Data Analytics for Faculty Success and Career Development

Dr. Alyson Grace Eggleston, Pennsylvania State University

Alyson Eggleston is an Associate Professor in the Penn State Hershey College of Medicine and Director of Evaluation for the Penn State Clinical and Translational Science Institute. Her research and teaching background focus on program assessment, STEM technical communication, industry-informed curricula, and educational outcomes veteran and active duty students.

Dr. Robert J. Rabb P.E., The Pennsylvania State University

Robert Rabb is the associate dean for education in the College of Engineering at Penn State. He previously served as a professor and the Mechanical Engineering Department Chair at The Citadel. He previously taught mechanical engineering at the United States Military Academy at West Point. He received his B.S. in Mechanical Engineering from the United Military Academy and his M.S. and PhD in Mechanical Engineering from the University of Texas at Austin. His research and teaching interests are in mechatronics, regenerative power, and multidisciplinary engineering.

Eric Donnell, The Pennsylvania State University

Data Analytics for Faculty Success and Career Development

Engineering and STEM education opportunities can often hinge on extra-departmental funding opportunities—institutional research centers and external grant competitions. As engineering programs seek to invest in the next generation of engineers, research administrators can operationalize research effort data to identify (1) near-term undergraduate and graduate student experiential opportunities; (2) top-performing teacher-scholars poised to lead student experiences; (3) features of teacher-scholars that can be predictive of early-stage interventions that support their success as fundable grantees. Data visualizations in service to engineering and STEM programs provide a high-context field of opportunity for administrators, faculty, and students, supporting the continued growth of the engineering career pathway. More than just pie-charts, intentional Power BI dashboards can provide a map of the engineering opportunity 'terrain' as well as identify the most impactful areas to intervene.

This paper reports on a research institute's multidimensional regression analysis at Penn State University that was used to support STEM-focused research outcomes and support continued mentoring and development for faculty. Impacting downstream student experiential learning opportunities as well as faculty development, we present this approach as a turnkey operation that is easy to implement in other institutional contexts. The developed research data architecture includes selected variables: scholar h-index; professional rank; collaboration team; interdisciplinarity; prior grant performance; project spending; high impact publications; and external funding outcomes. Selected processes for automating scholar data collection are described.

Results from initial work were tailored for implementation within the Colleges of Engineering and Medicine at Penn State University. This paper provides initial results from both case studies and explores the data-driven decision-making process in the context of STEM programs. Challenges and operational bottle-necks to automating the data collection are discussed with possible solutions outlined. The authors recognize the potential conflicts to privacy and preference that can emerge during the dashboard-building phase. As communication becomes increasingly visual, Power BI dashboards that are tailored for decision-makers are featured with an explanation of metric selection, audience selection, and focus—recognizing that unique audiences value different metrics.

Introduction

Horizon challenges in Higher Education (HE) identify data-driven decision making and initiatives to improve faculty development as important future 'rupture points' where both opportunity and challenge coincide [1]. HE Institutions are increasingly turning to Business Intelligence (BI) tools to capture Key Performance Indicators (KPIs), understand student, faculty, and staff populations; inform leadership; comply with funding and accreditation mandates; and predict future trends [2]. With a growing need for transparency, accountability to multiple stakeholders, and a value shift toward data-driven decision making, Business Intelligence (BI) Dashboards provide the information needed to ensure continuous improvement in service offerings and efficient processes [3].

A comprehensive seed funding and external funding data dashboard is essential for gaining deep, actionable insights into individual researchers' performance, enabling institutions to track funding trajectories, identify high-impact contributors, and strategically allocate resources. To ensure the dashboard is both analytically rigorous and practically valuable, these projects are led by a trained program evaluator and social scientist skilled in managing large datasets and translating complex data into meaningful insights for technical and strategic decision-makers.

The Role of Pilot Funding

Pilot funding is an institutional tool that addresses several concurrent challenges and priorities along four dimensions: research and innovation; faculty engagement; faculty development; and economic value capture. Operationally, good ideas begin as pilot grants where prototype testing and key collaborations can be instigated (research innovation). The creation of research teams (faculty engagement) and subsequent triaging of institutional supports for those teams serves to drive faculty professional development, upskilling new and mid-career faculty in novel competencies (faculty development). Finally, all institutional pilot programs conduct some kind of Return On Investment (ROI) analysis, providing key figures for institutional leadership to demonstrate the effectiveness of the programs and quantifying the economic value capture of Primary Investigators (PIs) who secure external grant support, yielding a 30-50% revenue stream for the institution. Pilot funding serves as the seed funding to ensure external grant capture is possible.

