МОНГОЛ УЛСЫН ЗАСГИЙН ГАЗАР
ЭРҮҮЛ МЭНДИЙН ЯАМ

МОНГОЛ УЛСЫН ЗАСГИЙН ГАЗАР
ЦАХИМ ХӨГЖИЛ, ХАРИЛЦАА ХОЛБООНЫ ЯАМ

unicef
НҮБ-ын Хүүхдийн Сан

ЦАХИМ ЭРҮҮЛ МЭНД

Диксон Памела Грахам

Дэлхийн Эрүүл Мэндийн Байгууллагын
Эрүүл мэндийн хамтын ажиллагааны
зөвлөлийн шинжээч,
Америкийн Нэгдсэн Улсын
Карнеги Меллон Их сургуулийн зөвлөх

Updates on Big Data, AI, and Digital Health: Use Cases and New and Emerging Regulatory Guidance

E-Health Convergences, Ministry of Health, Government of Mongolia
April 2024

Pam Dixon Executive Director
World Privacy Forum | www.worldprivacyforum.org
pdixon@worldprivacyforum.org
@privacyforum
@pamdi

Stucture of Presentation

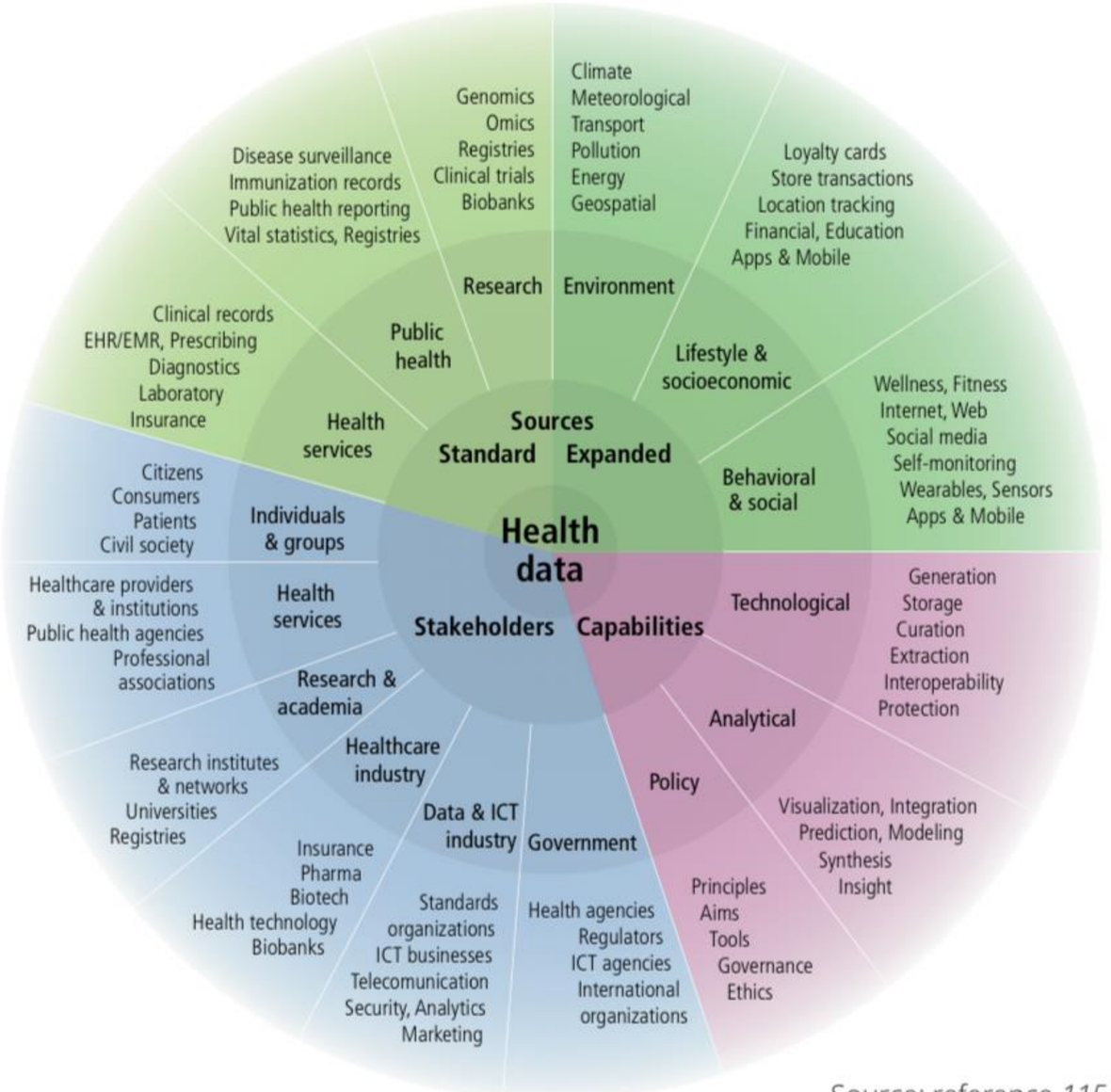
I. AI-enabled research and precision medicine

II. