



# Using a Business Analytics Tool to Visualize NIAID Extramural Data

Patricia Haggerty, PhD<sup>1</sup>; Kevin Footer<sup>2</sup>; Rachel Belyavski<sup>2</sup>; Aarti Gupta<sup>2</sup>; Teja Dogiparthi<sup>2</sup>; Sarah Qureshi<sup>2</sup>; Roy Wright<sup>2</sup>; and Stephen Xu<sup>2</sup>  
<sup>1</sup>Division of Extramural Activities Office of the Director and <sup>2</sup>Office of Cyber Infrastructure and Computational Biology Software Engineering Branch, National Institute of Allergy and Infectious Diseases

## INTRODUCTION

At NIAID, the Division of Extramural Activities (DEA) needs a comprehensive understanding of the Institute's extramural research grant and contract portfolio. With a very large database of extramural data, traditional query, spreadsheet analysis and charting tools require significant manual data conditioning and manipulation. This process is time consuming and prone to human error. DEA wanted a tool that could vastly improve the efficiency, accuracy and speed of analysis and visualization of e-stored data.

The *Extramural Data Dashboard (EDD)* is a data analysis and visualization tool to help translate e-stored extramural data into meaningful information to inform strategic planning, resource allocation and decision making. The tool is being developed in stages. The current stage focuses on NIAID competing grant application and award data. Planned future stages of the EDD include: non-competing data, PI-based data, comparative NIH IC-data, and contract proposal data. The poster shows examples of data visualizations and real analyses we have done using the EDD.

## OBJECTIVES

DEA wanted a comprehensive portfolio analysis and visualization tool to:

- assess current status, patterns and trends in the NIAID extramural research grant and contract portfolio;
- provide an accurate, fast and efficient way to visualize data extracted from IMPAC II, and to automate the data conditioning, summarization and charting we had been doing manually before;
- assess current and historical submission and award data, and analyze the portfolio at multiple levels of programmatic and budgetary interest; and
- incorporate future modules focused on PIs, contract proposals, and IC-comparative data.



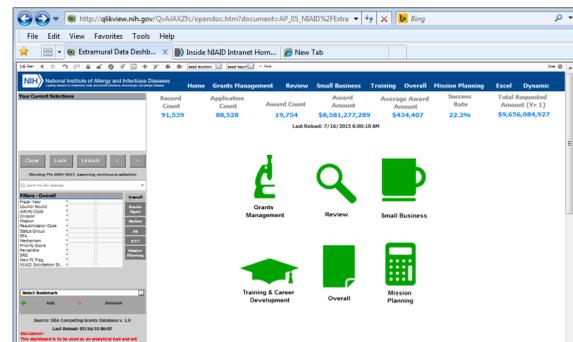
## METHODS

DEA and NIAID's Office of Cyber Infrastructure and Computational Biology developed the EDD using: 1) the QlikView platform for the dashboard display, 2) SQL Server @ 2012 to store the database, and 3) IMPAC II to retrieve the stored data.

- We began with a focus on NIAID-funded competing grant application and award data (Phase 1). Data from FY03 through the current FY were included.
- OER frozen reports were used to validate all historical (frozen) years.
- QVR was used to validate data from the current FY.

## PRODUCT DESCRIPTION

### EDD Landing Page



The EDD is divided into:

- Site: A Site can contain multiple projects and is the master domain
- Project: Projects act as folders to hold multiple workbooks
- Workbooks: Workbooks are similar to Excel files that may contain multiple tabs called views
- Views: Views are individual pages that contain the interactive dashboards that collectively share the story of the data



The Data Details:

- FY03-FY16:
  - ~92,000 records
  - 55 dimensions
  - ~ 5,000,000 cells of programmatic information
- 6 key metrics:
  - Record Count
  - Application Count
  - Award Count
  - Award Total Amount
  - Average Award Amount
  - Success Rate
- Data from IMPAC II are updated automatically on a daily basis.

## RESULTS

### Screen Shots from the EDD

- Visualization of FY03-FY10 R01-equivalent data



- Visualization of New PI Counts FY09-15



- New PI Counts FY09-15, Spreadsheet view

- EDD Data Dictionary

## LESSONS LEARNED

Business Analytics (BA) tools can be powerful resources for the NIH extramural community. At NIAID, the EDD is harnessing business analytics to interpret large volumes of e-stored data on research funding applications and awards. In the course of developing our Dashboard, we learned many important lessons. We share a few of these below:

- BA tools can help turn "mega data" into meaningful information to support decision making and planning.
- BA tools can help save an organization *time and money*. Analyses that would take hours, weeks or longer with traditional search, analysis, and charting tools, can be done instead in minutes or hours. Duplication of analyses required with traditional tools to reconcile data and reduce error are greatly decreased, or eliminated.
- A learning curve for developers and business users is to be expected in the NIH environment. The larger and more complex the source databases are, the more time and effort it takes to develop a tool that is accurate, efficient, and relatively easy for users to learn.
- We anticipate that BA tools are here to stay. One day, even at NIH, they will be ubiquitous. Given the power and sophistication of such tools, however, it is hard to imagine how radically they may change the way the NIH does business in the future.



For additional information please contact either:

Patricia Haggerty, PhD  
 Division of Extramural Activities  
 NIAID  
[haggertp@nih.gov](mailto:haggertp@nih.gov)

or

Software Engineering Branch, OCICB  
 NIAID  
[NIAIDBusinessIntelligence@niaid.nih.gov](mailto:NIAIDBusinessIntelligence@niaid.nih.gov)