Lane, Julia

Working Paper

Science metrics: The issues and new approaches

Working Paper Series des Rates für Sozial- und Wirtschaftsdaten, No. 159

Provided in Cooperation with:
German Data Forum (RatSWD)

Suggested Citation: Lane, Julia (2010) : Science metrics: The issues and new approaches, Working Paper Series des Rates für Sozial- und Wirtschaftsdaten, No. 159, German Data Forum (RatsWD), Berlin

This Version is available at:
http://hdl.handle.net/10419/43604

Standard-Nutzungsbedingungen:
Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.
Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.
Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Science Metrics:
The Issues and New Approaches

Julia Lane

September 2010
Working Paper Series of the German Data Forum (RatSWD)

The RatSWD Working Papers series was launched at the end of 2007. Since 2009, the series has been publishing exclusively conceptual and historical works dealing with the organization of the German statistical infrastructure and research infrastructure in the social, behavioral, and economic sciences. Papers that have appeared in the series deal primarily with the organization of Germany’s official statistical system, government agency research, and academic research infrastructure, as well as directly with the work of the RatSWD. Papers addressing the aforementioned topics in other countries as well as supranational aspects are particularly welcome.

RatSWD Working Papers are non-exclusive, which means that there is nothing to prevent you from publishing your work in another venue as well: all papers can and should also appear in professionally, institutionally, and locally specialized journals. The RatSWD Working Papers are not available in bookstores but can be ordered online through the RatSWD.

In order to make the series more accessible to readers not fluent in German, the English section of the RatSWD Working Papers website presents only those papers published in English, while the German section lists the complete contents of all issues in the series in chronological order.

Starting in 2009, some of the empirical research papers that originally appeared in the RatSWD Working Papers series will be published in the series RatSWD Research Notes.

The views expressed in the RatSWD Working Papers are exclusively the opinions of their authors and not those of the RatSWD.

The RatSWD Working Paper Series is edited by:

Chair of the RatSWD (2007/2008 Heike Solga; since 2009 Gert G. Wagner)

Managing Director of the RatSWD (Denis Huschka)
This paper documents the presentation slides of the 1st Distinguished Lecture of the German Data Forum (RatSWD), held on 15th September 2010 at the DIW Berlin.
Science Metrics: The Issues and New Approaches

Julia Lane

This presentation represents the views of the author and not of the institution she represents.

Overview

- Why Metrics Matter
- Conceptual Framework
  - The scientific challenge
  - The empirical challenge
- What’s Being Done in the US: STAR METRICS
  - What it is
  - Structure
  - Measuring outcomes: The Role of Incentives
  - Examining impact: The Role of Social and Domain Scientists
Why metrics matter

- Government
  - Advance basic science
  - Improve wellbeing of citizens
  => Affects level of funding
- Funding agencies
  - Want to identify and fund good science
  => Affects type of funding
- Academic institutions
  - Want to hire and retain good scientists
  - Want to demonstrate impact
  => Affects who does science

Administration Interest

- **Investment in Science**
  - American Recovery and Reinvestment Act
  - The National Academy of Sciences Speech, April 2009
- **Openness and transparency**
  - data.gov; open.gov; etc.
- **Evidence based policy**
  - Joint memo on “Science and Technology Priorities for the FY2012 Budget”: *Science of Science Policy* (is the only program listed by name – also in 2011)
- **Accountability**
  - ARRA Reporting Guidelines
  - *Putting Performance First: Replacing PART with a new performance improvement and analysis framework*
Administration Interest

Agencies, in cooperation with OSTP and OMB, should develop and sustain datasets to better document Federal science, technology, and innovation investments and to make these data open to the public in accessible, useful formats. Agencies should develop and regularly update their data sharing policies for research performers and create incentives for sharing data publicly in interoperable formats to ensure maximum value, consistent with privacy, national security, and confidentiality concerns.

Agencies should develop outcome-oriented goals for their science, technology, and innovation activities, establish timelines for evaluating the performance of these activities, and target investments toward high-performing programs in their budget submissions. Agencies should support the development and use of “science of science policy” tools that can improve management of their R&D portfolios and better assess the impact of their science, technology, and innovation investments.

