

Analysis of the Service System of the Data Management Plan Tool DMPTool: Postprint

Authors: Huang Guobin, Di Hongyang, Zhang Shasha, Zhao Qian

Date: 2023-08-26T00:00:00+00:00

Abstract

[Purpose/Significance] This study examines the representative data management planning tool DMPTool, analyzing its service system to introduce data management planning tools for domestic libraries developing scientific data services.

[Method/Process] Employing systematic testing and literature research methods, and based on a review of DMPTool's development history and current research status, this study summarizes the foundational services of its data management plan templates, cases, and guidelines; systematically outlines its service system encompassing data management plan services, template and guideline services, and user information services; analyzes its service approaches for researchers and institutions; and synthesizes the advantages and limitations of its application in scientific data services of American university libraries.

[Results/Conclusion] DMPTool aligns with the requirements of U.S. research funding agencies for data management plans and has been adopted by numerous university libraries as a key component of scientific data services. While DMPTool cannot be directly transplanted to domestic library scientific data services, its service system offers valuable insights for adaptation.

Full Text

Preamble

Data Management Plan Tool DMPTool: An Analysis of Its Service Architecture

Huang Guobin, Di Hongyang, Zhang Shasha, Zhao Qian
School of Government, Beijing Normal University, Beijing 100875

Abstract

[Purpose/Significance] This study examines DMPTool, a representative data management plan tool, analyzing its service architecture to introduce such tools to domestic libraries developing research data services. **[Method/Process]** Using system testing and literature review, this paper traces DMPTool's development history and research status, summarizes its service foundations (templates, examples, and guidance), outlines its service systems (data management plan services, template and guidance services, and user information services), analyzes its service methods for researchers and institutions, and identifies advantages and limitations in its application within U.S. academic library research data services. **[Result/Conclusion]** DMPTool aligns with U.S. research funding agencies' requirements for data management plans and has been adopted by numerous academic libraries as a key component of research data services. While DMPTool cannot be simply transplanted to domestic library research data services, its service architecture offers valuable lessons.

Keywords: data management plan; research data service; DMPTool

Classification Number: G250

DOI: 10.13266/j.issn.0252-3116.2018.04.005

1. Introduction

Government-funded research projects constitute a major source of scientific data production, and the data they generate attract significant attention from researchers due to their critical role in driving scientific discovery. On September 26, 2006, then-U.S. President Bush signed the Federal Funding Accountability and Transparency Act of 2006, legislatively mandating the disclosure of information about all entities receiving federal funds. Under this act, the National Institutes of Health (NIH) required researchers to submit data management plans (DMPs) when applying for high-value grants [1]. Subsequently, the National Oceanic and Atmospheric Administration (NOAA), the National Science Foundation (NSF), and other funding agencies implemented similar requirements [2]. A data management plan is a formal document outlining how researchers intend to store, manage, and share scientific data during and after research projects. Comprehensive DMPs are essential for securing funding and for effectively managing project data while mitigating potential issues. However, resources provided by funding agencies like NIH, NOAA, and NSF often fail to meet researchers' practical needs. Lacking theoretical knowledge and practical experience in research data management, researchers frequently turn to libraries for information and support. In response, libraries have expanded DMP reference materials and guidance services while designing and developing data management plan tools to help researchers efficiently create compliant DMPs.

2. DMPTool Development Status

Representative tools such as DMPTool, supported by the University of California Curation Center (UC3) at the California Digital Library (CDL), have matured in their service foundations, architecture, and delivery methods, earning adoption by researchers, institutions, and funding agencies across the U.S. and globally.

