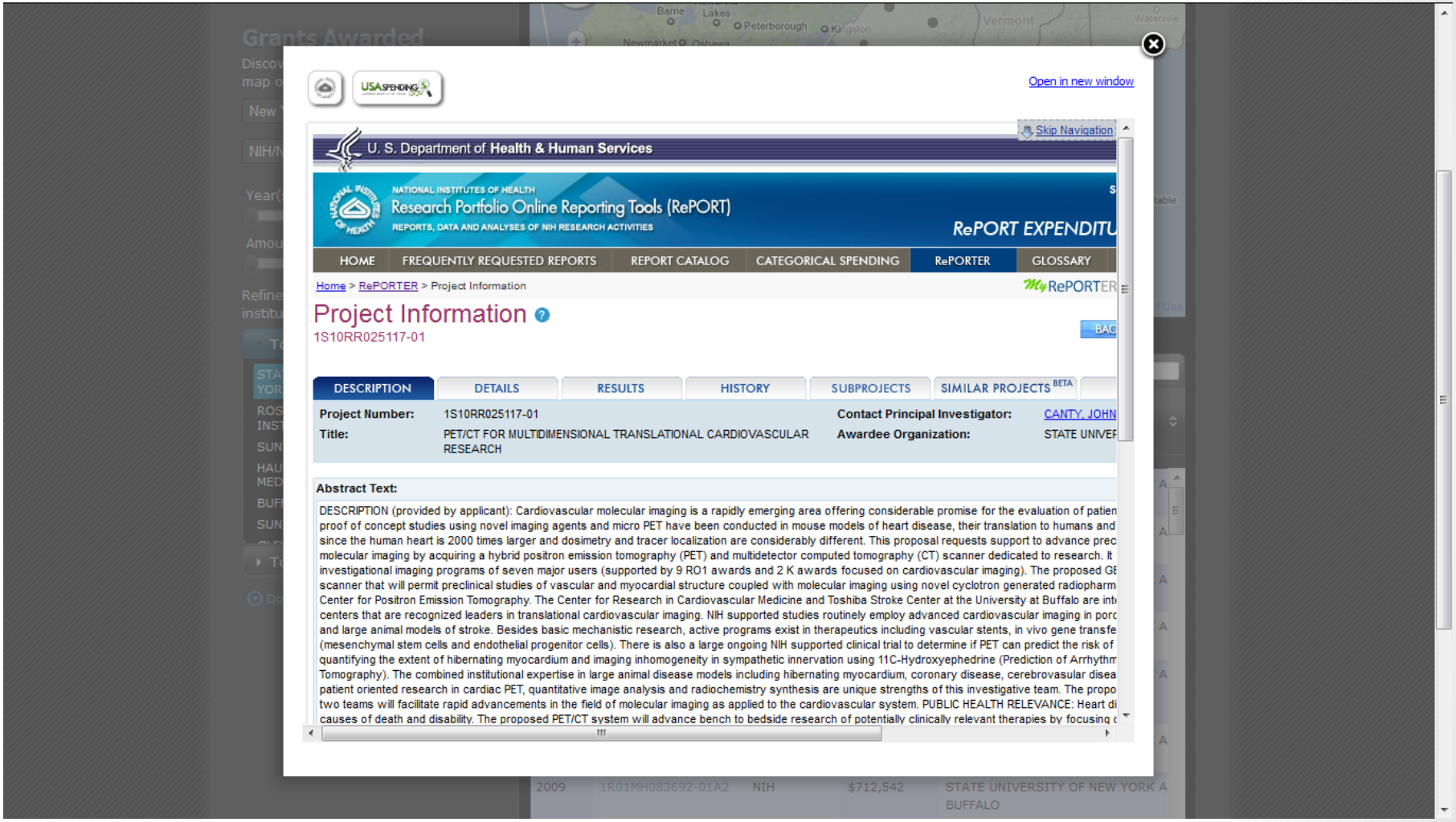


Place Based Portfolio Reporting: R&D Dashboard

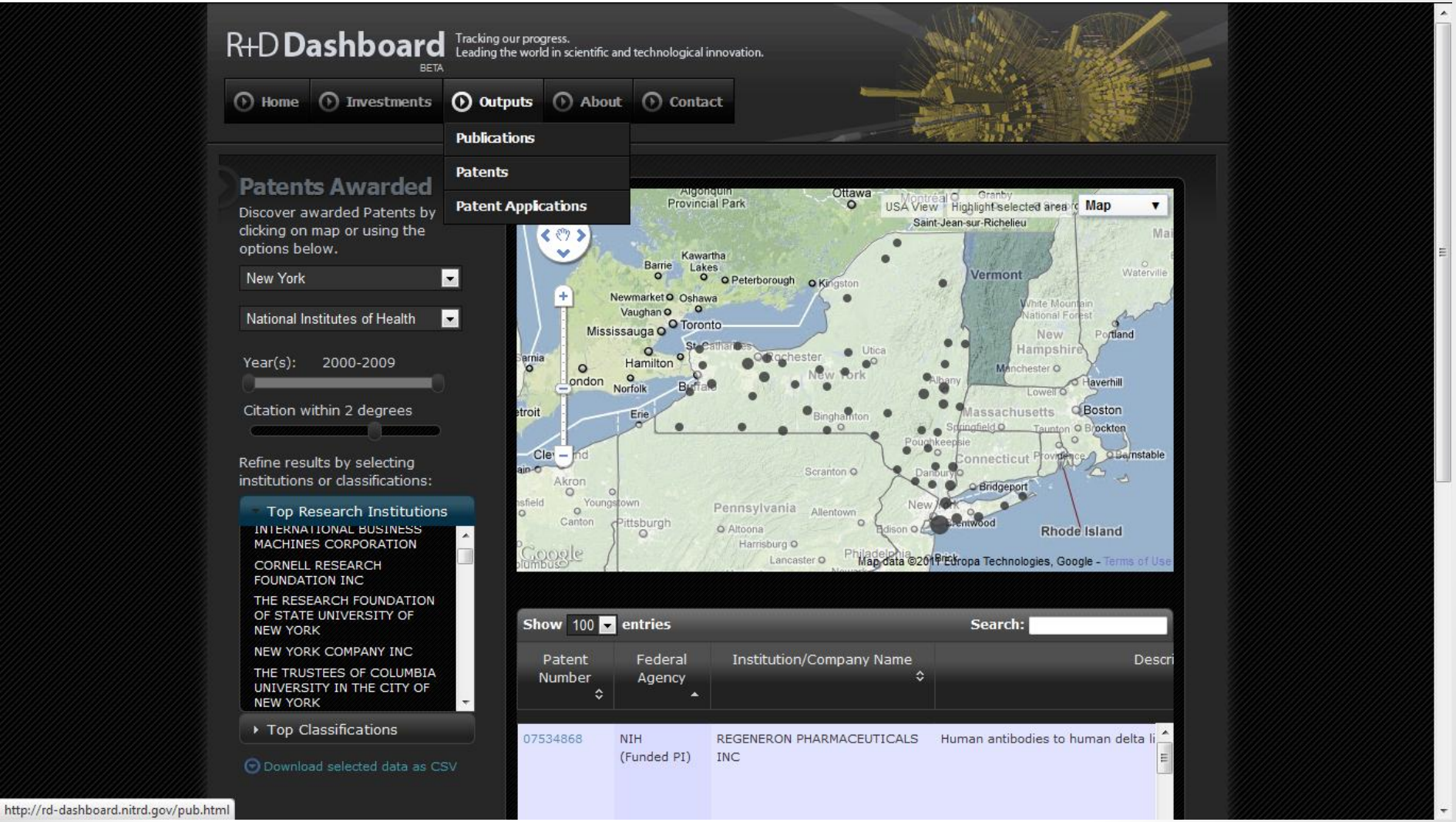
- For stakeholders
  - What research has been funded in my state/city?
  - Who are the researchers doing the research?
  - What are the results?
- For agencies/research institutions
  - Automated reporting of research portfolio
  - Automated documentation of institutional contribution
- For researchers
  - Minimal burden
  - Increased visibility to agencies, peers and stakeholders