FY12 Orszag-Holdren Memo, July 21 2010; reiterates August 4, 2009 memo; Science of Science Policy is only program mentioned by name

Congressional Interest
Public Interest

Jobs Matter

Mean what you say

What science is really worth

International Interest

EU2009.CZ

The global challenge

What science is really worth

Spending on science is one of the best ways to generate jobs and economic growth, say research advocates. But as Colin MacKean reports, the evidence behind such claims is spotty.
Scientists Can Provide a ‘Black Box’ Answer

Or…Start To Develop A Scientific Framework

- Science of Science Policy Interagency Task Group
- The SoSP Roadmap
  - Published in November, 2008
  - Four guiding themes
  - Ten key questions
- December, 2008 Workshop
  - Engage the current community of practice
  - Interactive evaluation of Roadmap
Research Challenge: Conceptual

Need to describe and measure the creation, transmission and adoption of knowledge

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Defining the Set of Hypotheticals or Counterfactuals</td>
<td>A Scientific Theory</td>
</tr>
<tr>
<td>2</td>
<td>Identifying Causal Parameters from Hypothetical Population Data</td>
<td>Mathematical Analysis of Point or Set Identification</td>
</tr>
<tr>
<td>3</td>
<td>Identifying Parameters from Real Data</td>
<td>Estimation and Testing Theory</td>
</tr>
</tbody>
</table>


Research Challenges: Conceptual

- How to describe creation of knowledge?
  - Unit of analysis
  - Input measures
- How to describe transmission?
  - Networks
  - Technology
- How to describe adoption?
  - Lags
  - Proximal causes
- What structural model?
  - Linear
  - Outcome measures
- Fundamental challenge: Establishing counterfactuals
  - Selection bias
  - Random assignment not an option
Research Challenges: Empirical

- **Data Infrastructure**
  - Science agencies have balkanized proposal and award administration systems
  - Unit of analysis is awards – while appropriate unit is individuals
  - Typically limited data on postdocs, graduate students, undergraduate students
  - Limited data on subawards
  - Information captured only during funding period
  - Information typically captured manually, sporadically and in unstructured format
  - Outputs not linked to inputs or infrastructure investments in a systematic way.
  - Data not captured on people who DON’T get funded, so difficult to establish counterfactual

- **Heterogeneous sources of outcomes**

- **Changing nature of scientific communication**

- **Scientific Attribution**
  - Name disambiguation
  - Global enterprise

If we can automate the DNA sequencing, we can describe science investments!
Science and Technology in America’s Reinvestment – Measuring the Effects of Research on Innovation, Competitiveness and Science

What is STAR METRICS?

1. Data Infrastructure to capture impact of science investments.
2. Collaborative identification of data and data sources
3. Explicit integration of domain and social scientists in development of metrics
Basic Approach

Creating the Frame

- Start with basic unit of analysis
  - Science is done by scientists. Need to identify universe of individuals funded by federal agencies (PI, co-PI, RAs, graduate students etc.)
- Capture Inputs using existing data

Measuring outcomes

- Scientific
- Social
- Economic
- Workforce
Based on Existing Record Reporting

STAR Pilot Project

Institution

Agency

Budget

Award

Record

Endowment Funding

Research Project

Financial System

Disbursement

HR System

Procurement System

Subcontracting System

Papers

Patents

Start-Up

Jobs, Purchases, Contracts Benefit Analysis

Detailed Characterization and Summary

Download

Acquisition And Analysis

Direct Benefit Analysis

Intellectual Property Benefit Analysis

Innovation Analysis

Agency

State Funding

Institution

Financial System

Disbursement

HR System

Procurement System

Subcontracting System

Personnel

Vendor

Contractor

Hire

Buy

Engage

Existing Institutional Reporting

State
Creating the Frame (and measuring jobs)

14 administrative data elements from awards, grants, HR or finance systems are provided to STAR Metrics on a quarterly basis...
- Award data
- Payroll Staff Information
- Non-Payroll Charges
- Sub-awards
- Indirect Cost Rate Proposal

...will yield these Quarterly pre-calculated reports...
- **Stimulus FTE Jobs** (ARRA) – with and without Overhead Job calculations
- **FTE Jobs and Positions** – All awards (with and without Overhead)
- **FTE Sub-awards** – All awards (with and without Overhead)
- **Vendor FTE’s (Jobs)** – All awards
- **Overhead Jobs** (calculated from Indirect Costs)
### Star Metrics Phase 1 – 14 Requested Data Elements

