

NCI-Viz:

Developing an agile tool for monitoring and visualizing funding outputs and understanding research portfolios

Duane Williams

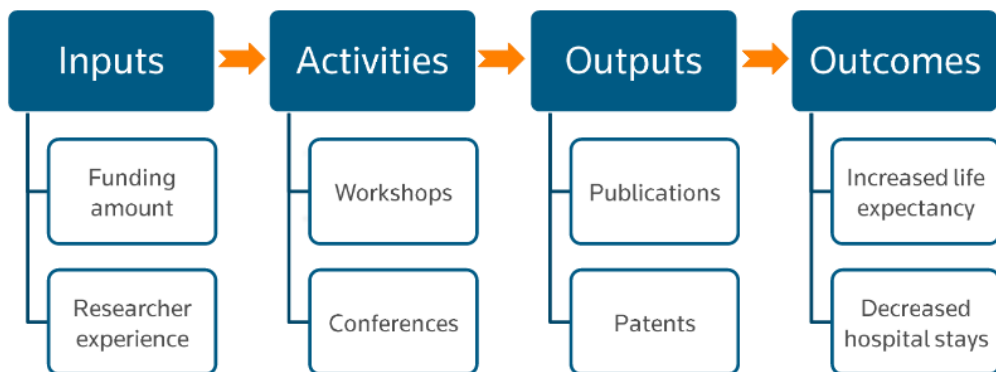
Thomson Reuters

Elizabeth Hsu, Danielle Dae, Larry Solomon, James Corrigan (NIH)

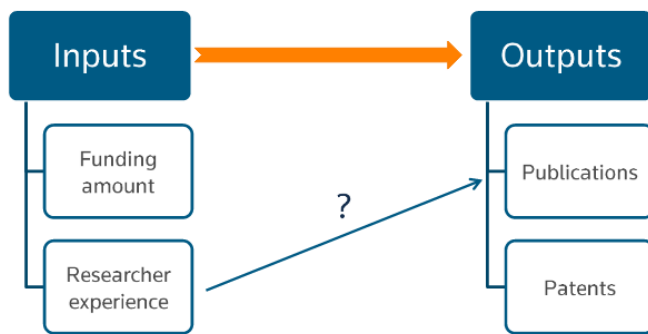
Joshua Schnell (Thomson Reuters)

Motivation for System Development

Logic model for common program evaluations:



Prior analysis on NCI grant portfolio to identify grant characteristics associated with higher productivity?



Motivation for System Development: Identifying Variables

Inputs

Topic Group	Input Variables (truncated)
Project characteristic	Funding Mechanism
	Average annual funding
	Priority score
	Number of Investigators
	Number of competitive renewal applications
Institution characteristic	Institution classification
	US state
	Medical school involvement
Personnel prior achievement	Average prior NIH awards per person
	Average prior publication count
	Average impact factor of prior publications
Personnel demographics	Average age of investigator at project start
	Investigator degree classification

Data Source: NIH Grants Database (IMPAC II)

Outputs (Bibliometrics+)

Topic group	Output Variables (truncated)
Productivity	Publications per dollar
	Publication count in the first 5 years per dollar
Publication quality	Average times cited (excluding self)
Journal quality	Average impact factor
Timing and trends	Time to first publication
	Average citation velocity
	Average cited half life
Enhanced return	Has patents
	Has publications in Science, Nature, or Cell
	Has publications cited by NCCN guidelines or ASCO advances

Data Source: Web of Science, MEDLINE, US Patent and Trade Office

Motivation for System Development: Findings from Prior Study

Return on Investment study (presented at AEA in 2013*)

Large-scale statistical analysis on NCI grant portfolio results

Somewhat Expected

- **Average annual funding:** more funding associated with higher return

Somewhat Unexpected

- **Institution Type:** Grants to medical schools produce fewer publications, but in higher ranked journals compared to research institutions. (production down ~10%, quality up ~10%)
- **Funding Mechanism:** even after accounting for all other factors within the input variable list (measures accounting for several aspects of the grant, institution, prior achievement, and demographics), some mechanisms are still associated with higher return per dollar.

Would the data aggregation and/or findings be useful for program staff?