Retaining Faculty

Pilot funding serves as a faculty retention [5]-[6] tool because it creates interdisciplinary teams, builds industry-academia collaborations, and provides novel opportunities to upskill faculty. Participation in pilot funding programs has been found to correlate with higher faculty satisfaction scores [7]-[8], as faculty report perceiving that the institution cares about their research, professional success, and supports their collaboration with other interdisciplinary teams, fostering a culture of innovation.

STEM Pathway Support

Both National Science Foundation and National Institutes of Health, as federal agencies that support national research infrastructure, have published priorities in supporting individuals throughout the STEM career pathway. An ancillary effect of external grant capture is the opportunity for undergraduate and graduate students to train with PIs while executing cutting edge research, resulting in continual training of the next generation of scientists and engineers. STEM Pathway support also has shown to be a useful recruitment and retention tool for individuals who are members of underrepresented minority groups and women [5]-[6]. Given that most external grants begin with institutional pilot seed funding, we can look at the presence of pilot funding opportunities as ‘first causes’ for these second- and third-order effects: faculty development; student research experiences.

Leveraging Data Visualizations and Dashboards for Decision-Making

Constructing a historically accurate dataset of prior pilot applicants, awardees, and project impacts, the evaluation team at Penn State University Clinical and Translational Science Institute created a dashboard that provided novel insights along temporal, fiscal, and demographic dimensions.

Organized by fiscal year of opportunity, initial efforts focused on identifying which Requests for Proposals (RFP) announcements were most productive in terms of downstream external grant capture and bibliometric impacts. In Figure 1 below, the Big Data RFP yielded $4.5M within 2 years of project completion with a total outlay of $650K.