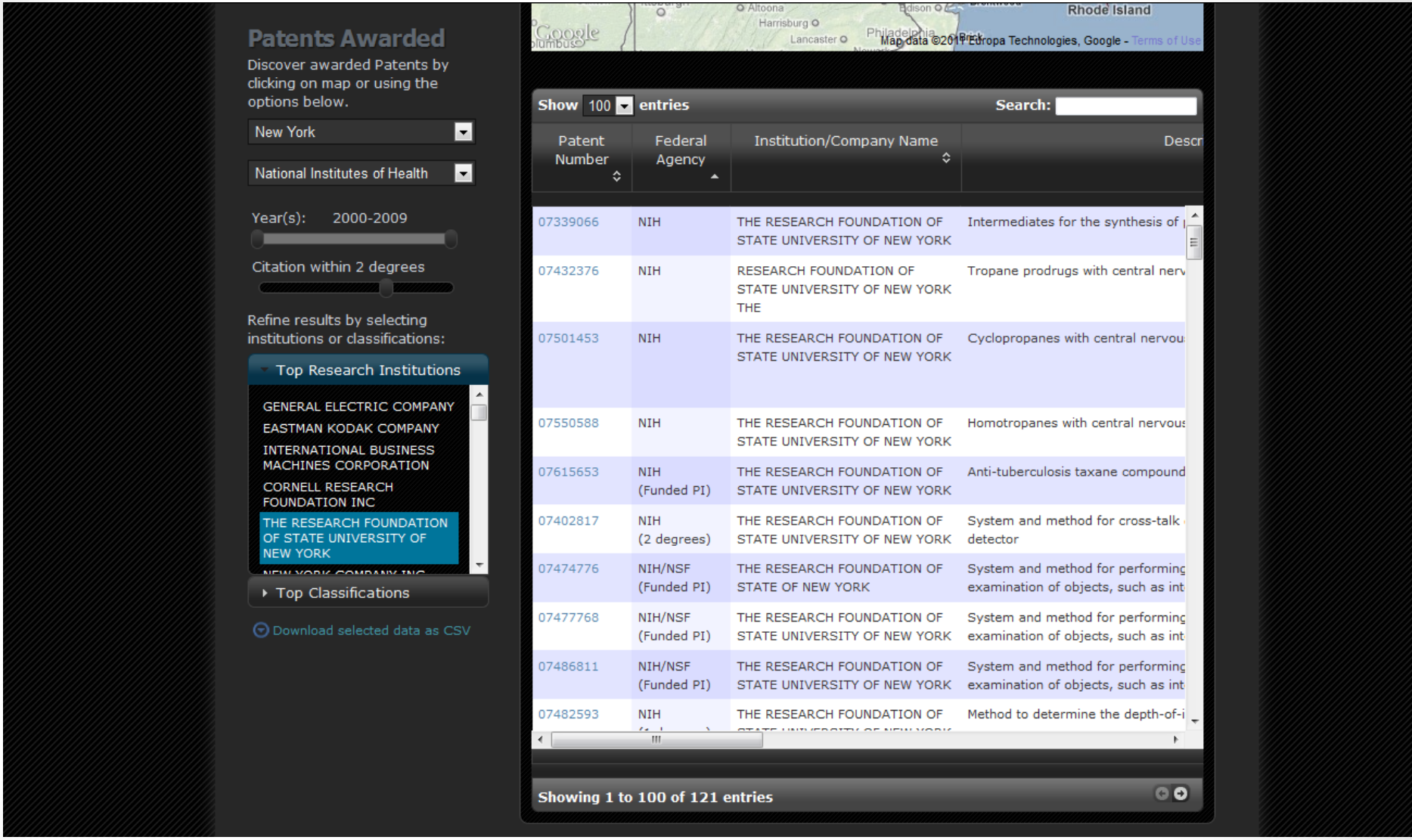
Place-based investments

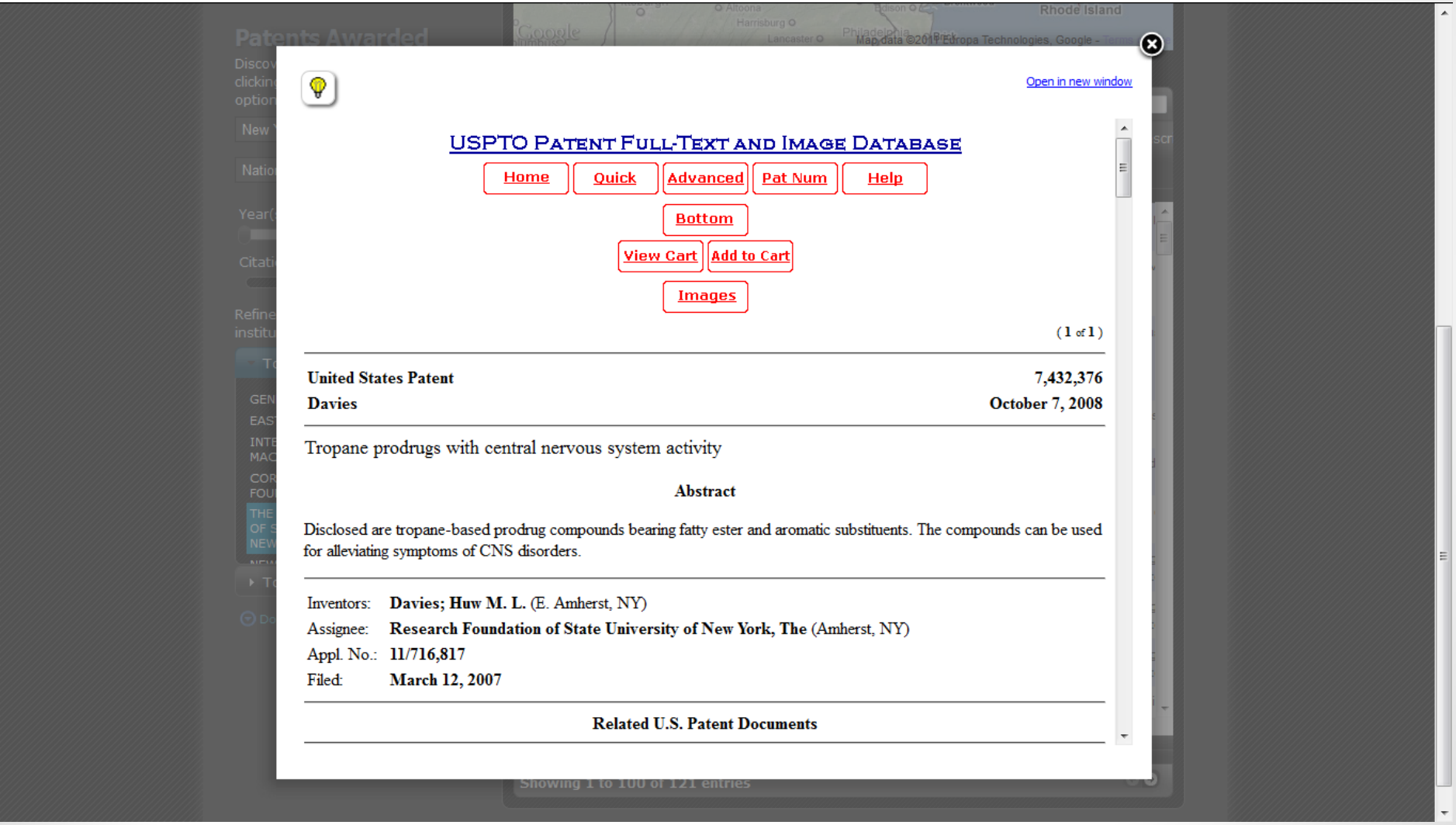



Open in new window
Skip Navigation
U. S. Department of Health & Human Services
NATIONAL INSTITUTES OF HEALTH
Research Portfolio Online Reporting Tools (RePORT)
REPORTS, DATA AND ANALYSES OF NIH RESEARCH ACTIVITIES
RePORT EXPENDITU
HOME
FREQUENTLY REQUESTED REPORTS
REPORT CATALOG
CATEGORICAL SPENDING
RePORTER
GLOSSARY
Home > RePORTER > Project Information
MyRePORTER
Project Information
1S10RR025117-01
DESCRIPTION
DETAILS
RESULTS
HISTORY
SUBPROJECTS
SIMILAR PROJECTS BETA
Project Number: 1S10RR025117-01
Title: PET/CT FOR MULTIDIMENSIONAL TRANSLATIONAL CARDIOVASCULAR RESEARCH
Contact Principal Investigator: CANTY, JOHN
Awardee Organization: STATE UNIVEF
Abstract Text:
DESCRIPTION (provided by applicant): Cardiovascular molecular imaging is a rapidly emerging area offering considerable promise for the evaluation of patien
proof of concept studies using novel imaging agents and micro PET have been conducted in mouse models of heart disease, their translation to humans and
since the human heart is 2000 times larger and dosimetry and tracer localization are considerably different. This proposal requests support to advance prec
molecular imaging by acquiring a hybrid positron emission tomography (PET) and multidetector computed tomography (CT) scanner dedicated to research. It
investigational imaging programs of seven major users (supported by 9 RO1 awards and 2 K awards focused on cardiovascular imaging). The proposed GE
scanner that will permit preclinical studies of vascular and myocardial structure coupled with molecular imaging using novel cyclotron generated radiopharm
Center for Positron Emission Tomography. The Center for Research in Cardiovascular Medicine and Toshiba Stroke Center at the University at Buffalo are int
centers that are recognized leaders in translational cardiovascular imaging. NIH supported studies routinely employ advanced cardiovascular imaging in porc