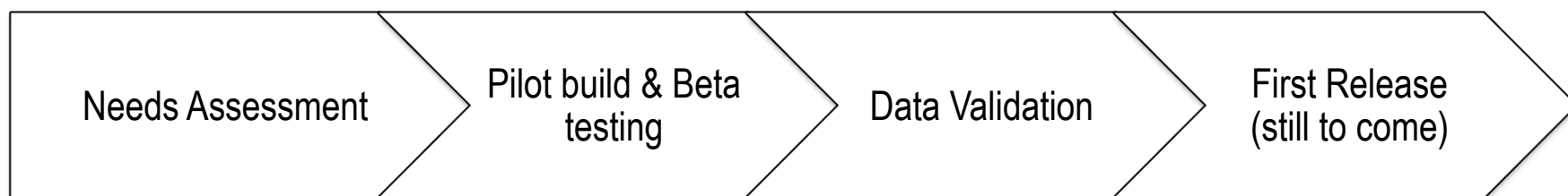
National Cancer Institute

*Data Modeling to Support Portfolio Analysis and Productivity Assessment

American Evaluation Association: Evaluation 2013, Evaluation Practice in the Early 21st Century, Washington, DC

Moving Findings off of the Shelf

Would the data aggregation and/or findings be useful for program staff?



Needs Assessment: Process

Goals

- Assess value of NCI providing results through a visualization tool
- Solicit suggestions for analyses and visualizations

Participants

- NCI program directors

Pilot Interviews

- To determine potential level of interest and usefulness of tool before pursuing focus groups

Focus Groups – divided into three portfolio topic areas

- Basic science research portfolio (Current visualization tool primarily focused on addressing needs identified in this area)
- Training portfolio
- Technology/tool development

Needs Assessment: Question Sample

Information Needs

- For what types of external information requests do you need to provide portfolio data?
- How were these data obtained, and how difficult was the process?
- What data have you been asked to produce on your portfolio over the last year?

Specific Prompts

- Would it be useful to have reporting/visualization layer that would represent NIH-collected administrative data in a variety of ways?
- What questions do you have that integrated bibliometric information might answer?
- What would be useful ways to represent/report on those data?

Day-to-Day Management

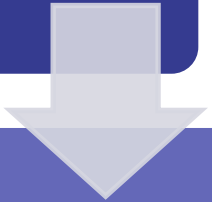
- What are the most important data elements?
- What tools do you have (NCI, NIH, or others) for pulling/displaying the data (e.g., QVR, NIHMaps, etc)?
- What information would you like to have at your fingertips but do not have readily available?

Needs Assessment: Sample Findings

Question	Response
Frequent external requests for portfolio data	“How much research does NCI/branch fund in a given area?”
Useful of additional tool for day-to-day grants management	Potentially very useful if it works effectively and mimics current workflow.
Current process for obtaining data	Varies significantly by individual and group.
Currently used NCI /NIH tools	Several tools used, but post-processing often needed. Variation in comfort levels in using available tools.

Grant-focused Philosophy

Pre-process Metrics for Grants and
Associated Inputs/Outputs
(People, Publications, Patents)



Web Interface for Querying Grant Text and
Attributes (Titles, Abstracts, Specific Aims)



Analyze by Grant and Other Properties in
Visualization Tool

Pilot Build: Components of the Tool

Structured Data & Statistics

Relevance-Overall	Project Number	Title	Abstract	Program Director	Funding Mechanism	Publication Count	Publications per \$M	Duration	Project Cost	Institution	WebLink
1	R01CA000000	C-MET 1	C-MET 1	PD #1	R01	19	5.264573934	11	\$4M	University #1	http://projectreporter.nih.gov/project
5	R21CA111111	C-MET 2	C-MET 2	PD #2	R21	0	0	3	\$370k	University #2	http://projectreporter.nih.gov/project
8	R01CA222222	C-MET 3	C-MET 3	PD #3	R01	2	1.563968664	5	\$1M	University #3	http://projectreporter.nih.gov/project
1	R01CA333333	C-MET 4	C-MET 4	PD #4	R01	7	5.871109027	5	\$1M	University #4	http://projectreporter.nih.gov/project
34	R15CA444444	C-MET 5	C-MET 5	PD #5	R15	0	0	3	\$421k	University #5	http://projectreporter.nih.gov/project
2	R21CA555555	C-MET 6	C-MET 6	PD #6	R21	0	0	3	\$507k	University #6	http://projectreporter.nih.gov/project