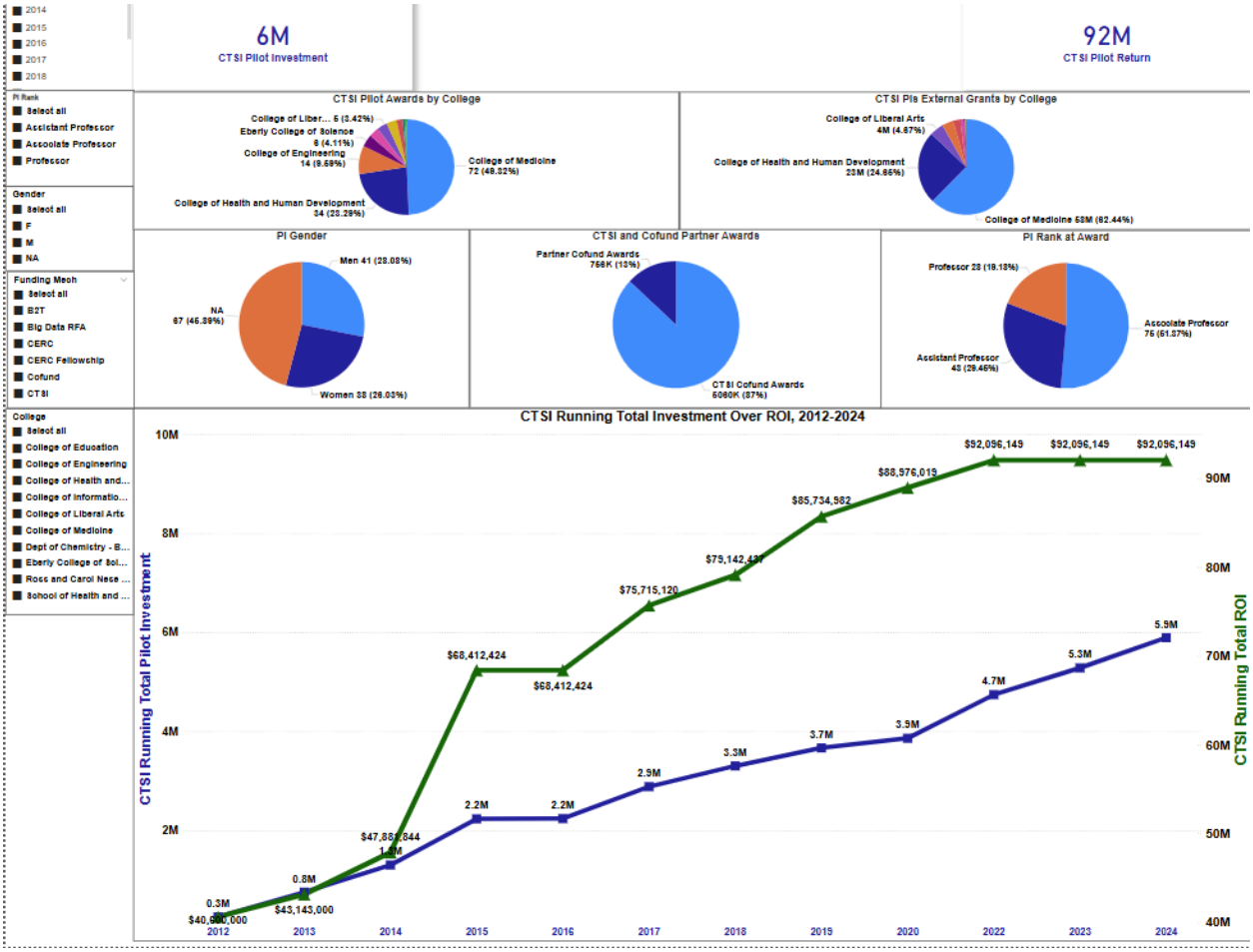

Figure 1: Pilots Dashboard – Historical Data.

Benefits for Key Decision Makers

BI Dashboards provide transparency and real-time insights around pilot funding outcomes. They also highlight useful resource allocations, inviting administrators to ask questions that drive Dashboard continuous improvements. For example, for the above dashboard, the evaluation team demonstrated through dynamic dashboard functionality that full professors provided over 60% of

the ROI capture and pilot awardees were roughly balanced among men and women 47% and 53%, respectively. However, full professors were not overrepresented in the pilot awardees, rather the majority funded were at the rank of associate professor (n=75) 51.4%, with the remainder split between assistant professors (n=43) at 29.5% and full professors (n=28) 19%. This is an important contrast with national trends, as Early Stage Investigators (ESIs) tend to be disadvantaged in capturing NIH funds especially women and researchers of historically underrepresented groups [7].

More than merely demographic information, historical year-over-year trends help guide decision makers as they consider continuous quality improvement strategies. For example, based on ROI alone, full professors seem to outperform lower faculty ranks, however more early-career faculty receive pilot awards than any other category. This finding shows us that pilot seed funding can indeed function as a faculty development tool in an Institution-specific context and we may consider viable co-mentoring intervention strategies between full and assistant professors to close the gap on external funding captures [9]. Co-mentoring models differ from traditional hierarchical mentoring in that they emphasize mutual learning and reciprocal leadership.

Figure 2 below provides the distribution of faculty rank across pilot funding awards from 2015-2024.

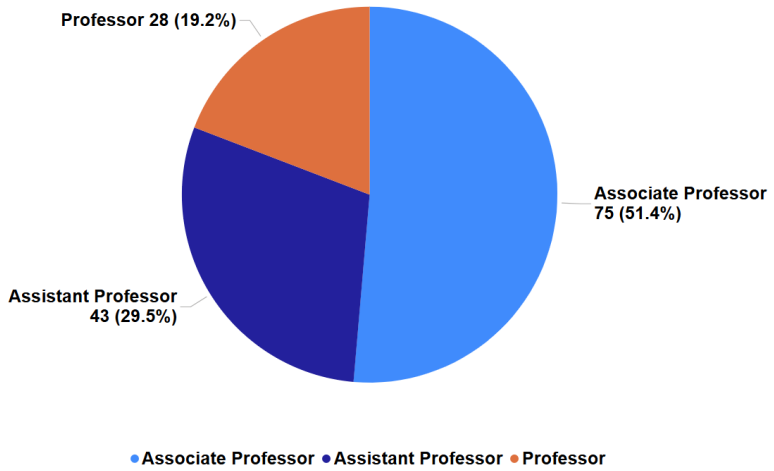

Figure 2: Faculty Rank Distribution Across Pilot Funding Awards.

Identifying top performers in our Research Institute's pilot seed funding program gives us opportunities to better understand how success is achieved in our broader institutional context. Additionally, top performers can be important mentors for early career faculty or first-time pilot fund applicants. In the dashboard below, faculty external funding capture is ranked and all dataset features are dynamically synced such that we can, e.g., drill down to look at women faculty scholars in 2020 who captured more than $1M in external grants.