DMPTool development began in January 2011 when UC3 collaborated with seven other organizations to design a tool that would integrate expertise across institutions and reduce the cost of data management in response to NIH and NSF requirements. The eight participating organizations included UC3, UCLA Library, UC San Diego Library, the University of Illinois at Urbana-Champaign Library, the University of Virginia Library, DataONE, the Digital Curation Centre (DCC), and the Smithsonian Institution. Each organization independently contributed to development, releasing a beta version in August 2011 and the official version in October 2011. In January 2013, the development partnership received funding from the Alfred P. Sloan Foundation, and in May 2014 released DMPTool version 2. This new version retained the original name but introduced significant updates, supporting more diverse user roles, more comprehensive workflows for creating and managing DMPs, richer reference resources, and broader compliance with funding agency requirements [2-4].

Concurrent with DMPTool's evolution, scholarly research on the tool has emerged domestically and internationally. Abroad, J. Starr et al. described UC libraries' tools and services for different stages of the research data lifecycle, noting DMPTool's role in providing reference information for DMP creation and strengthening library-researcher connections [5]. A. Sallans and M. Donnelly compared DMPTool with DMPOnline, identifying cultural and philosophical differences while acknowledging their shared goal of improving data management practices [4]. C. Strasser, S. Abrams, and P. Crus discussed version 2's development goals, process, features, and management, attributing its success to clear functional requirements, defined division of labor, user feedback from previous versions, and Sloan Foundation funding [2]. Domestically, Ma Jianling and Cao Yuezhen classified DMPTool as a data processing tool for DMP creation [6]. Wang Kai, Peng Jie, and Qu Baoqiang compared DMPTool and DMPOnline, examining their development paths, service methods, and requirement sources [7]. Wang Pu introduced DMPTool's development and objectives, offering recommendations for library services [8]. Wu Hairu traced DMPTool's history, noting its ability to help researchers quickly create DMPs, connect institutions with funders, and meet diverse customization needs [9]. However, existing research has focused primarily on overview rather than deep analysis of DMPTool's service architecture. This study aims to fill that gap by examining its service foundations, architecture, and methods to inform domestic library practice.

3. DMPTool Service Foundations

DMPTool's service architecture rests on three foundations: DMP templates, examples, and guidance, all aligned with funding agency requirements to support efficient DMP creation.

3.1 Data Management Plan Templates

DMP templates are structured outlines developed according to funding agency requirements to guide researchers in creating DMPs. DMPTool's templates are authored by partner organizations and consist of content frameworks and specific requirements for each section. Different templates vary in structure and requirements based on the authoring organization and funding agency. Content frameworks typically include: (1) data description (type, scale, objects, production methods); (2) data and metadata standards (file formats, metadata standards); (3) data access and sharing (timing, methods, privacy/intellectual property considerations); (4) redistribution and reuse (who can use the data and under what conditions); and (5) archiving and preservation (location, duration, associated resources). Specific requirements include instructions, mandatory/optional status, format constraints (text, numeric values with units, dates in YYYY-MM-DD format, or preset options), and other restrictions. By following these requirements, researchers can generate compliant DMP text [10-11].

3.2 Data Management Plan Examples

DMP examples are sample documents created and shared by researchers to provide reference for others. As of July 13, 2017, DMPTool had publicly shared 131 DMP examples based on 31 templates reflecting over 10 funding agencies' requirements [Figure 1: see original paper]. NSF-related examples dominated with 74 (56.5% of total), covering atmospheric and geospace sciences (12), social/behavioral/economic sciences (12), geosciences (10), computer and information science engineering (9), education and human resources (9), and biological sciences (6). Other examples addressed requirements from USGS, NEH Office of Digital Humanities, DOE, USDA, and NIH, spanning energy, education, agriculture, and geology.

The 121 creators of these examples represented over 50 organizations, primarily U.S. universities, reflecting DMPTool's focus on American users and the central role of universities in federally funded research. Most creators shared only one example; only two individuals shared multiple examples (one shared two, another three), while the rest shared single examples [12]. This suggests that a culture of sharing DMPs has not yet fully developed, with insufficient external incentives or mandates for researchers to share.