and large animal models of stroke. Besides basic mechanistic research, active programs exist in therapeutics including vascular stents, in vivo gene transfe
(mesenchymal stem cells and endothelial progenitor cells). There is also a large ongoing NIH supported clinical trial to determine if PET can predict the risk of
quantifying the extent of hibernating myocardium and imaging inhomogeneity in sympathetic innervation using 11C-Hydroxyephedrine (Prediction of Arrhythm
Tomography). The combined institutional expertise in large animal disease models including hibernating myocardium, coronary disease, cerebrovasular disea
patient oriented research in cardiac PET, quantitative image analysis and radiochemistry synthesis are unique strengths of this investigative team. The propo
two teams will facilitate rapid advancements in the field of molecular imaging as applied to the cardiovascular system. PUBLIC HEALTH RELEVANCE: Heart di
causes of death and disability. The proposed PET/CT system will advance bench to bedside research of potentially clinically relevant therapies by focusing
2009 1R01MH083692-01A2 NIH $712,542 STATE UNIVERSITY OF NEW YORK A BUFFALO

Place based results




Patents Awarded
Discover awarded Patents by clicking on map or using the options below.
New York
National Institutes of Health
Year(s): 2000-2009
Citation within 2 degrees
Refine results by selecting institutions or classifications:
Top Research Institutions
GENERAL ELECTRIC COMPANY
EASTMAN KODAK COMPANY
INTERNATIONAL BUSINESS MACHINES CORPORATION
CORNELL RESEARCH FOUNDATION INC
THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK
Top Classifications
Download selected data as CSV
Altoona
Harrisburg
Lancaster
Philadelphia
Rhode Island
Map data ©2011 Europa Technologies, Google - Terms of Use
Show 100 entries
Search:
Patent Number
Federal Agency