Slim web-interface for querying grants

SEARCH OPTIONS

Keywords To Include:

Keywords To Exclude:

From: To:

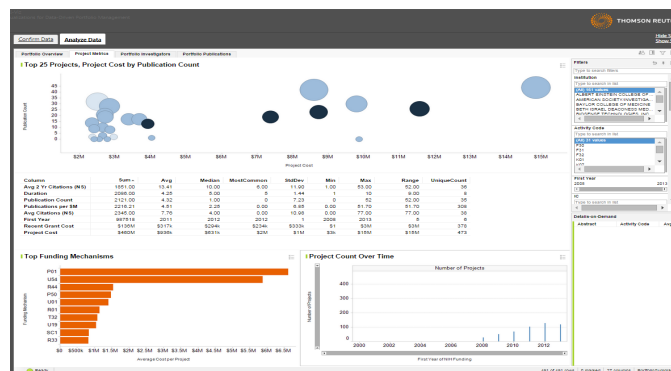
Program Officer:

Funding Mechanism:

Division:

IC:

Third Party Visualization Tool



Needs Assessment: Beta Testers

- In-depth hands-on testing with three program directors
- Feedback on interface, information and metrics, and visualizations
 - Refinement based on feedback
- Planned: in-depth data quality check

Web Interface for Querying Grants & Confirming Relevance

How many active grants related to the protein “c-Met” are funded by NCI’s Divisions of Cancer Biology (DCB) and Cancer Treatment and Diagnosis (DCTD)?

The screenshot shows a web interface titled "SEARCH OPTIONS". It contains several input fields and dropdown menus for filtering search results. The "Keywords To Include" field contains the text "c-met" or "cmet". The "Keywords To Exclude" field is empty. The "From" field is set to 1990 and the "To" field is set to 2013. There are four main filter categories: "Program Officer", "Funding Mechanism", "Division", and "IC". The "Funding Mechanism" dropdown is currently set to "ALL". The "Division" dropdown is currently set to "DCB" and "DCTD". The "IC" dropdown is currently set to "NCI". There are "Clear" and "Search" buttons at the bottom right of the form.

Keywords To Include:	Keywords To Exclude:	From:	To:	Program Officer	Funding Mechanism	Division	IC
"c-met" or "cmet"		1990	2013		ALL	DCB DCTD	NCI

Clear Search

- Boolean searching
- Program directors can search by their names
- Restrict searches by funding mechanism or division

Web Interface for Querying Grants & Confirming Relevance

How many active grants related to the protein “c-Met” are funded by NCI’s Divisions of Cancer Biology (DCB) and Cancer Treatment and Diagnosis (DCTD)?