<table>
<thead>
<tr>
<th>Description</th>
<th>Element ID</th>
<th>Item</th>
<th>Data Source</th>
<th>Unit of Analysis</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information on Scientists and Awards</td>
<td>1</td>
<td>De-identified Employee ID #</td>
<td>University</td>
<td>Individual</td>
<td>Job Metrics</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Federal Award ID #</td>
<td>University</td>
<td>Award</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>University Award ID #</td>
<td>University</td>
<td>Award</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Overhead charged</td>
<td>University</td>
<td>Award</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Occupational Classification</td>
<td>Individual</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Proportion of time allocated to award</td>
<td>Individual</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>FTE status</td>
<td>Individual</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information on Overhead</td>
<td>8</td>
<td>Proportion of overhead associated with salaries (from overhead cost proposal)</td>
<td>University</td>
<td>University</td>
<td>Job Metrics</td>
</tr>
<tr>
<td>Payments to vendors</td>
<td>2</td>
<td>Federal Award ID #</td>
<td>University</td>
<td>Award</td>
<td>Secondary Economic Impact</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>University Award ID #</td>
<td>University</td>
<td>Award</td>
<td>Secondary Economic Impact</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Duns #</td>
<td>Vendor</td>
<td></td>
<td>Vendor</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>Amount of Contract</td>
<td>Vendor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subcontracts and subawards</td>
<td>2</td>
<td>Federal Award ID #</td>
<td>University</td>
<td>Award</td>
<td>Secondary Economic Impact</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>University Award ID #</td>
<td>University</td>
<td>Award</td>
<td>Secondary Economic Impact</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Duns #</td>
<td>Subcontractor</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>Amount of Contract</td>
<td>Subcontractor</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Local Economic Impact**

**For UNIVERSITY OF MASSACHUSETTS DARTMOUTH**

Total Jobs (SIMULATED DATA)

Source: STAR Metrics - Jobs
Local Economic Impact
for UNIVERSITY OF MASSACHUSETTS
DARTMOUTH
Total Jobs (SIMULATED DATA)

<table>
<thead>
<tr>
<th>County Name</th>
<th>County Code</th>
<th>Sub-Awards &amp; Vendor Jobs</th>
<th>Award FTEs, Sub-Award &amp; Vendor Jobs</th>
<th>Total Jobs</th>
</tr>
</thead>
<tbody>
<tr>
<td>BARNSTABLE</td>
<td>1</td>
<td>76</td>
<td>76</td>
<td>76</td>
</tr>
<tr>
<td>BERKSHIRE</td>
<td>3</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>BRISTOL</td>
<td>5</td>
<td>100.7</td>
<td>861.4</td>
<td>931.1</td>
</tr>
<tr>
<td>DUKES</td>
<td>7</td>
<td>49.5</td>
<td>49.5</td>
<td>49.5</td>
</tr>
<tr>
<td>ESSEX</td>
<td>9</td>
<td>268.7</td>
<td>268.7</td>
<td>268.7</td>
</tr>
<tr>
<td>MIDDLESEX</td>
<td>17</td>
<td>123.8</td>
<td>123.8</td>
<td>123.8</td>
</tr>
<tr>
<td>NANTUCKET</td>
<td>19</td>
<td>5.8</td>
<td>5.8</td>
<td>5.8</td>
</tr>
<tr>
<td>NORFOLK</td>
<td>21</td>
<td>16.3</td>
<td>16.3</td>
<td>16.3</td>
</tr>
</tbody>
</table>

Source: STAR Metrics - Jobs

Initial Jobs Impact of Science Expenditures
for 5 universities
Total Jobs

Source: STAR Metrics - Jobs Q3 2009 - Q2 2010 (fuzz factor applied)
Note: Map excludes Alaska, Hawaii, and Puerto Rico.
Measuring Outcomes: The Role of Incentives

1. Reduce Burden
2. Leverage Existing Data
3. Describe Impact
Reducing Burden:
Use Existing Reports

NATIONAL SCIENCE FOUNDATION
Agency Information Collection Activities: Comment Request

AGENCY: National Science Foundation (NSF).

ACTION: Submission for OMB Review; Comment Request and Final Notice of a Uniform Research Performance Progress Report (RPPR) format.

