Received 25 August 2015
Revised 10 November 2015
Accepted 12 November 2015

Librarians and scientists partner to address data management

Taking collaboration to the next level

Andrea Medina-Smith, Kimberly A. Tryka, Barbara Silcox and Robert J. Hanisch
National Institute of Standards and Technology, Gaithersburg, Maryland, USA

Abstract

Purpose – This study aims to look at the changing way in which the Information Services Office (ISO) at the National Institute of Standards and Technology (NIST) provides services to NIST scientific and technical staff throughout their research and publishing cycles. These services include the more traditional services of a research library and publishing NIST technical reports and *The Journal of Research of NIST* and preserving and exhibiting scientific instruments and other artifacts. ISO has always prided itself on having a close relationship with its customers, providing a high level of service and developing new services to stay in front of NIST researcher needs. Through a concerted, strategic effort since the late 1990s, ISO has developed and promoted relationships with its key customers through its Lab Liaison Program.

Design/methodology/approach – This paper discusses the relationship ISO has developed with the Office of Data and Informatics (ODI), how this relationship was forged and how this collaboration will serve as a model for working with the other labs and programs at NIST. It will also discuss the risks and opportunities of this new collaborative service model, how ISO positioned itself to become an equal partner with ODI in the exploration of solutions to data management issues and the benefits of the relationship from ODI's perspective.

Findings – A pattern of strategic changes to the services and activities offered by the Lab Liaison program has put ISO in the position to collaborate as peers with researchers at NIST.

Originality/value – This study provides an overview of how ISO made strategic decisions to incorporate non-traditional services to support data management at NIST.

Keywords Case study, Libraries, Collaboration, Data management, Research libraries, Government libraries

Paper type Case study

Introduction

Much of the library literature on collaborations among librarians and the customers they serve addresses librarians and faculty collaboration in academic environments. These types of collaborations largely focus on improving student learning and instructional programs. In a research organization, productive, ongoing collaborations

The authors thank the researchers in ODI who provided comments about the collaboration. They also thank Susan Makar, the MML Librarian Liaison and Regina Avila, Digital Services Librarian, who read several versions of the paper and provided thoughtful and valuable insights.

Digital Library Perspectives
Vol. 32 No. 3, 2016
pp. 142-152
Emerald Group Publishing Limited
2059-5816
DOI 10.1108/DLP-08-2015-0010

between librarians and researchers may not occur as readily as they do in academic environments and the types of collaborations differ. While the key ingredients for successful collaborations may be similar across different types of organizations, the librarians and researchers at the National Institute of Standards and Technology (NIST) have found ways to collaborate that grew out of incremental, targeted, relationship-building and a broad and deep understanding of organizational needs and priorities. This paper discusses the evolving collaborations between the Information Services Office (ISO) and the laboratory programs and offices at NIST, focusing on the current collaboration with the Office of Data and Informatics (ODI) in the Material Measurement Laboratory (MML). The activities described in the paper illustrate the variety of ways librarians and researchers can engage in successful collaborations that are of mutual benefit.

Background

The NIST, founded in 1901, is a non-regulatory agency within the United States Department of Commerce. Its mission is "to promote USA innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life". NIST's seven laboratories (Physical Measurement Laboratory, MML, Engineering Laboratory, Information Technology Laboratory, the Communications Technology Laboratory, the Center for Nanoscale Science and Technology and the NIST Center for Neutron Research) conduct measurement, standards and technology research across a range of science and engineering disciplines. The scientists, engineers and guest researchers publish about 1,500 papers annually and generate large amounts of data from critically evaluated standard reference data (SRD) (www.nist.gov/srd/index.cfm) to analyzed data that underpins published research results used by academia, industry and other government agencies.

The ODI in the MML was created in 2014 to provide guidance, assistance and resources to researchers, so that MML data products are more discoverable, usable and interoperable. ODI also facilitates MML's compliance with the US government open-data policy by providing guidance and assistance in the best practices for archiving and annotating research and data outputs. The office is comprised of researchers with expertise in the subject areas MML covers, along with experts in enterprise data architecture design and development. The ODI supports national needs, such as the Materials Genome Initiative (MGI); ODI staff are also actively involved in external data initiatives, such as the National Data Service (NDS) and the Research Data Alliance (RDA).