Confirm results, then toggle to analysis

Remove search results that are not of interest

Confirm Data Analyze Data

29 results found

Check Text Remove Checked

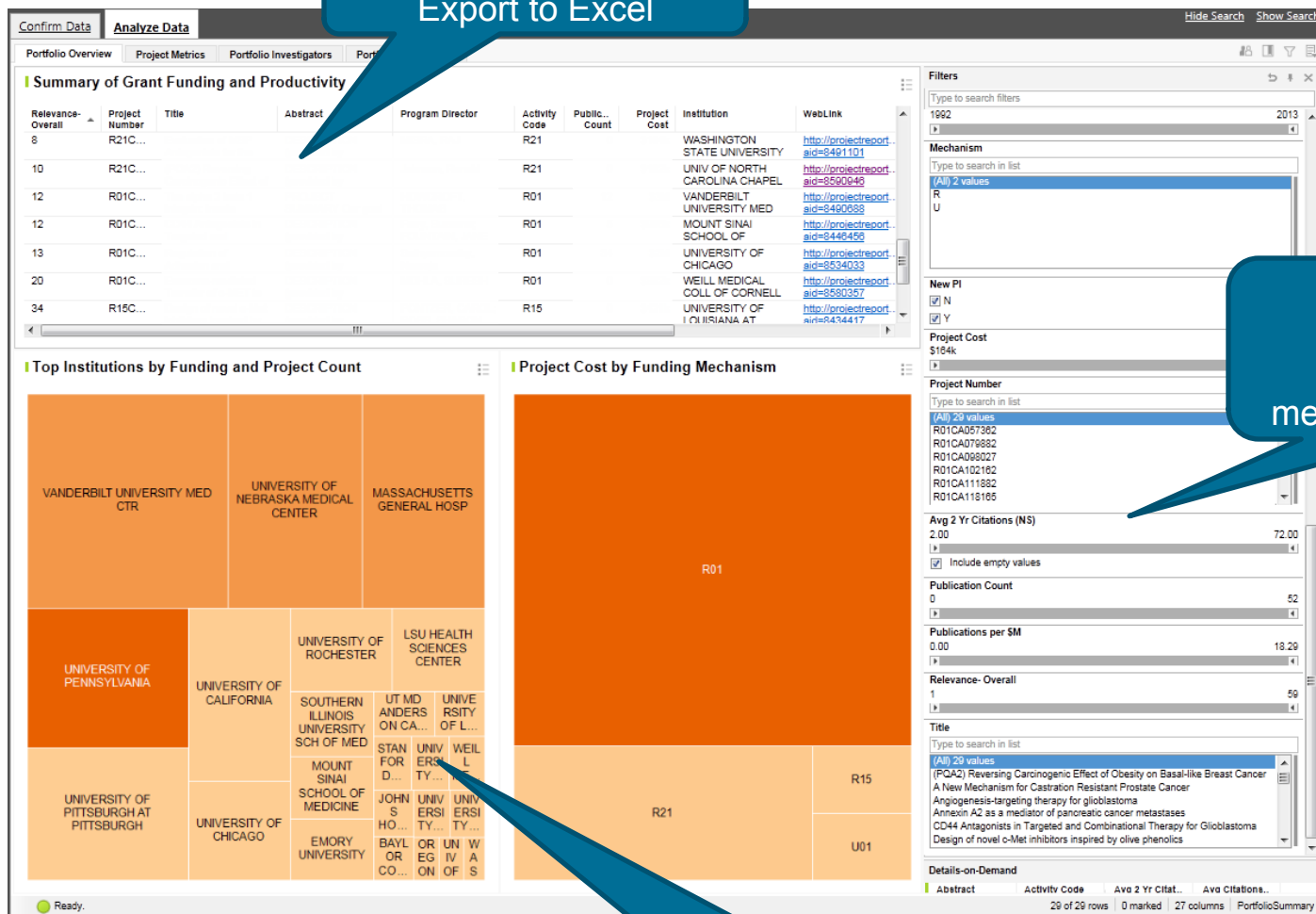
	Relevance	Project Number	Investigator	Title	Program Director	First Year	IC
<input checked="" type="checkbox"/>	59	R21CA158965		Synergism of c-Met and EGFR pathways: their therapeutic role in lung cancer		2012	CA
<input type="checkbox"/>	43	U01CA166894		A New Mechanism for Castration Resistant Prostate Cancer		2013	CA
<input type="checkbox"/>	36	R01CA079882		Targeting the HGF/c-Met Pathway in Lung Cancer		1999	CA

DESCRIPTION (provided by applicant): c-Met and Epidermal Growth Factor Receptor (EGFR) are receptor tyrosine kinases (RTKs) that work together to promote the progression of Non-Small Cell Lung Cancer (NSCLC). Our current studies focus on the use of VeraTag technology to quantitatively, accurately and objectively determine if co-activation, co-expression and proximity of c-Met and EGFR can predict patient prognosis. Research over the last

Preview abstract

Portfolio Summary

Project Details for
Export to Excel



Filters by project
attributes
(e.g., funding
mechanism, total cost)

