4-19-2016

Data Management Plan Requirements for Campus Grant Competitions: Opportunities for Research Data Services Assessment and Outreach

Andrew M. Johnson
University of Colorado Boulder, andrew.m.johnson@colorado.edu

Shelley Knuth
University of Colorado Boulder, shelley.knuth@colorado.edu

Follow this and additional works at: http://escholarship.umassmed.edu/jeslib

Part of the Scholarly Communication Commons

Recommended Citation

Johnson, Andrew M., and Shelley Knuth. . "Data Management Plan Requirements for Campus Grant Competitions: Opportunities for Research Data Services Assessment and Outreach." *Journal of eScience Librarianship* 5(1): e1089. http://dx.doi.org/10.7191/jeslib.2016.1089

Journal of eScience Librarianship
putting the pieces together: theory and practice

Full-Length Paper

Data Management Plan Requirements for Campus Grant Competitions: Opportunities for Research Data Services Assessment and Outreach

Andrew M. Johnson and Shelley Knuth

University of Colorado Boulder — Boulder, CO, USA

Abstract

Objective: To examine the effects of research data services (RDS) on the quality of data management plans (DMPs) required for a campus-level faculty grant competition, as well as to explore opportunities that the local DMP requirement presented for RDS outreach.

Methods: Nine reviewers each scored a randomly assigned portion of DMPs from 82 competition proposals. Each DMP was scored by three reviewers, and the three scores were averaged together to obtain the final score. Interrater reliability was measured using intraclass correlation. Unpaired t-tests were used to compare mean DMP scores for faculty who utilized RDS services with those who did not. Unpaired t-tests were also used to compare mean DMP scores for proposals that were funded with proposals that were not funded. One-way ANOVA was used to compare mean DMP scores among proposals from six broad disciplinary categories.

Results: Analyses showed that RDS consultations had a statistically significant effect on DMP scores. Differences between DMP scores for funded versus unfunded proposals and among disciplinary categories were not significant. The DMP requirement also provided a number of expected and unexpected outreach opportunities for RDS services.

Conclusions: Requiring DMPs for campus grant competitions can provide important assessment and outreach opportunities for research data services. While these results might not be generalizable to DMP review processes at federal funding agencies, they do suggest the importance, at any level, of developing a shared understanding of what constitutes a high-quality DMP among grant applicants, grant reviewers, and RDS providers.

Introduction

Since 2011 when the National Science Foundation (NSF) first instituted its data management plan (DMP) requirements, many government agencies and other research funders have adopted similar policies requiring applicants to submit DMPs with all grant proposals. In line with this trend, the Office of the Vice Chancellor for Research at the University of Colorado Boulder (CU-Boulder) added a DMP component to the proposal process for all internal campus grant competitions beginning in 2014. This was done in order to raise awareness of the importance of data management in general as well as to better prepare researchers for the growing number of research funder DMP requirements. The CU-Boulder requirement was first implemented for the popular Innovative Seed Grant (ISG) program, an internal grant competition open to all faculty that provides up to $50,000 in seed funding for interdisciplinary initiatives that have the potential to grow into larger collaborations or funding opportunities. During the first year of the ISG DMP requirement, faculty were required to submit a DMP conforming to local templates and guidelines with all proposals, but the DMPs were not used in the evaluation process. This changed in 2015 when faculty were notified via the initial announcement for the competition that DMPs would be included in overall scores for ISG proposals.

Because the ISG program is open to all faculty on campus, many applicants, particularly from the arts and humanities, were encountering a DMP requirement for the first time during the ISG application process. In order to provide support for these individuals and any other faculty members who needed assistance creating a DMP for the competition, the campus Research Data Services (RDS) organization (a collaboration led by the Office of the Vice Chancellor for Research, Research Computing, and the University Libraries) offered two brown-bag sessions that were open to all faculty on campus. One of these sessions was geared toward the arts and humanities faculty, while the other was targeted to faculty from the sciences and social sciences. In addition, the regular RDS channels for DMP consultations were available to all faculty prior to the ISG proposal deadline. These channels included email, telephone, and in-person consultations as well as reviews of DMPs within the DMPTool (http://dmptool.org) interface. This paper examines the entire ISG DMP requirement process, including analyses of DMP scores, as a potential assessment tool for evaluating current RDS support offerings as well as an outreach opportunity to promote data management best practices and RDS services.