The ISO provides professional scientific and technical information assistance to NIST research staff throughout their research and publishing cycles through the activities of three programs: the Research Library, the Digital Services and Publishing Group (DSPG) and the museum and history program. ISO is an award winning organization with a deeply rooted commitment to customer service excellence and a culture that encourages collaboration and knowledge-sharing[1]. The essential component of ISO's operating philosophy is the knowledge continuum, a concept first introduced in 1993, by a former ISO director. (Vassallo, 1999) The knowledge continuum represents the various stages of the research process from discovery to dissemination to preservation, operating in one continuous loop (Figure 1). It provides a holistic

framework for providing services to the NIST researchers, and it forms the basis for the ways ISO collaborates with them.

Most of DSPG's work maps to the dissemination and preservation components of the knowledge continuum. In the past, this group primarily assisted researchers through the editing, printing and publishing of the agency publications. Today, DSPG's major activities focus on implementing strategies for increasing the visibility of and long-term access to NIST research results and providing guidance to NIST researchers on publishing and research data management.

While academic libraries only recently ventured into publishing services, these services have been a major component of ISO's portfolio of services for the past 20 years. The organization's involvement in research data management is a natural extension of its role in publishing at NIST.

Developing a collaborative service model

ISO has actively nurtured customer relationships with the NIST laboratory programs over the past 18 years. Beginning with its 1998 strategic planning efforts, ISO laid the foundation for working with the NIST research community to plan, develop and deliver services that enable and add value to the dissemination of NIST research results (Office of Information Services, 1998). By 2015, ISO's strategic plan identified particular objectives and actions for collaborating and partnering with NIST researchers. This new language, specifically addressing collaboration, reflects the evolution of ISO's focus on the customer and the development of workforce competencies to advance this focus. Over the past 18 years, ISO took incremental steps to expand outreach, build on established relationships and become skilled in new areas, such as data management and data visualization. ISO has adapted to the changing research landscape by redefining its publishing and research support services and finding ways to more actively engage with the NIST researchers.

The current collaborations with ODI evolved from two components of ISO's services – the Lab Liaison Program and the publishing services integration with the public access mandates.

Lab Liaison Program

Librarians are assigned to NIST labs and programs to serve as a single point of contact with ISO. The Lab Liaisons establish and maintain close working relationships with managers and scientists in their assigned organization allowing them to provide customized services. These services include assessing the impact of the lab's work, providing guidance on the best methods for digitizing and preserving research support

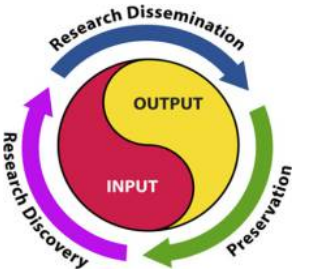

Figure 1.
The knowledge continuum

materials, compiling and analyzing industry data and conducting bibliometric studies of NIST papers. Additionally, librarians who specialize in research data management, digital services and publishing are assigned to particular laboratories to work alongside the primary Lab Liaison, to provide outreach and support in these emerging service areas. The role of the Lab Liaison has evolved from being an "ambassador" for use of the research library's collections and resources to being a "strategic research partner" contributing their expertise to scientists' and managers' research and planning activities.

Open access mandates

After the White House issued a series of memoranda and executive orders in 2013, federal agencies began to develop plans aimed at making the results of federally funded research freely available to the public and ensuring data are better managed (OSTP Memo; Executive Order; office of management and budget [OMB] Memo, 2013). NIST leadership needed to create formal responses to these directives, and they began the process by examining current publishing and data management practices of their researchers. ISO had the opportunity to play a major part in these early efforts because of its role in publishing at NIST and because staff are recognized as experts in information management. Some of these activities included:

- examining options and making recommendations to the NIST chief of staff regarding a repository to archive NIST publications;
- participating on the NIST Scientific Data Committee to provide expertise on metadata and data preservation;
- contributing to NIST's plan for providing public access to the results of federally funded research;
- contributing to NIST's data and publications policies;
- developing a metadata schema to describe NIST research data; and
- reviewing and creating metadata for NIST SRD, as a pilot process for entering metadata into the NIST enterprise data inventory (EDI).

ISO's involvement with the various committees and working groups addressing the public access issues helped to forge new relationships and enhance existing ones with members of the NIST community. Because of its knowledge of publishing, preservation and metadata, ISO was given specific responsibilities in the implementation of NIST's public access plan, including ensuring that NIST-funded publications are deposited in a publicly accessible repository and reviewing metadata entries describing NIST-generated data.

