Public Access at the United States Department of Energy

By Brian A. Hitson Director, Office of Scientific and Technical Information U.S. Department of Energy November 2015

On August 4, 2014, the U.S. Department of Energy (DOE) published the <u>DOE Public Access</u> <u>Plan</u>, which committed the Department to 1) deploying a web-based portal to make scholarly scientific publications resulting from DOE research funding publicly accessible and searchable at no charge to readers, and 2) instituting data management principles and requirements that apply to proposals for research funding submitted to all DOE program offices. This paper focuses on DOE's implementation of public access to scholarly publications.

Concurrent with the publication of the DOE Public Access Plan, the Department launched the DOE Public Access Gateway for Energy and Science, or <u>DOE PAGES^{Beta}</u>, which will provide free public access to accepted peer-reviewed manuscripts or published scientific journal articles within 12 months of publication. The DOE portal offering public access to scholarly scientific publications was developed and is maintained by the <u>DOE Office of Scientific and Technical Information (OSTI)</u>.

The DOE Public Access Plan and DOE PAGES^{Beta} are historic developments, the culmination of years of interagency deliberations spurred by Congress and public comments solicited by the White House Office of Science and Technology Policy (OSTP).

For nearly 70 years, OSTI has collected, preserved, and disseminated scientific and technical information (STI) emanating from research and development (R&D) by the Department of Energy and its predecessor agencies. DOE conducts more than \$11 billion a year in R&D, and OSTI helps ensure a return on those investments by making DOE-sponsored R&D results available in web-based searchable databases. These DOE databases include electronic full-text research reports; energy citations going back to the Manhattan Project era; e-prints (journal article pre-publication drafts, scholarly papers and more); DOE patents; energy science and technology software; multimedia videos about DOE and other science research; DOE non-text data collections; and DOE R&D accomplishments.

Despite the breadth of these collections, they generally have not included what is considered the "gold standard" of scientific communication – peer-reviewed journal articles or final accepted manuscripts resulting from agency funding. Until 2014, the National Institutes of Health (NIH) was the only U.S. federal science agency that offered broad public access to scientific publications resulting from its funding (as required by a law enacted in 2008); so, this gap in what OSTI offered was not unique to DOE.

OSTI designed DOE PAGES^{Beta} so that it leverages and adapts DOE's existing STI infrastructure and systems to encompass accepted manuscripts and also capitalizes on the public access achievements of publishers, laboratories, universities, and other institutions.

Background of Public Access in the United States

In 2005, the NIH announced its policy for providing public access to published articles from NIH-funded research. Three years later, the NIH public access policy was enacted into U.S. law. The policy requires all NIH-funded investigators to submit (or have submitted on their behalf) to the National Library of Medicine's PubMed Central an electronic version of their final peer-reviewed manuscripts upon acceptance for publication for posting within 12 months after they are published.

Subsequently, the America COMPETES Reauthorization Act of 2010 included a provision, Section 103, that directed the White House Office of Science and Technology Policy (OSTP) to establish a working group under the National Science and Technology Council (NSTC) "with the responsibility to coordinate Federal science agency research and policies related to the dissemination and long-term stewardship of the results of unclassified research, including digital data and peer-reviewed scholarly publications, supported wholly, or in part, by funding from Federal science agencies."

The NSTC in turn established the Task Force on Public Access to Scholarly Publications (PASP) to develop recommendations for Federal public access policies. The PASP Task Force was comprised of federal science agency representatives and, in spring 2012, produced a consensus set of objectives and recommendations. The PASP Task Force's vision was clear: "Peer-reviewed scholarly publications of the results of unclassified research supported wholly or in part by federal funding will be permanently stored and accessible to enable any individual or business to search, retrieve and analyze these publications in ways that maximize the impact and accountability of the federal research investment."

Following the task force report, OSTP continued dialogue with federal science agencies about implementing public access policies. Then, on February 22, 2013, OSTP Director Dr. John Holdren issued a memorandum to all federal agencies and department heads entitled, "Increasing Access to the Results of Federally Funded Scientific Research." The memo directed federal agencies that spend more than \$100 million a year on research and development to develop and implement plans to make peer-reviewed scientific journal articles and accepted manuscripts and scientific data in digital formats resulting from federally funded research investments publicly available in a timely fashion.

"Scientific research supported by the Federal Government catalyzes innovative breakthroughs that drive our economy," Dr. Holdren wrote in the OSTP memo. "The results of that research

become the grist for new insights and are assets for progress in areas such as health, energy, the environment, agriculture, and national security."