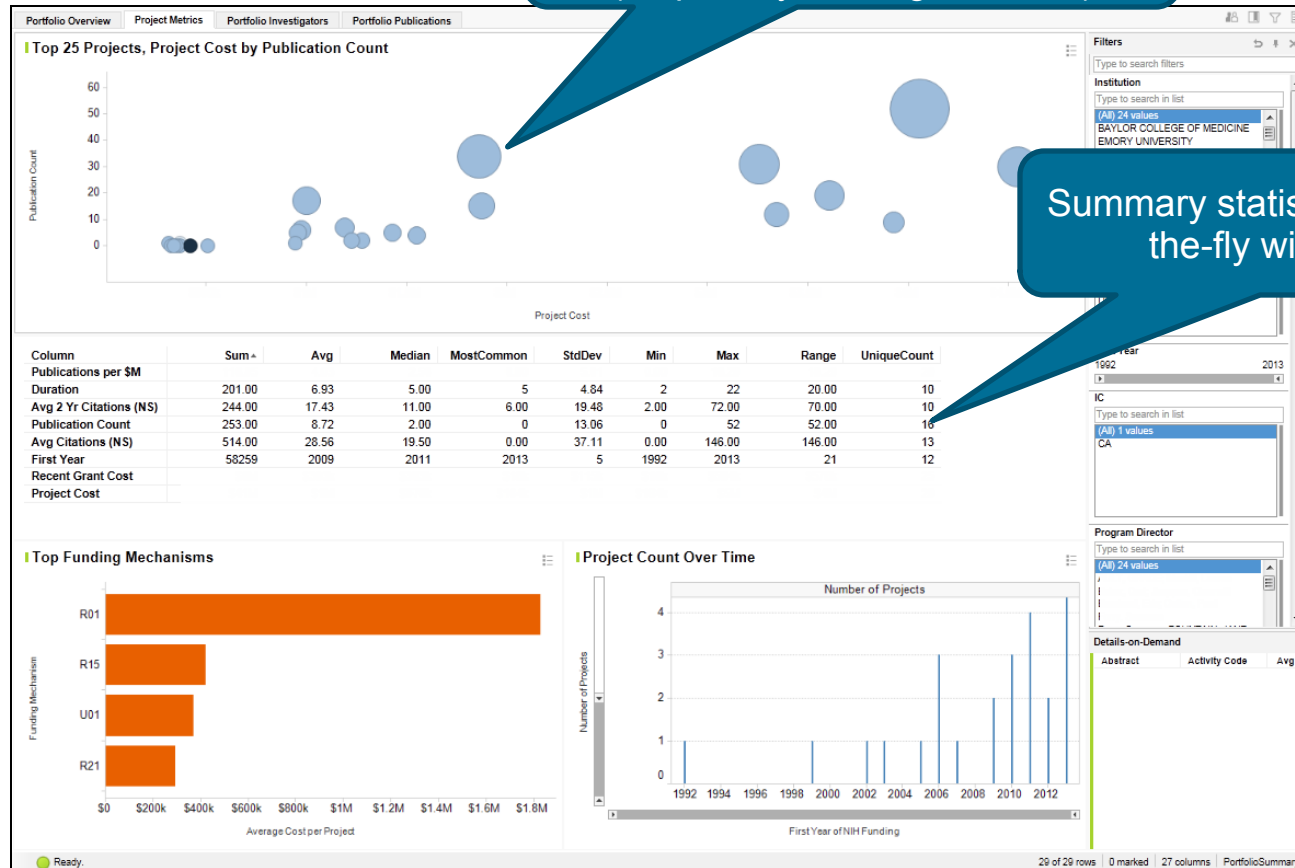
Overview of Institutions
and Funding Mechanisms