Information Services Office's strategy for developing data management services

ISO's participation on the NIST-level Scientific Data Lifecycle Management Working Group (SDLM WG), established in 2009 and which later became the NIST Scientific Data Committee (SDC), laid much of the groundwork for the future collaboration with ODI, as well as informing ISO's data management services. Membership on these two working groups provided ISO the opportunity to cultivate relationships with other organizations at NIST who also had a stake in agency-wide data management planning. ISO's

presence and active role in creating documentation covering a NIST extension to the OMB project open data metadata schema and a pilot data management plan tool were excellent ways for the organization to demonstrate expertise in the data management arena to the wider NIST community.

As ISO staff were participating in SDC activities, ISO management put together a team with members from both the research library and the DSPG to:

- raise awareness and understanding across ISO of data management issues, trends and terminology; and
- map out a role for ISO in the stewardship of NIST's research data.

The team's 2013 report to management made recommendations for short-term and long-term activities that addressed the following questions:

Q1. How is NIST responding to external drivers as they relate to research data management?

Q2. What needs are the laboratory programs articulating? What specific requests or needs have been expressed to ISO?

Q3. Are there specific groups or individuals within the NIST laboratory programs ISO should partner or collaborate with?

Q4. How is the research/special library community responding? Are there models of excellence/best practices (ISO Digital Data Working Group, 2013) to adopt or emulate?

The team's recommendations included methods for ongoing education of ISO staff, types of services ISO could offer to the NIST community and ways for marketing those services and the hiring of librarians with research data management expertise. The team's work and the actions taken by management to address the recommendations positioned the organization to jointly address data management challenges with the laboratory programs.

Jointly solving data management problems

The relationship between ISO and ODI transitioned to a new level in the summer of 2014 after ODI hired its inaugural director. Several projects with MML researchers were well underway because ODI started to take shape under the new director. Both ISO and ODI are involved in outreach to their respective constituencies related to the public access plan and other NIST-wide data management efforts, and they both have a stake in ensuring that NIST's data sets are well described, discoverable and preserved. It became clear early on that ISO and ODI could make considerable headway in addressing these issues by working together. What is unique about the newly forged relationship between ISO and ODI is that the librarians and researchers are working jointly to address open-ended problems. Each of the participants brings a particular set of skills to the table, creating a situation where the whole is more than the sum of its parts.

Examples of projects

Materials Genome Initiative Code Catalog[2]. The MGI[3] is a White House initiative aimed at speeding up the time and cost it takes to develop advanced materials (such as

strong but lightweight metals) that will be the basis of a new generation of products. NIST coordinates MGI activities across government agencies, academia and industry.

Early in 2014, the MML technical program director for MGI sought ISO's assistance with several data management issues, including an evaluation of a specific repository system (SIdora), examination of researcher workflow to recommend best methods for data capture and assisting with the development of a catalog of materials software models and codes. It was determined that the best place to start this collaboration was to develop the metadata schema for the catalog.

The work, performed primarily by ISO's metadata librarian assigned to MML, included multiple discussions with the two primary researchers working on the project to determine what types of metadata were needed for discovery of these resources. The metadata librarian also met with other NIST researchers to gather input about the types of information they would find useful in this catalog.

Some of the metadata fields used in this schema map to the Dublin Core[4], but most are specific to the description of software in general (such as codeLanguage and operatingSystem), software related to materials science (such as scale) or legal issues related to software (such as exportControls). ISO staff have used this schema to describe upwards of 75 products to populate the initial version of the catalog (Table I).

As the project progressed, additional discussions took place about the infrastructure required to make the catalog accessible inside and outside of NIST and how the catalog will be maintained over time. The ISO librarians also helped resolve technical issues, such as creating files needed by DSpace to define the new metadata fields and to upload metadata files.