The memo required each agency to submit a draft public access plan to OSTP within 6 months. The agencies' plans were expected to leverage existing archives; foster partnerships with journals and publishers; optimize search, archival and dissemination features so as to promote innovation in accessibility and interoperability and to ensure long-term stewardship of federally funded research results; provide for measuring and enforcing compliance; and identify resources within the existing agency budget to carry out the plan. The OSTP memo stipulated that agencies "shall use a 12-month post-publication embargo period as a guideline for making research papers publicly available" and noted that agencies could tailor their plans and that stakeholders may petition to change the embargo period.

The OSTP memo noted the important role of publishers in peer review and promoting the quality of research and encouraged public-private collaboration. "The Administration recognizes that publishers provide valuable services, including the coordination of peer review, that are essential for ensuring the high quality and integrity of many scholarly publications," Dr. Holdren wrote. "It is critical that these services continue to be made available."

In a February 22, 2013, interagency news <u>release</u> about the OSTP policy, then DOE Office of Science Director William Brinkman endorsed the initiative. "Collaboration, transparency and open access to scientific findings accelerate discovery and innovation," Dr. Brinkman said. "The Department of Energy has been working for years with our colleagues in other science agencies and our stakeholders to advance open access. So we fully support the goals of the OSTP memorandum and will work quickly to develop and implement policies and procedures so that peer-reviewed journal articles funded by the Office of Science are available to the public."

With agency-wide responsibility for providing public access to other forms of scientific and technical information, OSTI was charged to develop and socialize DOE's public access model and system. With stakeholder input, OSTI drafted the scholarly publications section of the DOE Public Access Plan. On August 4, 2014, with the concurrence of OSTP, DOE published its final DOE Public Access Plan on the DOE open government website, effectively launching implementation of the plan and DOE PAGES^{Beta}.

The DOE Public Access Plan

The February 2013 OSTP memo was the product of two interagency task forces (one on digital data and one on scholarly publications) and broad stakeholder input, including multiple public comments periods on the White House website. The memo reflected a balance of stakeholder interests, while focusing squarely on increasing public access, and contained

detailed requirements and objectives for each agency's plan, covering both scholarly publications and digital data.

Taking note of each and every requirement and objective in the OSTP memo, DOE's goal was to develop a plan that worked for both the DOE research community and its STI infrastructure. DOE is one of the most diverse agencies in the U.S. federal government, with 17 national laboratories and thousands of research grants. Thanks to the Department's wellorganized agency program for managing scientific and technical information, DOE's public access plan was vetted and reviewed relatively quickly for an agency of its size.

The DOE Public Access Plan addresses both "Public Access to Scientific Publications" and "Public Access to Scientific Data in Digital Formats." The responsibility for implementing public access to scholarly publications rests with OSTI. The Plan specifically defines the Scope; Requirements; Applicability; Roles and Responsibilities; Planning; Implementation; Metrics, Compliance, and Evaluation; Public Consultation Experience, Public Notice; Update and Re-Evaluation of the Plan; Timeline for Implementation; and Resources – all in accordance with the requirements set forth in the OSTP memo.

How DOE PAGES^{Beta} Works

DOE PAGES^{Beta} leverages existing agency infrastructure and author submission practices, complemented by the public access efforts of publishers. Specifically, DOE-funded authors at DOE national laboratories and grantee institutions provide article metadata and either a link to the accepted manuscript on their institution's repository or the electronic accepted manuscript itself to OSTI. Using this same distributed model, for publishers who wish to provide public access through their own websites, OSTI has partnered with the publisher consortium CHORUS -- the Clearinghouse for the Open Research of the United States – to ingest CrossRef-supplied metadata for articles authored by DOE-funded researchers. By relying on its existing infrastructure and author practices, DOE's implementation of public access to full text.

DOE PAGES^{Beta} uses a "best available version" algorithm in displaying links to full text. The following graphics illustrate two paths by which users can access full text. The published article held by the publisher, also known as the "Version of Record" or VoR, is considered the "best version." When a publisher provides access to the VoR, DOE PAGES^{Beta} will only display that link, as in Example 1. When a publisher does not provide access to the VoR, DOE PAGES^{Beta} will display a link to the accepted manuscript submitted by the author to DOE, as shown in Example 2. The author may deposit the accepted manuscript on an acceptable institutional repository or to OSTI directly.