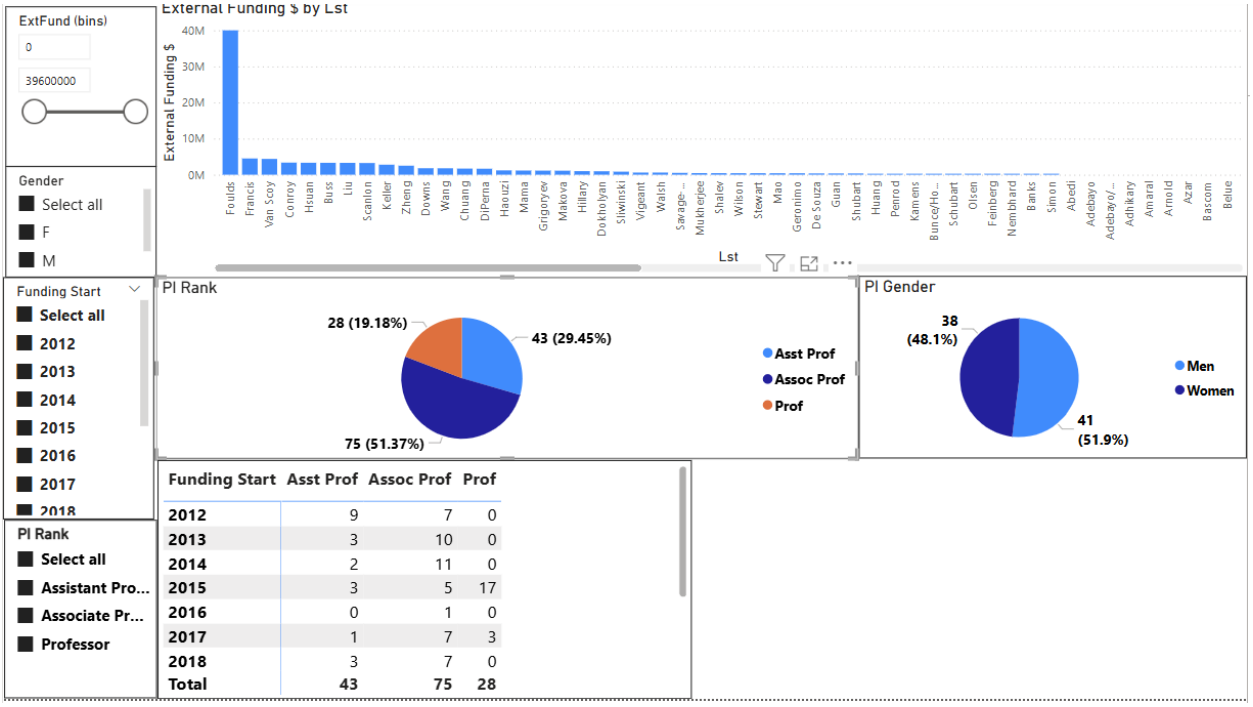

Figure 3: Faculty External Funding Success.

Challenges

BI Dashboards also present unique challenges related to communicating specific narratives to known and unknown audiences; data retrieval bottlenecks; and the need for constant, iterative user-centered design and maintenance. With the call for transparent and connected datasets [1] that conform to international data standards, there is renewed concern for security and privacy. For this reason, BI reporting can be 'fenced off' for privacy, releasing column-by-column information for users with specific access levels. Data creation and retrieval remain real challenges as record quality tends to diminish as the 'look-back' period increases.

Within the Clinical and Translational Research Institute, Standard Operating Procedures (SOPs) have been implemented to (1) increase faculty scholar reporting and (2) integrate cloud data into aggregate scholar impacts. Good relationships are fundamental to acquiring good data. Continuous Quality Improvement (CQI) efforts for the evaluation team have resulted in a faculty-scholar contact model that integrates faculty outreach, onboarding, personalized outreach, and finally regular automated reporting in a HIPAA-compliant platform, REDCap. Figure 2 below shows this phased design.