For DMPTool, examples serve dual purposes: they build an open, shared discovery environment where researchers can learn from peers' approaches to common data management questions, yet they also pose risks since examples are created

based on researchers' own interpretations without funding agency verification, potentially leading to non-compliance.

3.3 Data Management Plan Guidance

DMP guidance comprises explanatory resources developed by partner organizations to clarify funding agency requirements. Embedded within templates, guidance supplements template and example usage through template-level and requirement-level guidance. Template-level guidance appears across all requirements and includes textual instructions and resource links. Requirement-level guidance appears for specific requirements and includes textual instructions, links, recommended responses, and example answers. Textual instructions help researchers understand requirements; resource links point to relevant webpages or documents; recommended responses provide copyable text; and example answers offer reference material [13]. Guidance is characterized by its specificity (targeting DMPTool usage), explanatory nature (clarifying template requirements), and directionality (guiding response content).

4. DMPTool Service Architecture

DMPTool's service architecture comprises DMP services, template and guidance services, and user information services, all working toward defined service objectives.

4.1 Service Objectives

Early DMPTool versions aimed to provide researchers with a simple DMP creation method by offering funding agency requirements, explanations, and links to policies and guidance. Through user feedback, current versions have expanded objectives to include institution-specific information: resources and services for managing research data, help text for specific requirements, recommended copy-paste responses, and internal data management information and events [14-15]. This evolution reflects a shift from relying solely on funding agencies for guidance to emphasizing institutional responsibility for interpreting requirements, developing supportive templates and guidance, and encouraging DMP sharing within and across institutions.

4.2 Data Management Plan Services

DMP services encompass creation, management, and updating of DMP descriptive and content information; copying DMP information to quickly create new plans; exporting DMPs as PDF or DOCX files; deletion; and institutional review workflows where managers can approve or reject submitted DMPs based on funding agency compliance.

4.3 Template and Guidance Services

These services support authoring, management, and sharing of DMP templates and guidance. For authoring, organizations can create templates by completing basic information and specific requirements, embedding guidance, and activating templates. Templates have two visibility levels: public (available to all DMPTool users) and institutional (available only to organizational members) [11]. For management, organizations can add, delete, modify, sort, or group requirements during the drafting phase. Once activated, templates cannot be modified; revisions require creating a new template, deprecating the old one (marked “obsolete”), and specifying activation dates [17].

4.4 User Information Services

These services manage user profiles and subscription information. User profiles for researchers include username, email, ORCID, and API token; for organizations, they include full name, abbreviation, website URL, contact information, organization identifier, description, and auto-reply email templates for review requests. Subscription services enable email notifications for permission changes, co-creation invitations, visibility changes, and new comments or reviews [18].

5. DMPTool Service Methods

DMPTool offers different functionalities and methods for DMP creators (researchers) and managers (organizations).

5.1 For Researchers

DMPTool leverages institutional templates, examples, and guidance supplemented by public resources to guide researchers through self-service DMP creation, management, and sharing [19]. To create a DMP, researchers can either copy an existing public or institutional example or start from a blank template. During the descriptive information phase, DMPTool prompts researchers to enter the DMP name, associated project identifier, submission deadline, visibility level, and co-creators (name or email). Name and visibility are mandatory. During content entry, templates divide DMPs into sections with accompanying guidance and specific requirements. Researchers can complete sections sequentially or jump between them according to preference [Figure 2: see original paper]. The preview stage allows researchers to review the complete DMP, store it online, export as PDF/DOCX, or submit it for institutional review [16]. The service workflow is illustrated in [Figure 3: see original paper].

5.2 For Organizations

To facilitate management of templates, guidance, and researcher-created DMPs, DMPTool provides four administrative roles with distinct permissions: (1) Template Editor: create and edit institutional templates; (2) Guidance Editor: cre-

ate and edit guidance; (3) Plan Reviewer: review submitted DMPs; and (4) Institutional Administrator: highest authority, can perform all above functions and assign roles. Roles are not mutually exclusive—one administrator can hold multiple roles [20].