Example 1 – PAGES Query to Article on Publisher Website

Query:



Metadata in PAGES:

Beam dynamics studies for transverse electromagnetic mode type rf deflectors			
We have performed three-dimensional simulations of beam dynamics for transverse electromagnetic mode (TEM) type rf deflectors: normal and superconducting. The compact size of these cavities as compared to the conventional TM ₁₁₀ type structures is more attractive particularly at low frequency. Highly concentrated electromagnetic fields between the parallel bars provide strong electrical stability to the beam for any mechanical disturbance. An array of six 2-cell normal conducting cavities or a single cell superconducting structure is enough to produce the required vertical displacement at the target point. Both the normal and superconducting structures show very small emittance dilution due to the vertical kick of the beam.			of Record at IAB.15.022001
Authors:	Ahmed, Shahid; Krafft, Geoffrey A.; Deitrick, Kirsten; De Silva, Subashini U.; Delayen, Jean R.; Spata, Mike; Tiefenback, Michael; Hofler, Alicia; Beard, Kevin	for a way to improve these results? Let us know!	
Publication Date:	2012-02-01	CITATION FORMATS MLA APA Chicago Bibtex	
OSTI Identifier:	OSTI ID: 1074803		
Type:	Published Article		
Journal name:	Physical Review Special Topics. Accelerators and Beams		
Additional Journal Information:	Journal Volume: 15; Journal Issue: 2; Journal ID: ISSN 1098-4402	EXPORT METADATA EndNote Excel CSV XML	
Publisher:	American Physical Society (APS)		
Sponsoring Org:	USDOE		
Country of Publication:	United States	SAVE / SHARE THIS RECORD Send to Email	
Language:	English		

Article on Publisher Website:

Beam dynamics studies for transverse electromagnetic mode type rf deflectors Phys. Rev. ST Accel. Beams 15, 022001 – Published 14 February 2012 Shahid Ahmed, Geoffrey A. Krafft, Kirsten Deitrick, Subashini U. De Silva, Jean R. Delayen, Mike Spata, Michael Tiefenback, Alicia Hofler, and Kevin Beard						
PDF (free) Export Citation Citi	ng Articles (2)	C Like				
ABSTRACT AUTHORS REFERENCES	ABSTRACT We have performed three-dimensional simulations of beam dynamics for train the deflectors: normal and superconducting. The compact size of these cavities type structures is more attractive particularly at low frequency. Highly concert parallel bars provide strong electrical stability to the beam for any mechanic conducting cavities or a single cell superconducting structure is enough to pat the target point. Both the normal and superconducting structures show we vertical kick of the beam. DD: http://dx.doi.org/10.1103/PhysRevSTAB.15.022001 Image:	nsverse electromagnetic mode (TEM) type es as compared to the conventional TM_{110} ntrated electromagnetic fields between the al disturbance. An array of six 2-cell normal oroduce the required vertical displacement ery small emittance dilution due to the				

Example 2 – PAGES Query to Accepted Manuscript

Query:



Metadata in PAGES:

Top quark mas	s measurement using the template method at CDF	Citation Details		
We present a measur method. The data san collected with the CDF lepton+jets and two kii different jets-to- quark reconstructed top qua dilepton channel. The to the data yields a me	Free Publicly Accessible Full Text Accepted Manuscript (0.75 MB) Publisher's Version of Record 10.1103/PhysRevD.83.111101 Have feedback or suggestions			
Authors:	Aaltonen, T [Helsinki Inst. of Phys.]; Alvarez Gonzalez, B [Oviedo U.; Cantabria Inst. of Phys.]; Amerio, S [INFN, Padua]; Amidei, D [Michigan U.]; Anastassov, A [Northwestern U.]; Annovi, A [Frascati]; Antos, J [Comenius U.]; Apollinari, G [Fermilab]; Appel, J A [Fermilab]; Apresyan, A [Purdue U.]; Arisawa, T [Waseda U.; Dubna, JINR]	for a way to improve these results? Let us know! CITATION FORMATS		
Publication Date:	2011-06-03			
OSTI Identifier:	OSTI ID: 1016224	MLA APA Chicago Bibtex		
Report Number(s):	FERMILAB-PUB-11-205-E-PPD Journal ID: ISSN 1550-7998; arXiv eprint number arXiv:1105.0192; TRN: US1103031			
DOE Contract Number:	AC02-07CH11359	EXPORT METADATA EndNote Excel CSV XML SAVE / SHARE THIS RECORD Send to Email		
Туре:	Accepted Manuscript			
Journal name:	Physical Review. D, Particles, Fields, Gravitation and Cosmology			
Additional Journal Information:	Journal Volume: 83; Journal Issue: 11; Journal ID: ISSN 1550-7998			
Publisher:	American Physical Society (APS)			
Research Org:	Fermi National Accelerator Laboratory (FNAL), Batavia, IL			
Sponsoring Org:	USDOE Office of Science (SC)			
Country of Publication:	United States			
Language:	English			
Subject:	72 PHYSICS OF ELEMENTARY PARTICLES AND FIELDS; DECAY; FERMILAB COLLIDER DETECTOR; FERMILAB TEVATRON; LUMINOSITY; T QUARKS Experiment-HEP			