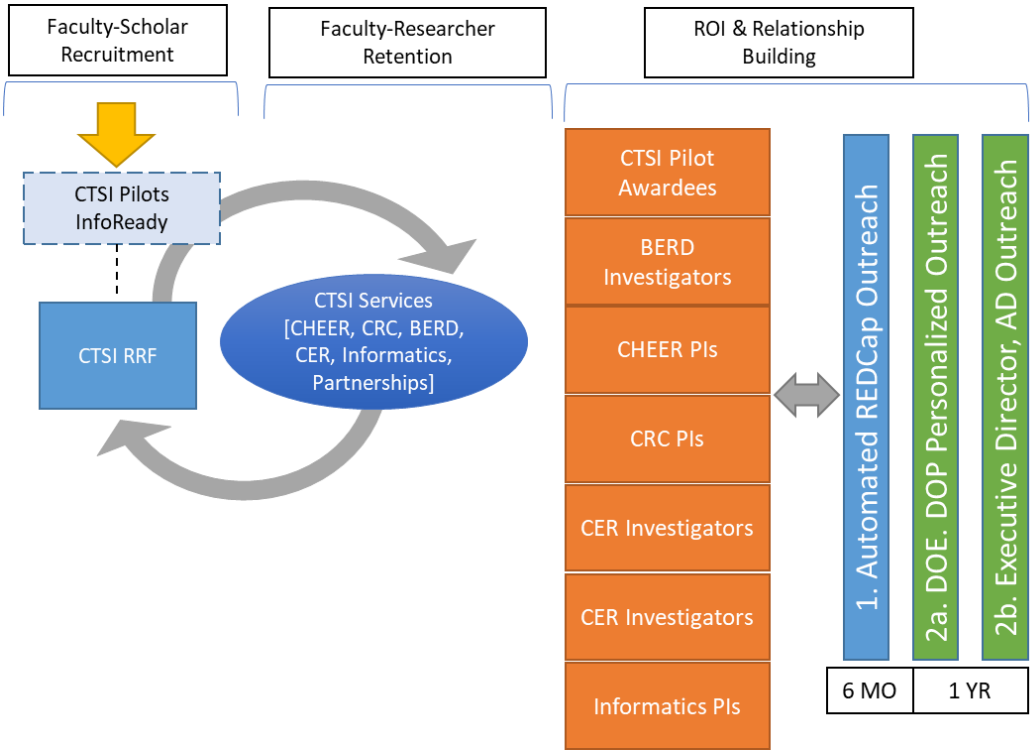

Figure 4: Faculty Recruitment, Retention, and Relationship-building Agents for Better Data Acquisition.

Data retrieval bottlenecks are especially prevalent in the institutional research sphere where faculty bibliometric data and metadata can be sparse and diffuse. To address this, internal reporting protocols must be developed and scaled with supporting staff. Institutions may begin pivoting toward machine-learning assisted data curation platforms like Overton or Dimensions [10]-[11]. Finally, good dashboard development is built by usability studies, user trend data, and -iterative user design (UX), a process that must be continuous to maintain functionality. NSF's internal Deep Insights Anywhere Anytime (DIA2) represents a research funding dashboard that meets these criteria [3].

Moving forward beyond bibliometrics and ROI, our Research Institute is interested in capturing impacts that drive innovation and research activity such as clinical trials, patents, and policy documents. We also want to develop deeper datasets on our pilot awardees, understanding their network of collaborators and where their work is often cited. In figure 5 below, an awardee's citation network is captured in VOSViewer, part of the Dimensions offering. Additionally, there are exportable-to-CSV dashboards that capture clinical trial, patent, and policy impacts as seen in Figure 6.