Literature Review

In response to DMP policies and requirements from research funders, many academic institutions have developed support services for DMP creation and consultation. A recent study found that 20.5% of 221 academic libraries surveyed already offered DMP consulting services with another 22.2% planning to offer such services at the time of the survey (Tenopir et al. 2012). One possible explanation for the prevalence of these types of services is that many funder data policies are general or ambiguous, which presents an opportunity for those with expertise in data management to provide guidance (Dietrich et al. 2012). The lack of clarity around DMP requirements and related policies is further exacerbated by the amount of variation, in terms of both expected content and format, across an increasing number of federal agencies (Thoegersen 2015).

Surveys of researchers also point to the need for DMP support services. Shortly after the NSF DMP requirement was announced, Cornell University conducted a survey of NSF researchers, which revealed that researchers had uncertainty about the requirement but were also receptive to assistance with understanding it (Steinhart 2012). A survey conducted at Carnegie Mellon University found that 73% of faculty respondents would be interested in DMP support for grant proposals, which was the highest level of interest in any of the service options presented (Van Tuyl and Michalek 2015). A need for increased DMP support services was also found in a survey of researchers conducted at Northwestern University (Buys and Shaw 2015).

Models for developing and providing DMP and other data management services vary widely from institution to institution (Raboin et al. 2013, Coates 2014). In addition to providing DMP consultation services, libraries also host workshops on DMPs and related topics. An evaluation of a data management workshop at Iowa State University found that participants were particularly interested in tools and resources for creating DMPs (O'Donnell and Bowen 2014). At the University of Minnesota, a DMP workshop that has been offered since late 2010 helps to satisfy a Principal Investigator eligibility requirement and has been successful in terms of building campus partnerships and demonstrating library expertise (Johnston et al. 2012). DMP support is offered not only for grant proposals but also for general good practice at Purdue University, and this includes support targeted at junior faculty who are just setting up their research laboratories (Sapp Nelson 2015).

In addition to providing support for creating DMPs, institutions can benefit from analyzing DMPs as a way to understand local researchers' data practices. A multi-institutional effort to formalize this type of analysis is currently underway with the goal of creating a rubric that can be used broadly by the community (Rolando et al. 2015). Results from an analysis of DMPs at the Georgia Institute of Technology showed that one third of the plans contained identical language used in at least one other plan suggesting a need to maintain consistent and current boilerplate text wherever appropriate (Parham and Doty 2012). A study at the University of Illinois at Urbana-Champaign (UIUC) analyzed 1,260 NSF DMPs and found that there was no statistically significant relationship between where data was stored and whether or not proposals were funded (Mischo et al. 2014). In addition, the UIUC study found that 44.1% of DMPs mentioned sharing data via some form of traditional scholarly publication, which the authors believe shows a disconnect with NSF expectations for data dissemination (Mischo et al. 2014). A study at the University of Minnesota Libraries found a similar result in their analysis of 182 NSF DMPs with 74% mentioning sharing data via traditional journal publication: The authors of that study suggest that library data services could help address this need for better understanding of data sharing in the context of federal funding agency requirements (Bishoff and Johnston 2015). An analysis of data management plans from accepted NSF proposals conducted at the University of Michigan reinforced the need for more education and outreach about DMPs and support services that was also seen in the results of the researcher surveys mentioned above (Nicholls et al. 2014).

While the literature provides many examples of DMP support services and analyses of the content of them, there is little objective data on the effects that these services have on the content and quality of the DMPs created by researchers who utilize them. This paper uses the results of a DMP requirement for a campus-level grant competition to examine such effects. In addition, the process surrounding this DMP requirement is explored as a novel approach to outreach about DMP support services, which is another area of need identified in the literature.

Methods

Nine volunteers from the campus Research Data Advisory Committee (RDAC) — a committee comprising faculty and other research data experts on campus who advise RDS efforts — each scored a randomly assigned portion of the total number of DMPs submitted with the 2015 ISG proposals. These reviewers used level of adherence to local DMP guidelines (https://data.colorado.edu/cudmpguidance) as the basis for a score ranging from 0 (poor) to 5 (excellent). These local guidelines, which were created by RDAC, were also advertised to ISG applicants in the announcement about the DMP requirement for the ISG competition as well as during brown bags and consultations. Each DMP was scored by three RDAC members, and the three scores were then averaged together to obtain the final DMP score. Two of the nine reviewers were the same RDS members who responded to faculty requests for consultations. Because of this, their individual scores for DMPs from faculty who requested consultations were not included in the analyses in order to eliminate a source of potential bias. This was done for four scores out of 246 total individual scores. In those four cases, only two reviewer scores were averaged together for the final DMP scores. Interrater reliability for all scores was measured using intraclass correlation.

