



**U.S. FOOD & DRUG
ADMINISTRATION**

DIGITAL TRANSFORMATION SYMPOSIUM

2023

Hosted by FDA's Office of Digital Transformation



Sharing Data Through the FDA Expertise and Research Portal

Ben Hope, Deputy Director, FDA Library, Office of Digital Transformation (ODT)

Alexis Norris, Chair, FDA Omics Working Group and Bioinformatician, Center for Veterinary Medicine (CVM)



Search the repository FDA Staff ▾ All ▾ 🔍

FDA Expertise and Research Portal

The FDA Expertise and Research Portal preserves and provides access to the research, expertise, and creative output of the FDA's staff.

[FDA Profiles](#) | [Output](#) | [FDA Organizations](#)

Our Researchers



JASON GORMAN

Principal Investigator



MAYUMI F MILLER

Veterinary Research Branch



ZHICHAO LIU

Principal Investigator



What is the Portal and why did we build it?

The Portal is a centralized, automated tool for finding FDA experts and linking to their research and data to support knowledge sharing and collaboration.

The FDA lacks a comprehensive, centralized, and automated source of researcher profiles and expertise data representing the collective research output across the Centers. Profiles are stored in multiple locations, manually created, and are often out-of-date or incomplete. FDA publications, data, and additional scientific and research outputs are similarly dispersed and not easily searchable.



How does the Portal align with the FDA IT Strategy?

The Portal directly aligns with the new FDA IT strategic goals:

1. Create a Shared OneFDA Ecosystem
2. Strengthen IT Infrastructure
3. Modernize Enterprise Services and Capabilities
4. Share Data for Mission Outcomes
5. Adopt AI and Mission-Driven Innovations
6. Cultivate Talent and Leadership



How does it work?

- Based on a commercial off-the-shelf (COTS) solution called **Esploro**, the Portal uses artificial intelligence (AI) to harvest data from trusted internal and external sources.
- Users can edit their profiles as needed, but they no longer need to update their research outputs manually.
- Persistent digital identifiers, such as the Open Researcher and Contributor ID ([ORCID](#)) and Digital Object Identifier ([DOI](#)) are used to increase data quality and share data with other systems.
- Planned integrations with systems like [openFDA](#) will connect researchers and publications to datasets, code, and additional scientific and research outputs.



How does the FDA benefit?

- AI harvesting from trusted internal and external sources helps minimize the data entry burden on FDA researchers.
- FDA staff has access to expertise across the Centers and to external collaborators.
- FDA has a searchable catalog of research and scientific assets.
- Data transparency enables compliance with federal requirements and laws, such as Open Science, HR 4174, and M-13-13.



How would the Public benefit?

- The Portal would be publicly accessible and would provide free and immediate access to FDA peer-reviewed publications.
- Faster access to expertise internally would contribute to better regulatory, scientific, and research decisions.
- Data approved for public access would be more readily searchable and accessible with openFDA integration.
- Additional research and scientific outputs would be available to the public.



Sharing Data – Expertise

- Expertise can be searched by keyword and refined using filters.
- Collaborators can be identified more readily.

The screenshot shows a search results page for the keyword "artificial intelligence". The search bar at the top includes a search icon, a dropdown menu for "Outputs", and a link to "Advanced Search".

Refine the results

Sort by Date-newest

Availability

- Peer-reviewed Journals 47
- Open Access 31

Resource Type

- Journal article 51
- Conference presentation 4
- Conference proceeding 2
- Book chapter 1
- Book 1

[Show more](#)

Research unit

- Division of Bioinformatics and Biostatistics 22
- CDRH - Center for Devices and Radiological Health 9
- Division of Regulatory Review and Research 6
- NCTR Contractor Special Employee Program 6
- Division of Imaging, Diagnostics and Software Reliability 5

[Show more](#)

Author/Creator

- Weida Tong 18
- Zhichao Liu 14
- Ruth Roberts 8
- Shraddha Thakkar 7

Journal Article | PEER REVIEWED

Classifying Free Texts Into Predefined Sections Using AI in Regulatory Documents: A Case Study with Drug Labeling Documents
 by Magnus Gray, Joshua Xu, Weida Tong and Leihong Wu
 Chemical research in toxicology
 07/24/2023
 DOI <https://doi.org/10.1021/acs.chemrestox.3c00028>; PMID 37487037
 The US Food and Drug Administration (FDA) regulatory process often involves several reviewers who focus on sets of information related to their respective areas of review. Accordingly, manufacturers that provide submission packages to regulatory agencies are instructed to organize the contents using a structure that enables the information to be easily allocated, retrieved, and...

Journal Article | OPEN ACCESS | PEER REVIEWED

Development of benchmark datasets for text mining and sentiment analysis to accelerate regulatory literature review
 by Leihong Wu, Si Chen, Lei Guo, Svitlana Shpyleva, Kelly Harris, Tariq Fahmi, Timothy Flanigan, Weida Tong, Joshua Xu and Zhen Ren
 Regulatory toxicology and pharmacology, Vol.137, pp.105287-105287
 01/2023
 DOI <https://doi.org/10.1016/j.yrtph.2022.105287>; PMID 36372266
 In the field of regulatory science, reviewing literature is an essential and important step, which most of the time is conducted by manually reading hundreds of articles. Although this process is highly time-consuming and labor-intensive, most output of this process is not well transformed into machine-readable format. The limited availability of data has largely constrained...