To avoid burdening organizations, DMP review is optional. Organizations can configure templates to require formal review, informal review, or no review. Formal review requires institutional approval before marking a DMP as complete, while informal review allows researchers to mark completion anytime except during active review. Both review types are conducted solely by institutional administrators without funding agency participation, and informal review does not mandate revisions, so neither guarantees full compliance with funder requirements [21].

6. DMPTool Application in U.S. Academic Libraries

DMPTool has been widely adopted by U.S. academic libraries as a key component of research data services, demonstrating distinct advantages and limitations.

6.1 Application Status

As of November 15, 2017, DMPTool had partnered with 216 organizations, including over 140 U.S. universities. Most partner universities integrate campus authentication, allowing users to log in via institutional credentials without separate registration. Common library practices include introducing DMPTool functionalities and providing access points, as seen at Harvard, Princeton, and Stanford libraries.

Other partners include funding agencies (NEH, NIH, NOAA, NSF), scientific data organizations (DataONE, DCC, IEDA), research institutions (Lawrence Berkeley, Los Alamos, and Oak Ridge National Laboratories), and international universities (Griffith University, Nanyang Technological University, University of Helsinki), plus medical organizations like GATACA and Mayo Clinic [22].

6.2 Application Advantages

DMPTool offers three key advantages: First, as an economical tool, it requires no hardware purchases, software installation, or usage fees—researchers simply register and log in, with many accessing via institutional accounts, avoiding financial burdens on libraries. Second, as a practical tool, it helps researchers address data management questions across the research lifecycle, create compliant DMPs, and enables libraries to manage DMP resources and deliver education. Third, as an open tool, its API provides access to partnership statistics, administrator data, public DMP information, user lists, and template lists, supplying raw data for studying DMP creation behaviors and informing service improvements [4, 23].

6.3 Application Limitations

Three main limitations exist: First, DMPTool primarily serves U.S. researchers, with templates, examples, and guidance reflecting American funder requirements, hindering adoption by libraries in other countries and regions. Second, DMPTool has not achieved universal recognition from U.S. funding agencies; while some provide links, it remains a third-party tool dependent on partner organizations for templates and guidance. Third, online data storage poses security risks, and the tool lacks integration with services for other research data lifecycle stages.

7. Conclusion

DMPTool emerged from researchers' DMP creation and management needs, building a service architecture founded on templates, examples, and guidance, with services for DMPs, templates/guidance, and user information, delivered through distinct methods for researchers and organizations. While well-suited to U.S. funding requirements and difficult to directly transplant to domestic libraries, DMPTool embodies valuable concepts, processes, and methods for DMP creation, management, and sharing that merit consideration by Chinese libraries seeking to develop tools and services tailored to domestic researchers' needs and deliver comprehensive research data services throughout the data lifecycle.

References

- [1] NIH. Federal funding accountability and transparency act [EB/OL]. [2017-08-06]. https://grants.nih.gov/grants/public_{accountability}/ffata.htm.
- [2] STRASSER C, ABRAMS S, CRUSE P. DMPTool 2: expanding functionality for better data management planning [J]. *International journal of digital curation*, 2014, 9(1): 324-330.
- [3] University of California. About the DMPTool [EB/OL]. [2017-08-06]. <https://blog.dmptool.org/about-the-dmptool/>.
- [4] SALLANS A, DONNELLY M. DMPTool and DMPOnline: different strategies toward a shared goal [J]. *International journal of digital curation*, 2012, 7(2): 123-129.
- [5] STARR J, WILLETT P, FEDERER L, et al. A collaborative framework for data management services: the experience of the University of California [J]. *Journal of eScience librarianship*, 2012, 1(2): 109-114.
- [6] 马建玲, 曹月珍. 研究数据管理工具发展研究 [J]. *图书馆学研究*, 2014(15): 40-47.
- [7] 王凯, 彭洁, 屈宝强. 国外数据管理计划服务工具的对比研究 [J]. *图书与情报*, 2015(3): 103-109.
- [8] 王璞. 英美两国制定数据管理计划的政策、内容与工具 [J]. *新世纪图书馆*, 2015(5): 69-72.