Institution/Company Name
Descr
07339066 NIH THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK Intermediates for the synthesis of
07432376 NIH RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK THE Tropane prodrugs with central nerv
07501453 NIH THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK Cyclopropanes with central nervou
07550588 NIH THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK Homotropanes with central nervous
07615653 NIH (Funded PI) THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK Anti-tuberculosis taxane compound
07402817 NIH (2 degrees) THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK System and method for cross-talk detector
07474776 NIH/NSF (Funded PI) THE RESEARCH FOUNDATION OF STATE OF NEW YORK System and method for performing examination of objects, such as int
07477768 NIH/NSF (Funded PI) THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK System and method for performing examination of objects, such as int
07486811 NIH/NSF (Funded PI) THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK System and method for performing examination of objects, such as int
07482593 NIH THE RESEARCH FOUNDATION OF Method to determine the depth-of-i
Showing 1 to 100 of 121 entries


Patents Awarded
Rhode Island
Map data ©2011 Europa Technologies, Google -
Open in new window
USPTO Patent Full-Text and Image Database
Home
Quick
Advanced
Pat Num
Help
Bottom
View Cart
Add to Cart
Images
(1 of 1)
United States Patent
7,432,376
Davies
October 7, 2008
Tropane prodrugs with central nervous system activity
Abstract
Disclosed are tropane-based prodrug compounds bearing fatty ester and aromatic substituents. The compounds can be used for alleviating symptoms of CNS disorders.
Inventors: Davies; Huw M. L. (E. Amherst, NY)
Assignee: Research Foundation of State University of New York, The (Amherst, NY)
Appl. No.: 11/716,817
Filed: March 12, 2007
Related U.S. Patent Documents