In order to evaluate the effectiveness of the DMP support that RDS provided at the brown-bag sessions and via individual consultations, unpaired t-tests were used to compare the means of DMP scores for faculty who attended a brown-bag session or requested a consultation with those who did not. Unpaired t-tests were also used to compare the means of DMP scores for faculty whose ISG proposals were funded with faculty whose proposals were not funded. Upon submission to the ISG competition, each faculty member self-assigned their proposal to one of six disciplinary categories: Arts and Humanities, Basic Physical Sciences, Biomedical Sciences, Engineering and Applied Sciences, Geological and Environmental Sciences, and Social Sciences and Professional Schools. One-way ANOVA was used to compare mean DMP scores among these six disciplinary categories.

Results

A total of 82 proposals were submitted to the 2015 ISG competition, and all 82 contained DMPs. The mean score for all DMPs was 3.11 (SD = 1.32). Inter-rater reliability was strong with an intraclass correlation of 0.72. See Figure 1 for a comparison of mean scores by broad disciplinary category. The Arts and Humanities category received 10 submissions with a mean DMP score of 2.50 (SD = 2.07), which was the lowest of the six disciplinary categories. The Engineering and Applied Sciences category had the next lowest mean DMP score of 2.84 (SD = 1.25) across 33 submissions. The Biomedical Sciences category received 14 submissions with a mean DMP score of 3.00 (SD = 1.33) while the Social Sciences and Professional Schools category had nearly the same mean score at 3.02 (SD = 1.29) across nine submissions. The Basic Physical Sciences, with 10 submissions, and Geological and Environmental Sciences, with six submissions, had the highest mean DMP scores of 3.42 (SD = 0.92) and 4.00 (SD = 1.25), respectively. The differences in mean DMP scores across disciplines were not significant according to the results of the one-way ANOVA test, $F(5, 76) = 1.21$, p = .31. Mean DMP scores for the 28 faculty whose ISG proposals were ultimately funded (N = 3.36, SD = 1.42) were higher than those of the 54 faculty whose proposals were not funded (N = 2.82, SD = 1.30); however, this difference was also not significant, $t(80) = 1.72$, $p = 0.09$.