New elements	Description of new elements
accessURL	URL to download the code or software itself
codeLanguage	Programming language the code was written in
contact	Contact information for questions including name and email/mailing address
cost	Free or paid?
distribution	The technical modality which states how the software is acquired. Generally, this will be a URL, but could be a DVD or other physical medium
exportControls	Any export controls that apply to this software that may restrict its distribution
implementation	Directions for running the code, including installation, dependencies, etc. The protocol to take a user from installation to execution of the code
inputsOutputs	Examples of the expected inputs and outputs of the code
landingPageURL	Homepage or landing page for code
methodAlgorithm	The general method the software takes to solve the question at hand. Select from controlled vocabulary
opSystemName	Indication of target operating system
opSystemVersion	Version of operating system indicated in OpSystemName
scale	Physical scale which the code models. Select from list
validation	Description of validation procedures used for the software
validationData	Description of data used to validate the software
verification	The verification processes used on the software
versionDate	The year the particular version of the code described by the record was published

Note: The second column offers short definitions of these new elements

Table I.
"New elements" lists elements created to address specific issues related to software and materials science

Material Measurement Laboratory data management planning tool. As part of plans for providing public access to the results of federally funded research, NIST mandated that its researchers create data management plans effective October 2014. ODI created a web-based data management planning tool called *Minerva* (after the Roman goddess of wisdom) that allows MML researchers to comply with this mandate. *Minerva* is broader than similar applications because it not only captures information related to the four areas typically required by a data management plan (descriptions of the activity, data types created, preservation and storage information and level of public access) but also captures specific information about the data products. The information related to data products is based on an expanded version of the Project Open Data metadata scheme that was developed at NIST to more accurately describe scientific data. Once *Minerva* was in alpha testing, ISO assisted with usability testing and suggested changes to the layout and language used in the tool, as well as helping to refine the frequently asked questions for the website. Figure 2 shows the hierarchy of projects and activities in *Minerva*.

Minerva went live in the spring of 2015 and, so far, has been populated with just over 200 records authored by more than 150 researchers. With this many records, the ODI staff involved in the MML data management planning tool implementation felt that it was time to take a methodical look at the information entered into *Minerva* to determine whether it was effectively serving the laboratory's needs. This examination included a look at what data sets were represented in the data management plans, what fields were not being used, whether fields were being used correctly and what information might need to be normalized, such as keywords or descriptions of instruments. This work began in the summer of 2015 led by ISO's research data librarian with support from a NIST Summer Undergraduate Research Fellow and members of ODI. Initial results of this work include a more human-readable version of reports for ODI staff and research group leaders to review, analysis of the use of various fields (such as keywords and rights) to inform potential changes in the user interface and a network graph showing that there are both a large number of people who work on individual projects and a large number who are enmeshed in tight networks of collaboration. Future planned collaborative

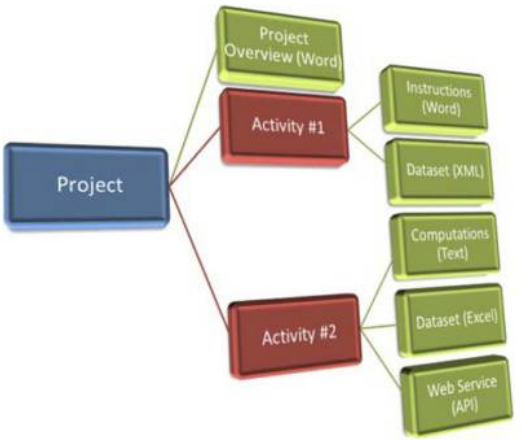

Figure 2.
Project, activity, and distribution hierarchy from the MML data management plan tool

work includes examining data management plan metadata for common uses of instrumentation and to improve sharing of sound data management practices.

Standard Reference Data impact study. ODI is in the process of reviewing NIST's SRD program and asked ISO to conduct an analysis to determine the impact of individual SRDs. SRD is defined as:

> […] quantitative information related to measurable physical or chemical property of a substance or system of substances of known composition and structure, which is critically evaluated as to its reliability (Standard Reference Data Act).

NIST develops and distributes SRD used by industry, academia and government agencies.

The purpose of this analysis was to assist ODI in the decision-making process, regarding the revamping of the NIST SRD program and to help identify which of 88 SRDs should be augmented or retired. ISO's study included both a citation analysis of the scholarly literature based on data collected from *Web of Science*[5][6] and *Google Scholar* and an analysis of US and world patents that have either cited or mentioned each SRD (Table II). ISO provided ODI with a report describing the methodology of the study, findings from the study and recommendations for additional analyses to further aid in the understanding of citing patterns and impacts. This work was a collaborative effort by ISO's MML Lab Liaison and the metadata librarian assigned to MML, with the assistance of a reference librarian from the NIST research library. The results of ISO's study were used and acknowledged by the ODI director in his presentation to the SRD program review committee. ISO and ODI will continue to jointly assess the impact of the SRDs.