Accepted Manuscript Hosted at DOE Laboratory:

FERMILAB-PUB-11-205-E-PPD

Top quark mass measurement using the template method at CDF

T. Aaltonen,²¹ B. Álvarez González^v,⁹ S. Amerio,⁴¹ D. Amidei,³² A. Anastassov,³⁶ A. Annovi,¹⁷ J. Antos,¹² G. Apollinari,¹⁵ J.A. Appel,¹⁵ A. Apresyan,⁴⁶ T. Arisawa,⁵⁶ A. Artikov,¹³ J. Asaadi,⁵¹ W. Ashmanskas,¹⁵ B. Auerbach,⁵⁹ A. Aurisano,⁵¹ F. Azfar,⁴⁰ W. Badgett,¹⁵ A. Barbaro-Galtieri,²⁶ V.E. Barnes,⁴⁶ B.A. Barnett,²³ P. Barria^{ce},⁴⁴ P. Bartos,¹² M. Bauce^{aa},⁴¹ G. Bauer,³⁰ F. Bedeschi,⁴⁴ D. Beecher,²⁸ S. Behari,²³ G. Bellettini^{bb},⁴⁴ J. Bellinger,⁵⁸ D. Benjamin,¹⁴ A. Beretvas,¹⁵ A. Bhatti,⁴⁸ M. Binkley^{*},¹⁵ D. Bisello^{aa},⁴¹ I. Bizjak⁹⁹,²⁸ K.R. Bland,⁵ B. Blumenfeld,²³ A. Bocci,¹⁴ A. Bodek,⁴⁷ D. Bortoletto,⁴⁶ J. Boudreau,⁴⁵ A. Boveia,¹¹ B. Brau^a,¹⁵ L. Brigliadori²,⁶ A. Brisuda,¹² C. Bromberg,³³ E. Brucken,²¹ M. Bucciantonio^{bb},⁴⁴ J. Budagov,¹³ H.S. Budd,⁴⁷ S. Budd,²² K. Burkett,¹⁵ G. Busetto^{aa},⁴¹ P. Bussey,¹⁹ A. Buzatu,³¹ C. Calancha,²⁹ S. Camarda,⁴ M. Campanelli,³³ M. Campbell,³² F. Canelli¹²,¹⁵ A. Canepa,⁴³ B. Carls,²² D. Carlsmith,⁵⁸ R. Carosi,⁴⁴ S. Carrillo^k,¹⁶ S. Carron,¹⁵ B. Casal,⁹ M. Casarsa,¹⁵ A. Castro⁷,⁶ P. Catastini,¹⁵ D. Cauz,⁵² V. Cavaliere^{ce, 44} M. Cavalli-Sforza,⁴ A. Cerri⁷,²⁶ L. Cerrito⁹,²⁸ Y.C. Chen,¹ M. Chertok,⁷ G. Chiarelli,⁴⁴ G. Chlachidze,¹⁵ F. Chlebana,¹⁵ K. Cho,²⁵ D. Chokheli,¹³ J.P. Chou,²⁰ W.H. Chung,⁵⁸ Y.S. Chung,⁴⁷ C.I. Ciobanu,⁴² M.A. Ciocei^{ce, 44} A. Clark,¹⁸
G. Compostella^{aa, 41} M.E. Convery,¹⁵ J. Conway,⁷ M.Corbo,⁴² M. Cordelli,¹⁷ C.A. Cox,⁷ D.J. Cox,⁷ F. Crescioli^{bb},⁴⁴
C. Cuenca Almenar,⁵⁹ J. Cuevas⁹ R. Culbertson,¹⁵ D. Dagenhart,¹⁵ N. d'Ascenzo^t,⁴² M. Datta,¹⁵ P. de Barbaro,⁴⁷ S. De Cecco,⁴⁹ G. De Lorenzo,⁴ M. Dell'Orso^{bb},⁴⁴ C. Deluca,⁴ L. Demortier,⁴⁸ J. Deng^c,¹⁴ M. Deninno,⁶ F. Devoto,²¹ M. d'Errico^{aa, 41} A. Di Canto^{bb},⁴⁴ B. Di Ruzza,⁴⁴ J.R. Dittmann,⁵ M. D'Onofrio,²⁷ S. Donati^{bb},⁴⁴

Some publishers provide free access to the accepted manuscript but not to the Version of Record. In those cases, for publishers who actively facilitate authors' accepted manuscript deposit requirements, DOE PAGES^{Beta} will display links to both the institution-hosted and publisher-hosted accepted manuscript.