Figure 1: Comparison of mean DMP scores for six broad disciplinary categories.
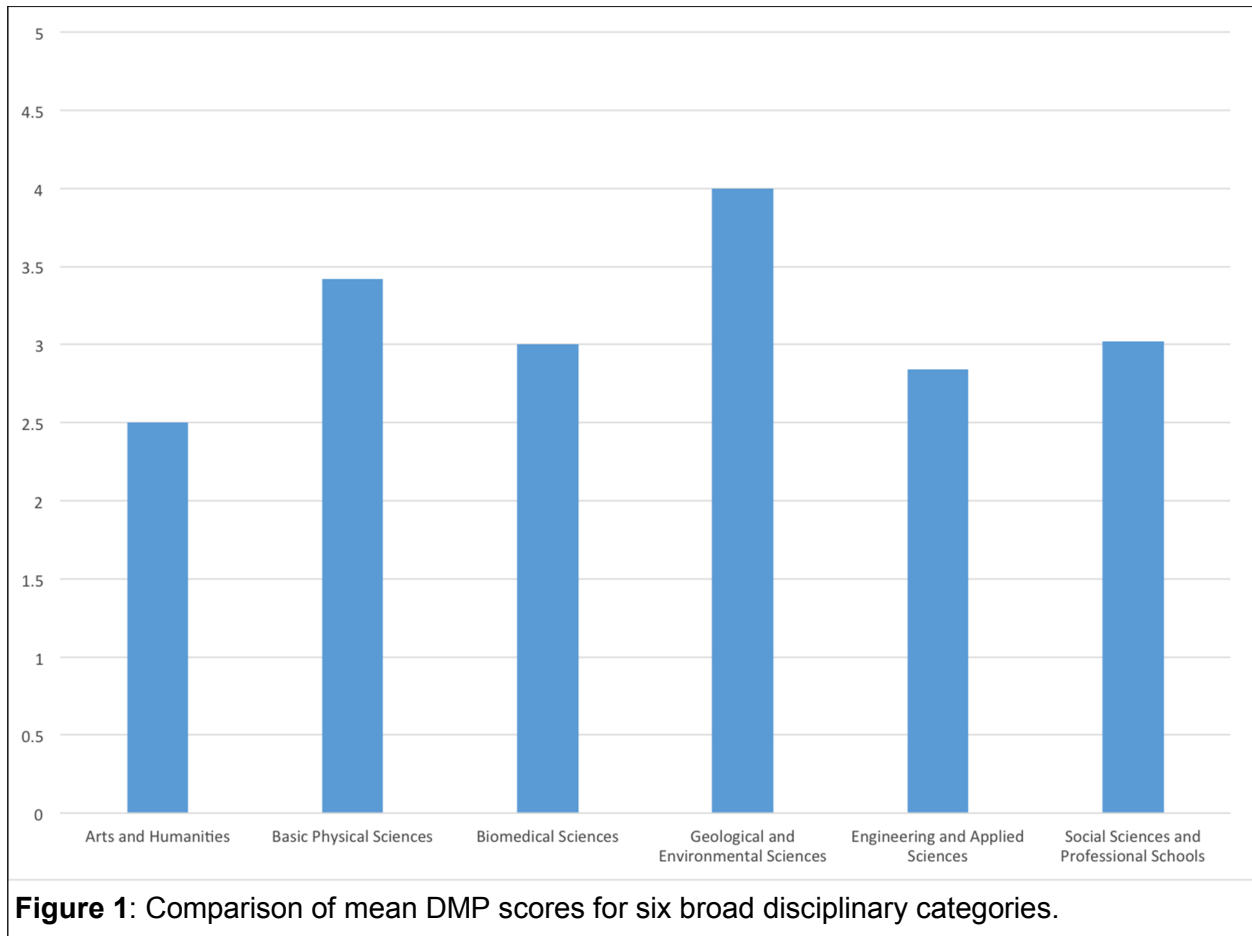
Of the 82 faculty members who submitted a proposal to the ISG competition, 16 (19.5%) either requested a consultation or attended a brown-bag session. While only seven (8.5%) faculty members who attended a brown-bag session ended up submitting a proposal, attendance records show that at least 14 other individuals attended those events as well. A total of nine (11%) faculty who submitted ISG proposals consulted with RDS members about their DMP via phone, email, in-person meeting, or DMPTool review. See Figures 2 and 3 for comparisons of mean DMP scores for faculty who sought a consultation or attended a brown bag versus those who did not. Mean DMP scores for faculty who requested a consultation (N = 4.00, SD = 0.98) were significantly higher ($t(73)$ = 2.52, p = .01) than those who did not seek any assistance (N = 2.84, SD = 1.33). Mean DMP scores for faculty who attended a brown-bag session (N = 3.24, SD = 1.63) were also higher than mean DMP scores for faculty who did not utilize any assistance (N = 2.84, SD = 1.33); however, this difference was not significant, $t(71)$ = 0.73, p = 0.47.

Figure 2: Comparison of mean DMP scores for faculty who requested a consultation with those who did not seek assistance (*p < 0.05).