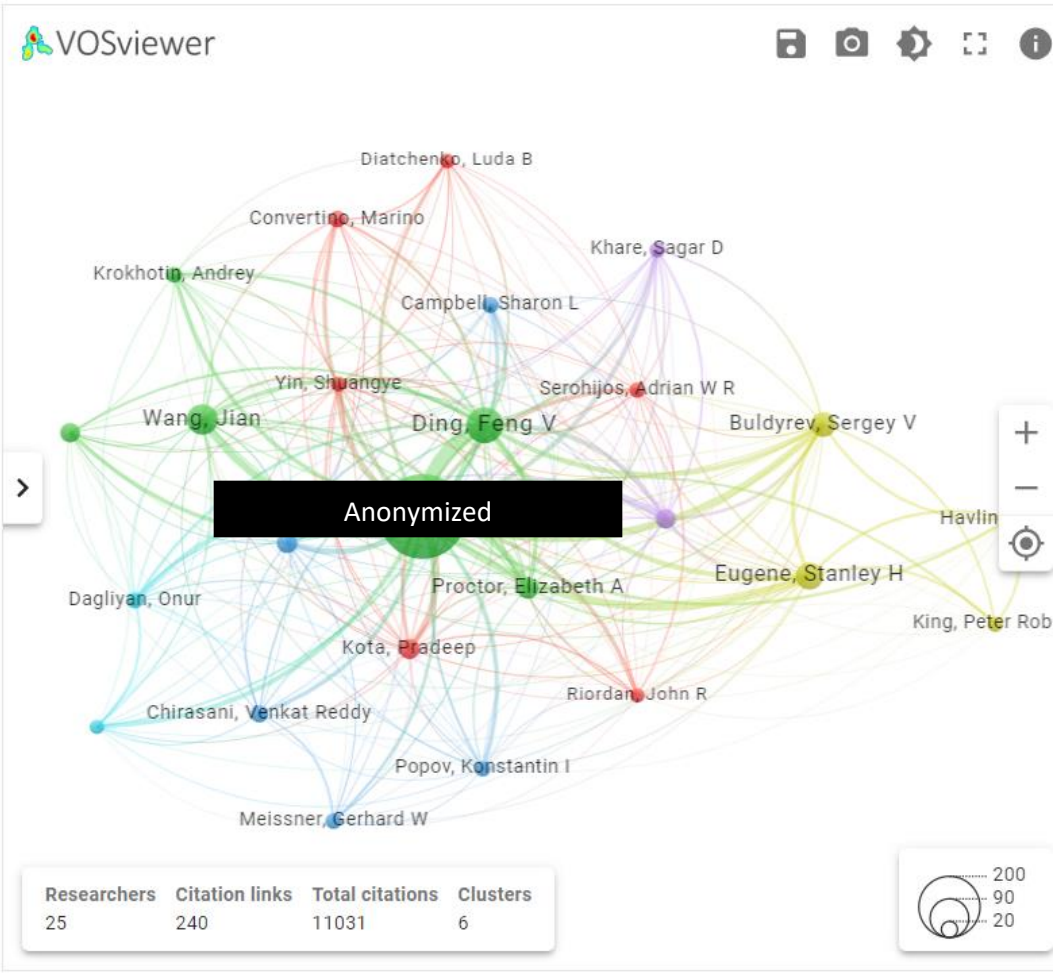

Figure 5: Example Pilot Awardee Citation Network Graph – VOS Viewer.

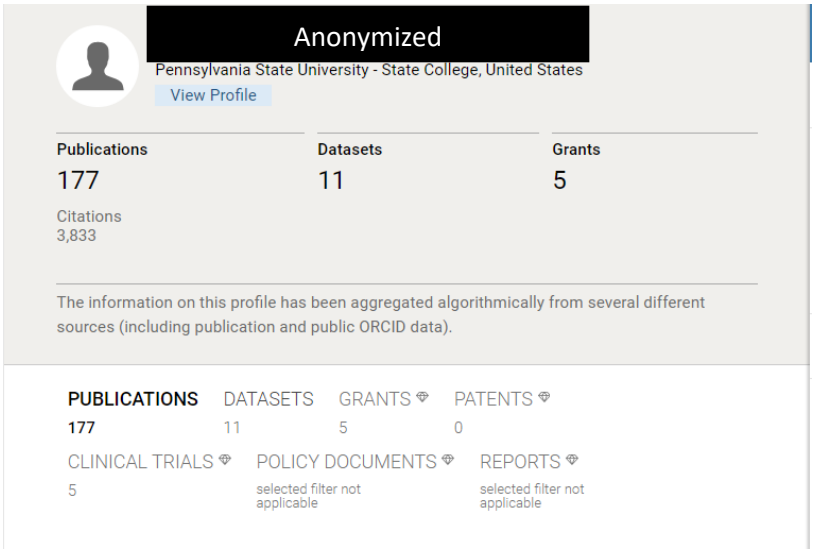

Figure 6: Dimensions Faculty Impact Dashboard.