Typically, DOE PAGES^{Beta} will provide access to full text after a 12-month administrative interval.

When launched on August 4, 2014, DOE PAGES^{Beta} contained an initial collection of ~6,500 accepted manuscripts and journal articles as a demonstration of its functionality and eventual expanded content. Since the 2014 launch, both DOE and CHORUS have made progress in their respective communities, enabling public access to approximately 25 percent of in-scope accepted manuscripts or articles in the first year of implementation. In addition to working with CHORUS, DOE also participates in the academia-based consortium SHARE – the SHared Access Research Ecosystem – and anticipates that this partnership will further contribute toward comprehensiveness, particularly from DOE's grantee research community.

Other U.S. Federal Agencies' Public Access Plans

In addition to DOE, most other major federal science agencies have now issued an approved public access plan. Like DOE, some agencies are leveraging their existing STI infrastructures, while others are partnering with another agency in implementing a public access model. For example, the National Science Foundation (NSF) is partnering with DOE

in developing a PAGES-like solution that includes author submission requirements along with CHORUS and publisher participation. NASA is partnering with NIH in using the PubMed Central model.

In general, agencies are adopting a 12-month embargo period before providing free access to full text, consistent with the general guidance of the OSTP memo. Authority for agencies to disseminate accepted manuscripts resulting from federal funding emanates primarily from existing language in the U.S. Code of Federal Regulations. In the vernacular of "Green" versus "Gold" open access, agencies' implementation of public access could be viewed as a form of Green OA. That is, the responsibility for complying with public access is on the author or his/her institution because that is the entity with whom the government has a financial relationship. As a condition of receiving federal research funding, the author agrees to support the agency's public access requirement by depositing his/her accepted manuscript, either in his/her institution's repository or in a central repository. In recognition of existing publishing business models, agencies are then choosing to provide public access after a 12-month embargo period.

While some publishers are transitioning to or exploring Gold OA options for authors (i.e., where the author or his/her institution pays a fee to provide immediate free access to the article) and while federal research contracts and grants generally do not prohibit the use of research funding to pay Gold OA fees, agencies' public access plans have not endorsed Gold OA as the pathway to achieve public access. In DOE's case, cost is a major consideration for its public access model. By relying on existing infrastructure and authority, the cost of DOE's public access implementation is significantly less than any other model or open access approach.

Stakeholder Feedback

Feedback from public access stakeholders reflects the interests of their particular communities. Some publishers have advocated for a variable embargo period based on usage and citation patterns and differences across various scientific disciplines. Some open science advocates have argued for broader reuse rights in agency policies and shorter embargo periods.

The OSTP memo encouraged public-private collaboration to explore the potential of creative reuse opportunities, such as text and data mining applications. Under its programmatic and technology committees, ICSTI will hold two workshops on text and data mining in Philadelphia, Pennsylvania, on February 20, 2016. These workshops will serve to demonstrate the state of text and data mining capabilities, as well as any current shortcomings and limitations.

What's Next

As U.S. federal science agencies deploy their public access plans, it is reasonable to anticipate a consistently upward trajectory in the number and percentage of peer-reviewed articles resulting from U.S. federal research funding that will become free and publicly accessible. As this corpus grows, so will opportunities for virtually aggregating and linking individual agency collections and associated supplementary information and research data. Specifically, the federated search engine Science.gov, which already searches over 200 million pages of U.S. federal science agencies' databases and websites, is poised to become the "go to" search engine for all federal public access content. And, internationally, WorldWideScience.org, a federated search engine covering scientific databases across 70+ countries, is poised to integrate this newly-accessible corpus and extend its reach and impact globally. (OSTI is the Operating Agent for both Science.gov and WorldWideScience.org, and ICSTI is a primary sponsor of WorldWideScience.org.)

Researchers and the public also increasingly want and need seamless access to datasets and supplemental material associated with research publications. This serves to support research reproducibility as well as text and data mining applications. It is also reasonable to anticipate that agency public access tools will increasingly provide linkages between publications and their related data, software, images, multimedia, and other supplemental material.