Journal Article | OPEN ACCESS | PEER REVIEWED

DeepCausality: A general AI-powered causal inference framework for free text: A case study of LiverTox
 by Xingqiao Wang, Xiaowei Xu, Weida Tong, Qi Liu and Zhichao Liu
 Frontiers in artificial intelligence, Vol.5, pp.999289-999289
 12/06/2022
 DOI <https://doi.org/10.3389/frai.2022.999289>; PMID 36561659
 Causality plays an essential role in multiple scientific disciplines, including the social, behavioral, and biological sciences and portions of statistics and artificial intelligence. Manual-based causality assessment from a large number of free text-based documents is very time-consuming, labor-intensive, and sometimes even impractical. Herein, we proposed a general causal inference framework...

Journal Article | OPEN ACCESS | PEER REVIEWED

DeepCarc: Deep learning-powered carcinogenicity prediction using model-level representation (vol 4, 757780, 2021)
 by Ting Li, Weida Tong, Ruth Roberts, Zhichao Liu and Shraddha Thakkar
 Frontiers in artificial intelligence, Vol.5
 11/28/2022
 DOI <https://doi.org/10.3389/frai.2022.1046668>; PMID 36518910

Sharing Data – Expertise



Alexis Norris

Bioinformatician, CVM

*Genomics | Bioinformatics |
Computer Programming |
Computational Biology |
Genetics | Animal Genetics or
Breeding | Transgenic
Animals | Genome Editing |
precisionFDA | High
Performance Computing |
Cloud Computing | Omics |
Transcriptomics*

Expertise

Bioinformatics reviewer (CVM) and current Chair of FDA Omics Working Group. Experience with bioinformatics analysis of data generated to characterize genome edited animals. Data is typically next generation sequencing (NGS) -- Illumina, Nanopore, PacBio. Analysis is primarily focused on variant calling, de novo genome assembly, and gene expression.

Links

[FDA Omics Working Group](#)

[GitLab \(FDA\)](#)

[Google Scholar
Publications](#)

[LinkedIn](#)

Organizational Affiliations

Division of Animal Bioengineering and Cellular Therapies, Office of New Animal Drug Evaluation, CVM - Center for Veterinary Medicine, FDA



Sharing Data – Publications



Overview **Output**

Highlights (1) [+ ADD WORK](#) [EXPORT ALL](#)

All assets (35)

- Publication (32)
- Posted Content (2)
- Conference/Event (1)

JOURNAL ARTICLE | OPEN ACCESS | PEER REVIEWED

Template plasmid integration in germline genome-edited cattle

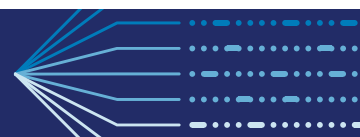
by Alexis L Norris, Stella S Lee, Kevin J Greenlees, Daniel A Tadesse, Mayumi F Miller and Heather A Lombardi

Published 02/2020

Nature biotechnology, 38, 2, 163 - 164



Alexis Norris



Sharing Data – Datasets

[← Back](#) | JOURNAL ARTICLE |  OPEN ACCESS |  PEER REVIEWED

Template plasmid integration in germline genome-edited cattle

Alexis L Norris, Stella S Lee, Kevin J Greenlees, Daniel A Tadesse, Mayumi F Miller and Heather A Lombardi [Show details for 6 authors](#)

Nature biotechnology, Vol.38(2), pp.163-164
02/2020
DOI: <https://doi.org/10.1038/s41587-019-0394-6>
PMID: 32034391

[Share](#) [Export](#)

Abstract

[Files and links \(1\)](#)

Metrics

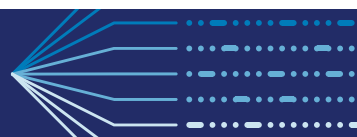
Details

Abstract

[Animals](#) / [Cattle - genetics](#) / [Genome](#) / [Germ Cells - metabolism](#) / [Mutagenesis, Insertional - genetics](#) / [Plasmids - genetics](#) / [Templates, Genetic](#)

Files and links (1)

	Genome Sequencing Data View
Data	The Carlson et al. whole-genome sequencing data are available from NCBI SRA
Open	



Benefits to FDA staff (user perspective)

- Find people (internal subject matter experts)
 - Research: potential collaborations
 - Review: potential consults
- Find communities (e.g., working groups)
- Find datasets
 - Leverage existing datasets to answer regulatory questions as they arise
- Find software/tools
 - Especially those that already exist internally (e.g., FDA GitLab)



Benefits to FDA staff (user perspective)

- Transparency to scientific community and external stakeholders (e.g., industry, trade groups, consumer advocacy groups)
 - Awareness of publications
 - Encourage reuse of FDA-generated datasets
 - Sharing FDA-developed tools/code



Looking ahead

- Integrations with internal and external trusted systems
- Collaborations with FDA Centers, publishers, and vendors
- Additional outputs minted with DOIs
- Open Science compliance
- Reporting and Analytics





Questions?