Interns and research fellows. ISO and ODI have worked jointly on two applications to host a fellow/intern at NIST to work on data management activities, such as analyzing the information in the MML DMP tool, assisting with further development of the MML DMP tool and helping to lay the groundwork for a registry of materials science repositories. In writing these proposals, ISO and ODI had discussions about how an individual could be successfully mentored by more than one organization and what competencies and domain expertise each organization would contribute to the intern or fellow's experience. Although only one of the two proposals was accepted, the

NIST SRD	No. of citations
NIST Chemistry WebBook–SRD 69	100 (Google,WIPO)
JANAF Thermochemical Tables–SRD 13	83 (USPTO)
NIST Reference Fluid Thermodynamic and Transport Properties Database (REFPROP)–SRD 23	79 (WIPO)
NIST Atomic Spectra Database–SRD 78	59 (Google)
NIST/EPA/NIH Mass Spectral Library with Search Program–SRD 1a	55 (WIPO)
Biological Macromolecule Crystallization Database–SRD 21	35 (USPTO)
NIST REFLEAK: NIST Leak/Recharge Simulation Program for Refrigerant Mixtures–SRD 73	31 (Google)
X-ray Photoelectron Spectroscopy Database XPS–SRD 20	26 (Google)
FIZ/NIST Inorganic Crystal Structure Database–SRD 84	22 (WIPO)
IUPAC/NIST Solubility Data Series–SRD 106	20 (USPTO)
NIST Polycyclic Aromatic Hydrocarbon Structure Index–SRD 204	20 (WIPO)

Table II. The number of citations for SRD found in patents using three tools: Google patents, US patents and trademark database (USPTO) and World intellectual property organization patent scope (WIPO)

experience set the stage for continued conversations about challenging topics and how ODI and ISO can tackle them together.

Metadata consultation. Like the work done to support the MGI Code Catalog, metadata consultations are an effort to become involved in a project during the planning and development phases rather than at the end. In the case of metadata attached to scientific data sets, the metadata also serve an additional crucial purpose – explaining and contextualizing the data. Typically, ISO's metadata consultation has been with the development of data dictionaries that document the metadata schema used to describe the dataset or project at hand.

A data set is not particularly useful if the fields it contains are not properly described or contextualized. A contrived example would be two data sets that both contained measurements of the viscosity of a substance as it was being heated. If the person using those data sets did not realize that in the first set the temperatures were given as degrees Celsius and in the other as degrees Fahrenheit they would reach conclusions that were meaningless. One way to make sure this does not happen is for researchers to adequately describe their data via a data dictionary or metadata schema.

ISO and ODI collaborated on the development of a data dictionary in response to a NIST sponsored "app challenge". (http://nistdata.devpost.com/) The call was to create a mobile application that used least one of six SRD sets, curated by ODI, that NIST made public in machine-readable format. (All of these data sets had previously been publically available, but they were mediated by a web form, making searches easy, but making it hard to download the entire data set in one go). While some of these data sets were self-explanatory, some required better documentation for this challenge. In particular, the ISO librarians suggested ways to take the documentation of the SRD 101 – Computational Chemistry Comparison and Benchmark DataBase – which described a relational database and create documentation that was both more readily editable and which could be output in manners that were appropriate to both a machine being able to read it (for example XML or JSON) and a human being able to read it (PDF).

Ongoing conversations. While ISO and ODI have jointly worked on several projects in the last year, an essential component of the collaboration has been the methods used for listening and learning from one another. Scheduled meetings take place on a monthly basis. Although these meetings typically have no formal agenda, they allow each group to share information on the status of projects, as well as bring each other up-to-date about other interactions they might be having with internal or external contacts. Most importantly, the meetings provide a venue for a free exchange of ideas and relationship building. Informal meetings, held in researcher or librarian offices, occur when prompted by a particular question that has arisen over the course of a project. More often, the librarians and researchers engage in casual conversations when they see one another in the hallway or cafeteria. These casual conversations help to build trust and offer another opportunity to gain insights about the work carried out by both organizations.