Figure 3: Comparison of mean DMP scores for faculty who attended a brown bag with those who did not seek assistance.
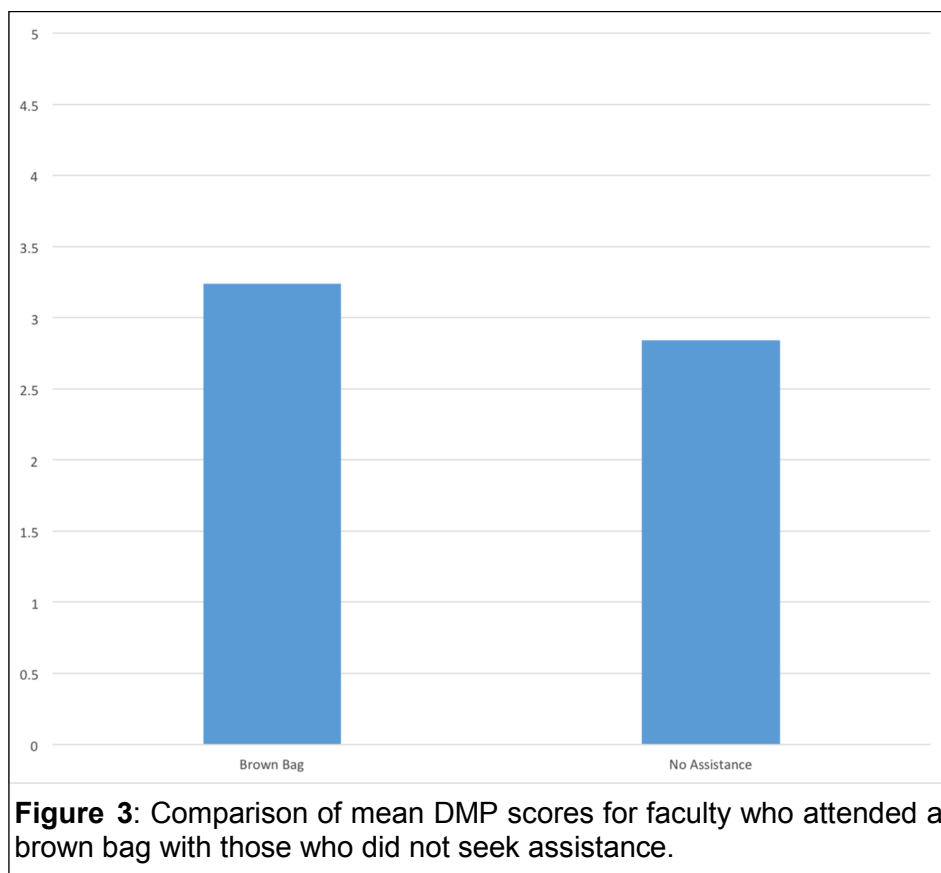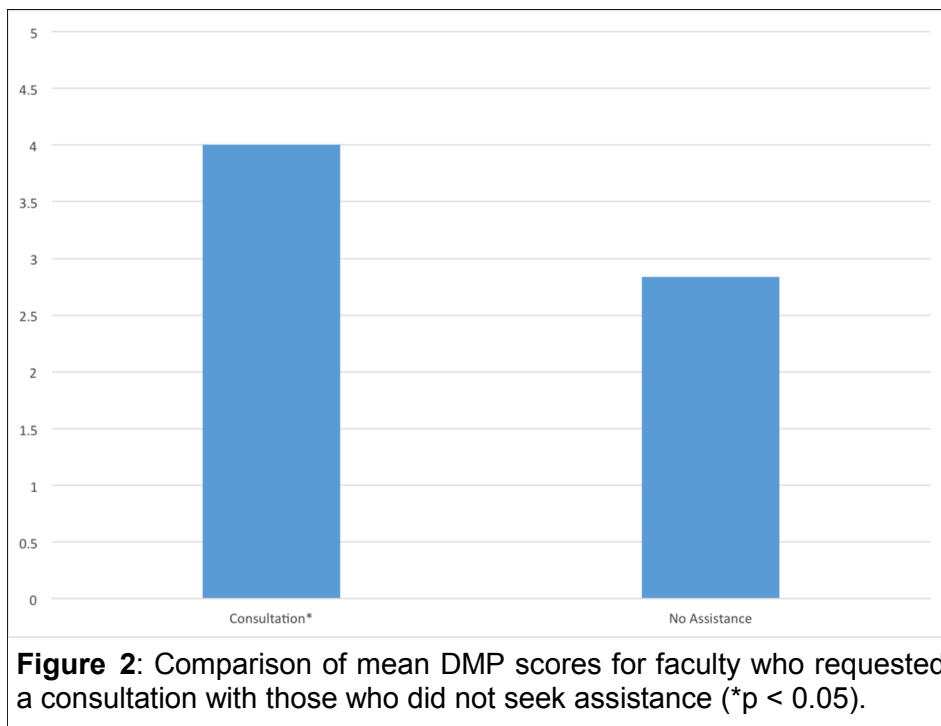
Discussion

The implementation of a DMP requirement for the ISG competition provided a number of RDS outreach and assessment opportunities, both expected and unexpected. Because the competition was open to all faculty on campus, faculty who do not typically submit grant proposals to research funders with DMP requirements were introduced to the concept of a DMP for the first time. This presented an opportunity for general outreach about research data management best practices to faculty who would likely not encounter or seek out RDS services for DMP support. Unexpectedly, some faculty who had written DMPs for NSF and other funding agency proposals in the past mentioned to RDS personnel during consultations and the brown-bag sessions that they were taking the ISG requirement more seriously than the requirements for their past grant proposals because of its prominence in the Office of the Vice Chancellor for Research's announcements about the competition. Brown-bag attendance was about what was expected. While nine total faculty requesting DMP consultations might not seem like an overwhelming number, it is rare for RDS to receive multiple DMP consultation requests related to a single funding opportunity.