- [9] 吴海茹. 加州数字图书馆数据管理计划工具研究及思考 [J]. 情报杂志, 2014(12): 203-206, 169.
- [10] University of California. Data management general guidance [EB/OL]. [2017-08-06]. https://dmptool.org/dm_{guidance}.
- [11] STRASSER C. Template overview [EB/OL]. [2017-08-06]. <https://github.com/CDLUC3/dmptool/wiki/Template-Overview>.
- [12] University of California. Public DMPs [EB/OL]. [2017-08-06]. https://dmptool.org/public_{dmps}.
- [13] STRASSER C. Resources chart [EB/OL]. [2017-08-06]. <https://github.com/CDLUC3/dmptool/wiki/Resources-Chart>.
- [14] UC3. DMPTool 2 administrator webinar #2 [EB/OL]. [2017-08-06]. <https://www.slideshare.net/UC3/dmptool2-administrator-webinar-2>.
- [15] UC3. Data management plans: tips, tricks and tools [EB/OL]. [2017-08-06]. <https://www.slideshare.net/UC3/data-management-plans-tips-tricks-and-tools>.
- [16] University of California. Quick start guide [EB/OL]. [2017-08-06]. <https://dmptool.org/quickstartguide/>.
- [17] SIMMSS. Template creation [EB/OL]. [2017-08-06]. <https://github.com/CDLUC3/dmptool/wiki/Template-Creation>.
- [18] SIMMSS. Institution profile customization [EB/OL]. [2017-08-06]. <https://github.com/CDLUC3/dmptool/wiki/Institution-Profile-Customization>.
- [19] SHREEVES SL. Presenting the new and improved DMPTool [EB/OL]. [2017-08-06]. <http://hdl.handle.net/2142/49957>.
- [20] STRASSER C. Administrator roles [EB/OL]. [2017-08-06]. <https://github.com/CDLUC3/dmptool/wiki/Administrator-Roles>.
- [21] SIMMSS. Review functionality [EB/OL]. [2017-08-06]. <https://github.com/CDLUC3/dmptool/wiki/Review-Functionality>.
- [22] University of California. Partner institutions [EB/OL]. [2017-11-15]. https://dmptool.org/partners_{list}.
- [23] SIMMSS. API [EB/OL]. [2017-08-06]. <https://github.com/CDLUC3/dmptool/wiki/API>.

Author Contributions

Huang Guobin: Topic selection, writing, and revision;
 Di Hongyang: Comprehensive data collection, processing, analysis, and writing;
 Zhang Shasha: Domestic and international literature review, abstract writing;
 Zhao Qian: DMPTool online testing and documentation collection.

Analysis of Service Architecture of Data Management Plan Tool DMPTool

Huang Guobin, Di Hongyang, Zhang Shasha, Zhao Qian
School of Government, Beijing Normal University, Beijing 100875

Abstract: [Purpose/significance] The research takes the significant data management plan tool, DMPTool, as the object, and analyzes its features for domestic libraries developing research data services. [Method/process] Based on the system test method and literature research, the article reviewed the development history and the research status of DMPTool, analyzed its service basics of templates, examples and resources, service systems of data management plan services, template and resource services, and users' profile services, and service methods for researchers and institutions. The research also summarized the advantages and limitations of DMPTool application in American academic libraries' research data services. [Result/conclusion] DMPTool is compatible with American funders' requirements for data management plans, and is used as important component of research data services by many academic libraries. Though DMPTool cannot be simply introduced into domestic libraries' research data services, its features deserve to be learned.

Keywords: data management plan; research data service; DMPTool

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.