Cost-benefit analysis

ISO and ODI have approached their collaboration with considerable self-awareness, asking each other along the way whether the benefits of collaboration outweigh the costs and/or risks. Because the ODI director had positive experiences working with librarians on issues surrounding data, publications and citation analysis in his previous position, he came to NIST with the expectation that the librarians would be similarly

interested and engaged. So it was quite natural for him to seek out interactions with the ISO. The many benefits of collaboration include the expertise librarians bring in metadata definition and curation; combining this with the knowledge of domain experts creates much stronger products than either could create alone. A number of library and information science schools are working on projects to understand how research is managed and disseminated; this information will not only influence the coursework of future information professionals but can also guide the practices of current practitioners. Librarians are directly engaged with publishers and so are in a unique position to influence the ecology of scholarly publishing. The ODI/ISO collaboration exposes to the NIST research community areas of expertise offered by the librarians that might not have been recognized otherwise.

ODI found that it could take on projects with more complex requirements because it was able to rely on ISO staff members to be a part of the solution rather than an occasional resource to draw on. ISO is planning to use the collaboration with ODI as a model for other potential partnerships elsewhere in the NIST organization.

Of course, collaboration and coordination also come at a cost. There are meetings, proposals, formal teaming agreements and personnel management issues to deal with. There is the fear, real or perceived, that any failures could lead to lack of support for future collaborative efforts. Our experience is that open and ongoing discussion of the collaboration, with clarity in terms of project participants, roles and responsibilities, successfully mitigates risks and assures that benefits warrant the costs involved.

Conclusion

ISO attributes its successful and productive relationships with NIST researchers to several factors:

- consistent strategic focus on seeking out opportunities to listen, learn and work with NIST researchers and programmatic staff;
- recruiting staff with talent and willingness to engage with researchers in a meaningful way, beyond the typical reference interview;
- creating an environment that encourages experimentation; and
- staying abreast of the technologies, trends, and issues in scholarly communication.

ISO expects to continue working with ODI over the next year, while also focusing on the data management needs of the other laboratory programs. Some initial work has involved presenting to the management of the Physical Measurement Laboratory on the NIST implementation of the public access requirements and discussing with the Engineering Laboratory the need for and development of a taxonomy to describe and connect NIST publications and data. ISO is also making forays in the area of data visualization and intends to seek out collaborative opportunities with researchers in this arena.

The collaborative model with ODI has been successful because it enabled the ISO librarians to learn more about research data management, while immersed in projects alongside the researchers. They earned the well-deserved recognition as domain experts in information management and continue to be sought out for this expertise.

ISO management is currently in discussions with the director of ODI on ways to continue the close collaboration between the two organizations in fiscal year 2016. This collaboration will help accelerate the development of ODI's data management

infrastructure while providing ISO with an opportunity to expand skills in research data management and accelerate its role in data analytics and research impact assessment.

Notes

1. The organization received the Federal Library of the Year Award from the Federal Library and Information Center Committee (FLICC) of the Library of Congress in recognition of its innovative practices and customer focus in 2003, 2008, and 2013.
2. This project started just prior to the establishment of ODI but was later moved under the auspices of this office.
3. www.mgi.gov
4. http://dublincore.org
5. Web of Science – Science Citation Index Expanded produced by Thomson Reuters.
6. The identification of any commercial product or trade name does not imply endorsement or recommendation by the National Institute of Standards and Technology.

References

ISO Digital Data Working Group (2013), ISO Digital Data Working Group Final Report, 1 November 2013.

Office of Information Services (1998), Office of Information Services Strategic Plan.

OSTP Memo; Executive Order; OMB Memo (2013), Office of Management and Budget (OMB) Memorandum for the Heads of Executive Departments Agencies of May 9, 2013, Open Data Policy – Managing Information as an Asset.

Standard Reference Data Act [P.L. 90-396; 15 USC 290-290f].

Vassallo, P. (1999), "The knowledge continuum: organizing for research and scholarly communication", *Internet Research: Electronic Networking Applications and Policy*, Vol. 9 No 3, pp. 232-242.

Website

ODI Website: www.nist.gov/mml/odi/index.cfm (accessed 7 July 2015).

Further reading

Executive Order of May 9, 2013, Making Open and Machine Readable the New Default for Government Information.

Office of Science and Technology Policy (OSTP) Memorandum for the Heads of Executive Departments and Agencies of February 22, 2013, Increasing Access to the Results of Federally Funded Scientific Research.

Corresponding author

Andrea Medina-Smith can be contacted at: andrea.medina-smith@nist.gov