The ISG DMP requirement also provided a way to assess current RDS outreach and support services. Because RDS was involved with the entire DMP requirement process for the ISG competition, it was possible to track the number of people who requested consultations or attended brown-bag sessions relative to the total number of ISG applicants. It is currently not possible for RDS to do this type of tracking for federal funding opportunities that require DMPs. Having some control over the ISG DMP process allowed RDS to measure the effect of its consultation and brown-bag offerings on DMP scores, which is not possible in the context of DMPs for grant proposals to federal agencies. The significant difference in mean DMP scores for faculty with whom RDS consulted versus faculty who did not seek any assistance suggests that RDS services were effective in improving the quality of ISG DMPs.

The lack of significant differences among the six broad disciplinary categories suggests that familiarity with existing federal funding agency DMP requirements might not have helped applicants with the ISG DMP process. Despite high numbers of NSF-funded researchers in disciplines like Engineering and Applied Sciences, many of whom likely would have created a DMP prior to the ISG competition, the mean score for that discipline was not much higher than the mean score for Arts and Humanities, a discipline with far less funding from agencies with DMP requirements. That said, the highest mean DMP score did come from a discipline, Geological and Environmental Sciences, with a great deal of NSF funding and strong cultures of data sharing in some sub-disciplines, while the lowest mean score came from the discipline that most likely has the least amount of NSF funding. The higher, yet not significant, mean score for faculty whose proposals were funded versus those whose proposals that were not funded is not all that surprising given that the DMP score was a component of the overall ISG scoring process.

While there were clear benefits to RDS having some control over the DMP process for the ISG competition, this level of coordination among DMP creators, DMP reviewers, and RDS providers is also a key limitation for generalizing the results of the analyses. For example, the high interrater reliability could be explained by having clear guidelines and expectations that RDAC worked together to create and then use for the scoring process. Likewise, guidance that RDS provided during consultations and brown-bag sessions was based on a similar shared

understanding of what constitutes a high-quality DMP. It is unlikely that this level of coordination between guidance provided to DMP creators and expectations for DMP reviewers occurs for DMPs accompanying grant proposals to federal funding agencies, and even if it did, the specific criteria might differ from what was used during the ISG competition. For example, while there is evidence that RDS DMP support services were effective in improving the quality of DMPs as defined by the criteria used for the ISG competition, it is not clear that the same services would improve the quality of a DMP in the eyes of an NSF reviewer.

These findings suggest that there is value in coordination among guidelines, review criteria, and support services for DMPs, which are all created and used by different individuals from various institutions in the case of federal funding agency requirements. If funding agencies were to take additional steps to get all individuals involved in the DMP creation and evaluation process on the same page, then the desired effects of DMP requirements (e.g., better data management) could be more easily achieved.

Conclusion

While the results described in this paper might not be generalizable to DMP review processes at federal funding agencies, they do suggest the importance of developing a clear and shared understanding of what constitutes a high-quality DMP among grant applicants, grant reviewers, and providers of research data services for any DMP requirements, whether local or at the research funder level. By having a strong influence on the entire DMP evaluation process from start to finish, CU-Boulder RDS was able to demonstrate that its DMP support services significantly improved the quality of DMPs submitted to the 2015 ISG competition. These assessment results also suggested that direct consultations were more effective than brown-bag attendance, which could indicate a need to emphasize support services that are tailored to individual researchers rather than one-size-fits-all approaches. Future iterations of the ISG DMP process or similar efforts at other institutions could also use this assessment model to evaluate the relative effectiveness of other types of support services. In addition to allowing for novel forms of assessment of RDS services, the ISG DMP requirement provided unique opportunities to reach out to faculty in departments who do not typically have to write DMPs for grant proposals. Even for faculty who were already familiar with DMPs for NSF proposals, for example, the ISG requirement drove home the emphasis that the campus is placing on data management best practices, which is intended to reinforce the messages coming from federal agencies and other research funders.

Supplemental Content

Data File
An online supplement to this article can be found at http://dx.doi.org/10.7191/jeslib.2016.1089 under "Additional Files".

Acknowledgement

Advisory Committees for their work developing local DMP guidelines and for their assistance with scoring the ISG DMPs.