Long-Term Strategies

BI Dashboards do not yet provide multivariable predictive models of data sets. Evaluation teams can use dashboards to assemble likely models of institutional data and use advanced statistical approaches to confirm data relationships. Innovation systems, defined as "networks, ecosystems, common elements" that co-produce or correlate with increased publications within collaborative groups concentrated in a particular domain [2], are a traditional bibliometric analysis. These bibliometric analyses invariably identify interdisciplinary teams as a driver for innovation. To test this, the evaluation team used multivariate logistic regression models for a series of demographic and productivity measures related to pilot awardees (n=135) in order to identify predictors of (1) pilot seed funding and (2) external grant capture. Dimensions modeled include:

1) Repeat applicant status
2) Gender
3) Interdisciplinary team
4) PI h-index

Preliminary findings from a prior publication investigating predictive factors for applicant selection for pilot seed funding and, ultimately, external grant capture [12] show the importance of interdisciplinary teams. Being a repeat applicant offered no advantage to selection. Repeat applicants were not significantly more or less likely to receive funding. Gender was not significant but women applicants had higher log-odds of receiving funding than men. Team status was highly significant with applicants not part of an interdisciplinary team being much less likely to receive funding. An applicant's h-index, modeled as a continuous variable here and assumed to be in normal distribution, does not affect funding odds in this model; however, in subsequent model analyses the evaluators acknowledge that h-index distribution is often not normal and so merits replacement with a long transformation value. Table 1 provides the model summary in [12]. Figure 7 provides a chart of selected variables and odds ratios where the dashed line at OR=1 represents "no effect." Predictors that do not cross 1 are considered statistically significant. Only Team(no) shows a clearly negatively significant effect on the odds of selection for pilot seed funding.

Table 1: Team Interdisciplinarity as Predictor of Seed Funding Selection.

Parameter		DF	Estimate	Odds Ratio	Standard Error	Wald Chi-square	Pr > ChiSq
Intercept		1	3.98	53.70	1.24	10.31	0.001
Repeat	no	1	0.18	1.21	0.61	0.09	0.763
Repeat	yes	0	0		-	-	-
Gender	F	1	0.79	2.22	0.50	2.45	0.117
Gender	M	0	0		-	-	-
Team	no	1	-4.936	0.007	1.077	20.89	**0.0001**
Team	yes	0	-		-	-	-
hindex		1	0.009	1.009	0.12	0.65	0.419

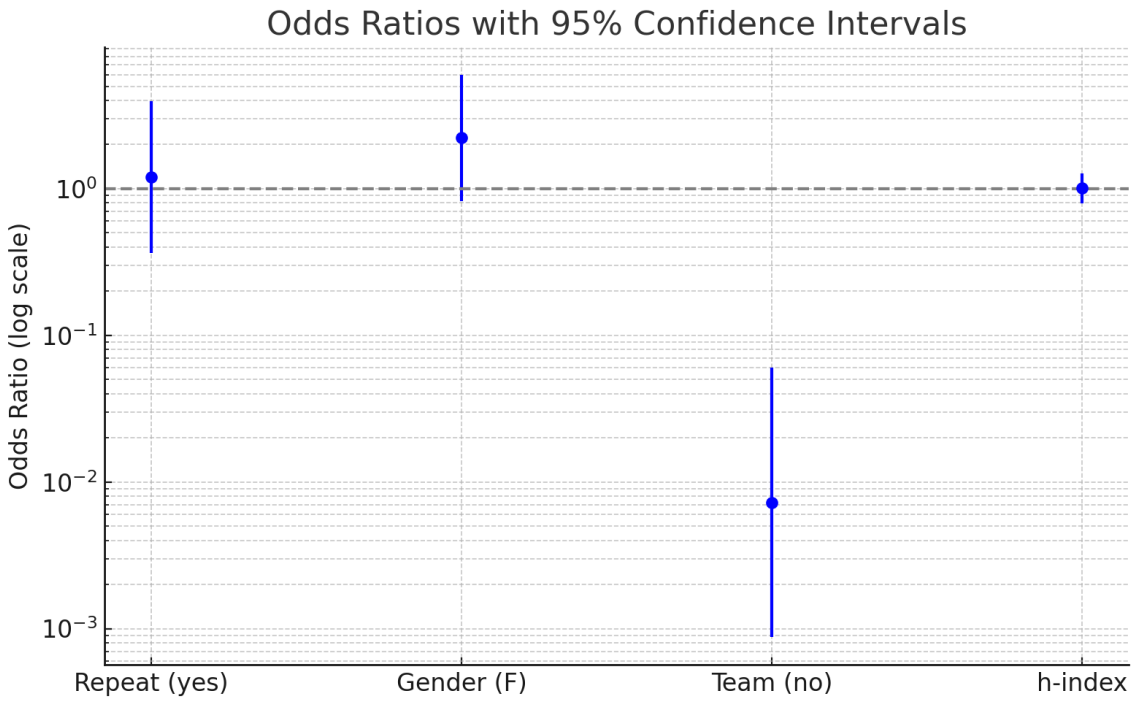

Figure 7: Odds Ratios by Seed Funding Predictor.