Reducing Scientist Burden

SciENCV is a Science Experts Network and CV system intended to create value for researchers, agencies and research institutions by providing an automated approach to collecting information on researchers and their accomplishments. This system will:

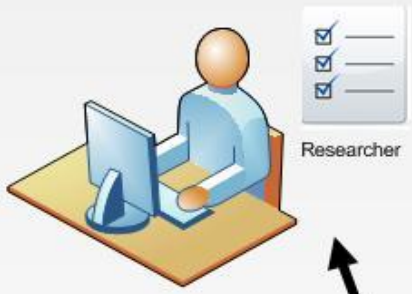
Streamline application preparation and reporting
Provide a resource for locating scientific expertise
Support analysis of results of research investments

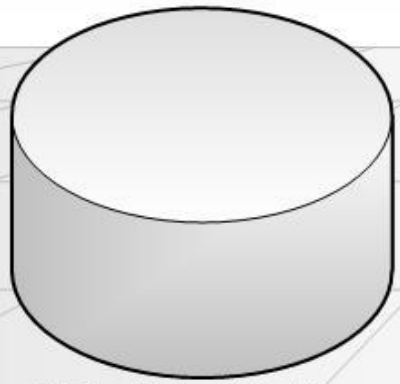
SciENCV is driven by two working groups from the National Science and Technology Council's Committee on Science: Research Business Models and Science of Science Policy, based on input from the research community.

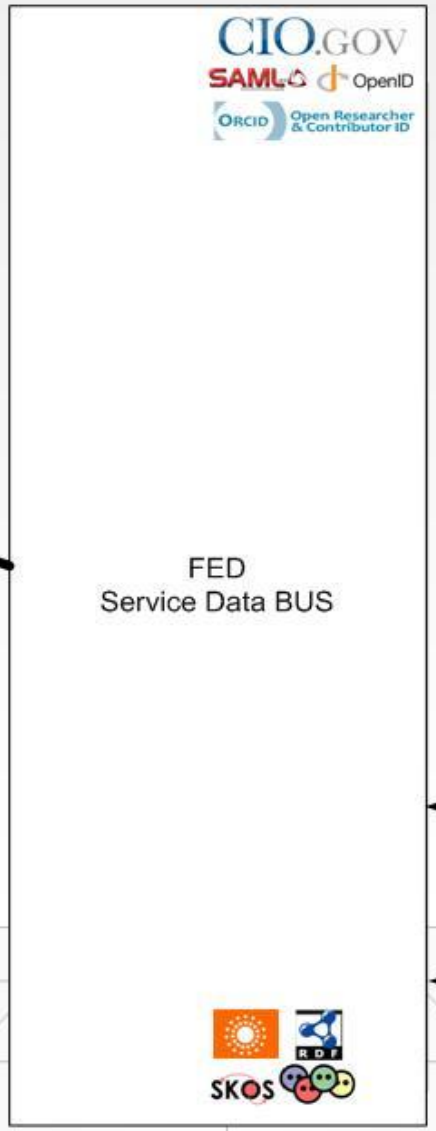
US FEDERAL RESEARCH PROFILE SYSTEM

STAR METRICS TECHNICAL MEETING

April 22-23 2011 Brazil-US Technical Meeting







SALM JR., JOSÉ
FRANCISCO

SciEN CV Platform Next steps

- Federal Demonstration Partnership Pilot
- Lattes http://www.slideshare.net/rpacheco/sti-national-information-system-platform-the-brazilian-case-of-lattes
- http://www.slideshare.net/rpacheco/sti-information-systems-brazilian-initiatives-frequently-asked-questions
- ORCID grant

The pitfalls

- Need to paint full picture of scientific outcomes
  - => engagement of scientific community critical
  - => Open and transparent process
- Data misuse
  - => careful presentation of results
- Data quality
  - => full collaboration
  - => extensive use of pilots
- Confidentiality
  - => researcher, institution and agency controls

Ultimate Goals

- Fully fledged academic field
- Fully fledged analytical tool set in government
  - Science policy in same analytical tier as tax policy
- Common empirical infrastructure available to all universities and science agencies to quickly respond to State, Congressional and OMB requests
- Common scientific infrastructure for researchers to develop and study science policy

Thank you

- Comments and questions?