Disclosure

The author(s) report no conflict of interest.

References

Bishoff, Carolyn, and Lisa Johnston. 2015. "Approaches to Data Sharing: An Analysis of NSF Data Management Plans from a Large Research University." *Journal of Librarianship and Scholarly Communication* 3(2), p.eP1231. http://doi.org/10.7710/2162-3309.1231

Buys, Cunera M., and Pamela L. Shaw. 2015. "Data Management Practices Across an Institution: Survey and Report." *Journal of Librarianship and Scholarly Communication* 3(2), p.eP1225. http://doi.org/10.7710/2162-3309.1225

Coates, Heather. 2014. "Building Data Services From the Ground Up: Strategies and Resources." *Journal of eScience Librarianship* 3(1): e1063. http://doi.org/10.7191/jeslib.2014.1063

Dietrich, Dianne, Trisha Adamus, Alison Miner, and Gail Steinhart. 2012. "De-Mystifying the Data Management Requirements of Research Funders." *Issues in Science and Technology Librarianship* 2012 (Summer). http://doi.org/10.5062/F44M92G2

Johnston, Lisa, Meghan Lafferty, and Beth Petsan. 2012. "Training Researchers on Data Management: A Scalable, Cross-Disciplinary Approach." *Journal of eScience Librarianship* 1(2): e1012. http://doi.org/10.7191/jeslib.2012.1012

Mischo, William, Mary Schlembach, and Megan O'Donnell. 2014. "An Analysis of Data Management Plans in University of Illinois National Science Foundation Grant Proposals." *Journal of eScience Librarianship* 3(1): e1060. http://doi.org/10.7191/jeslib.2014.1060

Nicholls, Natsuko H., Sara M. Samuel, Leena N. Lalwani, Paul F. Grochowski, and Jennifer A. Green. 2014. "Resources to Support Faculty Writing Data Management Plans: Lessons Learned from an Engineering Pilot." *International Journal of Digital Curation* 9(1): 242–252. http://doi.org/10.2218/ijdc.v9i1.315

O'Donnell, Megan, and Bonnie Bowen. 2014. "Data Management Brownbag Evaluation Summary." *Library Reports*. Paper 12. http://lib.dr.iastate.edu/libreports/12

Parham, Susan Wells, and Chris Doty. 2012. "NSF DMP Content Analysis: What Are Researchers Saying?" *Bulletin of the American Society for Information Science and Technology* 39(1): 37–38. http://doi.org/10.1002/bult.2012.1720390113

Raboin, Regina, Rebecca Reznik-Zellen, and Dorothea Salo. 2013. "Forging New Service Paths: Institutional Approaches to Providing Research Data Management Services." *Journal of eScience Librarianship* 1(3): e1021. http://doi.org/10.7191/jeslib.2012.1021

Rolando, Lizzy, Jake Carlson, Patricia Hswe, Susan Wells Parham, Brian Westra, and Amanda L. Whitmire. 2015. "Data Management Plans as a Research Tool." *Bulletin of the American Society for Information Science and Technology* 41(5): 43–45. http://doi.org/10.1002/bult.2015.1720410510

Sapp Nelson, Megan. 2015. "Data Management Outreach to Junior Faculty Members: A Case Study." *Journal of eScience Librarianship* 4(1): e1076. http://doi.org/10.7191/jeslib.2015.1076

Steinhart, Gail, Eric Chen, Florio Arguillas, Dianne Dietrich, and Stefan Kramer. 2012. "Prepared to Plan? A Snapshot of Researcher Readiness to Address Data Management Planning Requirements." *Journal of eScience Librarianship* 1(2): e1008. http://doi.org/10.7191/jeslib.2012.1008

Tenopir, Carol, Ben Birch, and Suzie Allard. 2012. "Academic Libraries and Research Data Services: Current Practices and Plans for the Future." *Association of College and Research Libraries (ACRL) White Paper*. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf

Thoegersen, Jennifer. 2015. "Examination of Federal Data Management Plan Guidelines." *Journal of eScience Librarianship* 4(1): e1072. http://doi.org/10.7191/jeslib.2015.1072

Van Tuyl, Steve, and Gabrielle Michalek. 2015. "Assessing Research Data Management Practices of Faculty at Carnegie Mellon University." *Journal of Librarianship and Scholarly Communication* 3(3), p.eP1258. http://doi.org/10.7710/2162-3309.1258