Discussion

Our analysis highlights the value of BI dashboards and predictive models in better understanding the pilot seed funding terrain that our faculty must navigate to secure external funding. Building out data infrastructure and BI dashboards create strategic tools for advancing Penn State's research goals and mission, while also providing valuable insights to faculty and their mentors as they plan for continued career development. Moving toward predictive models sets the stage for key insights that are sensitive to an institutional context—in this case, the primacy of interdisciplinary teams for securing initial seed funding.

Moving forward, BI dashboards also allow decision makers to steer pilot funding priorities to closely track with the changing goals of federal funding agencies. By integrating key metrics, such as team strengths and applicant funding histories, a higher resolution footprint of research impact against particular grant funding mechanisms can be established.

Finally, research administration BI dashboards facilitate continuous evaluation processes while arming leaders with the information they need to adapt programs to shifting needs of the institution, funding agency, or national security priorities. With these tools, research administrators are better equipped to steer innovation and maximize the impact of early-stage funding.

References

[1] *2024 EDUCAUSE Horizon Action Plan: Unified Data Models.* https://library.educause.edu/resources/2024/3/2024-educause-horizon-action-plan-unified-data-models

[2] W. Strielkowski, A. Samoilikova, L. Smutka, L. Civín, and S. Lieonov, "Dominant trends in intersectoral research on funding innovation in business companies: A bibliometric analysis approach," in *Journal of Innovation & Knowledge*, 7(4), 100271, 2022.

[3] A. Molnar, A. F. McKenna, Q Liu, M. Vorvoreanu, and K. Madhavan, "Using visualization to derive insights from research funding portfolios," in *IEEE Computer Graphics and Applications*, 35(3), 91-c3, 2021.

[2] L. Shaulska, L. Yurchyshena, and Y. Popovskyi, " Using MS power BI tools in the university management system to deepen the value proposition," in *2021 11th International Conference on Advanced Computer Information Technologies (ACIT)*, pp. 294-298, IEEE, September 2021.

[3] M. D. Tamang, V. K. Shukla, S. Anwar, and R. Punhani, "Improving business intelligence through machine learning algorithms," In *2021 2nd International Conference on Intelligent Engineering and Management (ICIEM)*, pp. 63-68, IEEE, April 2021.

[4] R. Heyard and H. Hottenrott, "The value of research funding for knowledge creation and dissemination: A study of SNSF Research Grants," in *Humanities and Social Sciences Communications*, 8(1), 1-16, 2021.

[5] R. G. Baldwin, C. J. Lunceford, and K. E. Vanderlinden, "Faculty in the Middle Years: Illuminating an Overlooked Phase of Academic Life," in *The Review of Higher Education* 29(1), 97-118, 2005. https://dx.doi.org/10.1353/rhe.2005.0055.

[6] Y. J. Xu, "Gender disparity in STEM disciplines: A study of faculty attrition and turnover intentions," in *Research in Higher Education*, 49(7), 607–624, 2008.

[7] C. C. Collins, J. Dolata, E. Nonguierma, M. Shediac-Rizkallah, A. R. Sehgal, and J. D. Thornton, "Early-Stage Investigators' Experiences with an NIH Pilot Award Program" in *Journal of Clinical and Translational Science*, 1–25, 2025. doi:10.1017/cts.2025.2

[8] K. Morozumi, T. Patel, P. Kerr, M. B. Cassely, T. Carey, J. Buse, G. Dave, "Evaluating a CTSA-funded pilot grant program." *Journal of Clinical and Translational Science*, 5(1), e63, 2021. doi:10.1017/cts.2020.557

[9] J. D. Nelson, M. E. Dunn, and Y. A. Levites Strekalova, "Pilot implementation of a co-mentoring circles program for the clinical research professionals: Evidence for formative evaluation and logic model," in *Journal of Clinical and Translational Science*, 8(1), e21, 2024. doi:10.1017/cts.2023.712

[10] Overton. https://www.overton.io/

[11] Dimensions. https://www.dimensions.ai/

[12] A. Eggleston, and C. Kantor. "Pace and Pitch: Predictive Factors for Seed Funding and Development," in *Journal of Clinical and Translational Science*, 8(s1), 53–54. 2024. doi:10.1017/cts.2024.169