National Cancer Institute

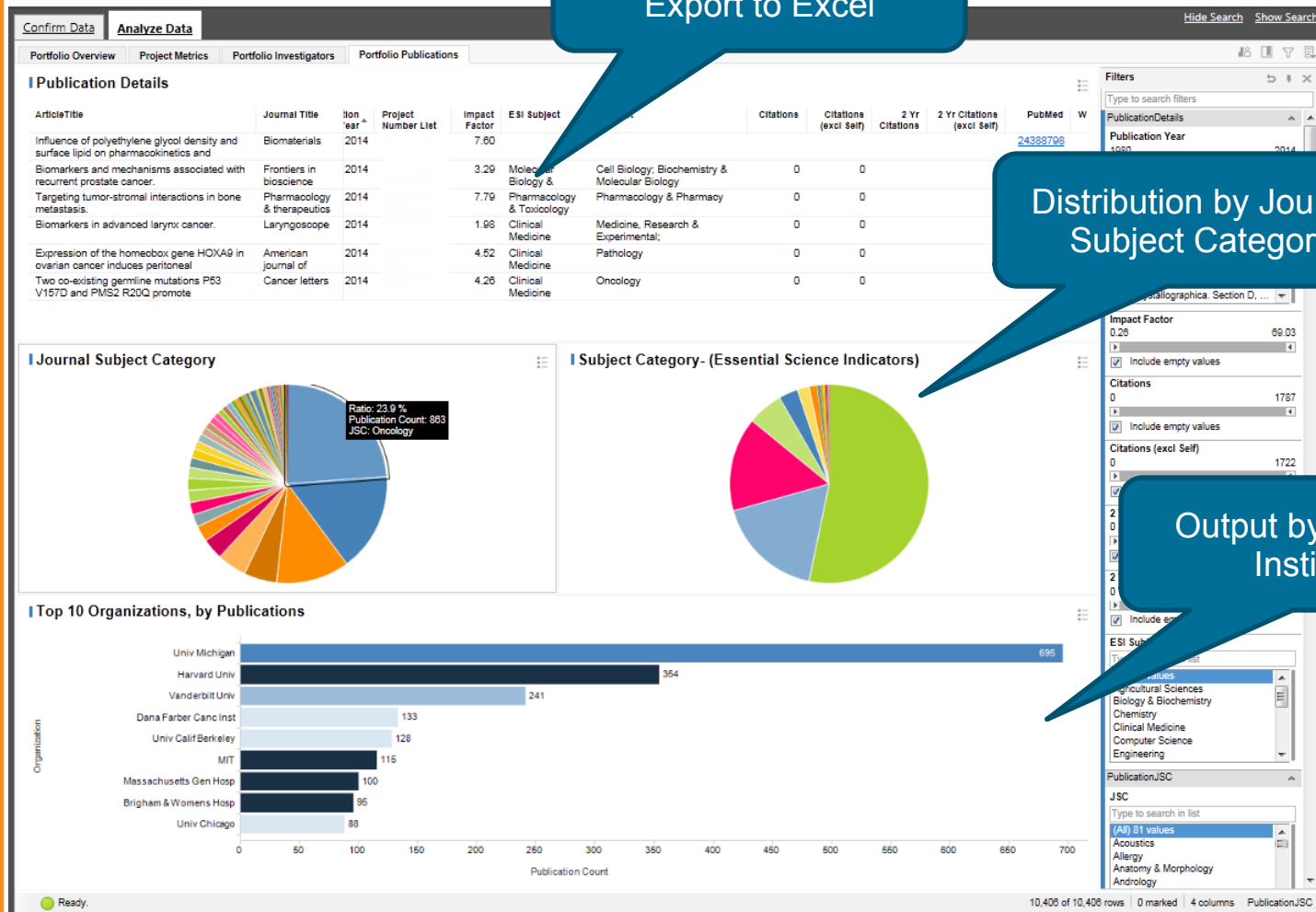
Analyze Productivity and Impact of c-Met Related Grants: Which projects are most productive?

Easy to interpret views of grant
productivity
(outputs by funding amounts)

Summary statistics updated on-
the-fly with filtering.



Understand the Nature and Quantity Publications Supported by Grants within the Portfolio



Acknowledgments

National Cancer Institute

1. Elizabeth Hsu
2. Danielle Daee
3. Larry Solomon
4. Karl Poonai
5. Anne Menkens
6. Christine Siemon
7. Sherri Schully
8. Jim Corrigan

IDA Science and Technology Policy Institute

1. Brian Zuckerman
2. Judith Hautala
3. Sherrica Holloman

Thomson Reuters

1. Ronan Sorensen
2. Josh Schnell
3. Matthew Fouch
4. Holly Wolcott
5. Pawel Sulima
6. Vincent Huang
7. Duane Williams



Questions?