SUMMARY: Effective with publication of this Notice in the Federal Register, agencies will be able to utilize a new uniform format for reporting performance progress on Federally-funded research projects. The Research Performance Progress Report (RPPR) will directly benefit award recipients by making it easier for them to administer Federal grant and cooperative agreement programs through standardization of the types of information required in interim performance reports—thereby reducing their administrative effort and costs. The RPPR will also make it easier to compare the outputs, outcomes, etc. of agencies or program-specific component, if necessary, to meet programmatic requirements, although agencies should minimize the degree to which they supplement the standard components. Such agency- or program-specific requirements will require review and clearance by OMB.

Agencies also may use other OMB-approved reporting formats, such as the Performance Progress Report (PPR), if those formats are better suited to the agency’s reporting requirements, for example, for research centers/institutes, clinical trials, or fellowship/training awards or in connection to reporting on program performance, through mechanisms such as the Performance Assessment Rating Tool.

On behalf of the RRM Subcommittee, the National Science Foundation (NSF) has agreed to serve as sponsor of this new format. We anticipate this being the final notice before the format and instructions are finalized. The general public and Federal agencies, however, are invited to comment on the proposed final format during the 30 day public comment period. The Government-wide RPPR is posted on the NSF Web site at www NSF.gov.

FOR FURTHER INFORMATION CONTACT: For information on the RPPR, contact Joan Feldman: Head, Policy Office, Division of Institutional & Support, National Science Foundation; 4201 Wilson Blvd; Arlington, VA 22230; e-mail: jfeldman@nsf.gov; telephone: (703) 292-8242; fax: (703) 292-0171.

For further information on the NSTC RRM Subcommittee, contact Diane DiEuliis, at the Office of Science and Technology Policy, 725 17th Street, NW, Washington, DC 20503; e-mail: ddieuliis@ostp.gov; telephone: 202-

Reducing Burden:
The Brazilian Experience

Ministério da Ciência e Tecnologia

Lattes Database is a Curriculum and institutions database of Science and Technology areas in Brazil.

More info

SCOPUS
ISI WEB OF KNOWLEDGE
Research and Students

Researchers and Institutions

Contact
Other Database

Copyright CNPq 2009.
“Facebook for Scientists”

- Information in VIVO can be used to create
  - Biosketches
  - Vitas
  - Annual reports
  - Department and research group web sites
- Information can be used to populate profiles in collaborative tools – portals, wikis, …

Leverage Existing Data:
- e.g. Developing Patent Database

Institutional Support:
- NSF SciSIP: 0830287, 0965259
- HBS: Department of Research
1. Knowledge Diffusion
Three links out
(Singh 2005)

2. Sources of Links
- Student graduation
- Inventor mobility
- Direct collaboration
(Fleming 2007)
Capturing Outcomes

Scientists create tags on their websites, collaborate through VIVO, or register through a LATTES-like process.

New approaches discussed and validated with FDP.

STAR METRICS
1. Inhales information from scientists
2. Creates Progress Report for scientists to validate
3. Exhales information to agency reports

Agencies identify fields that can be inhaled from STAR METRICS.

Practical Application

Accelerating Innovation Research (AIR)

PROGRAM SOLICITATION
NSF 10-608

National Science Foundation
Directorate for Engineering
Industrial Innovation and Partnerships

Letter of Intent Due Date(s) (required) (due by 5 p.m. proposer’s local time):
December 01, 2010

Full Proposal Deadline(s) (due by 5 p.m. proposer’s local time):
February 01, 2011
Current Status

- NIH, NSF and OSTP MOU signed, DOE and EPA joining
- Partnership with Federal Demonstration Partnership, and engagement with AAU, APLU, COGR
- Over 100 academic institutions at various degrees of participation
- European Union engagement and emulation
What does this entail?

- Partner with PIs to
  - develop flow-based annual and final reports/biosketches
    - [http://ideas.repec.org/e/pla36.html](http://ideas.repec.org/e/pla36.html)
    - [http://citeseerx.ist.psu.edu/](http://citeseerx.ist.psu.edu/)
  - Visualizations of networks and impact
  - Collaborative tagging of research outputs etc.…
- Partner with university administrators to develop flow-based impact of science funding
Ultimate Goals for Development of Science Metrics

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

Why metrics matter

- You can`t manage what you can`t measure
- And what you measure is what you get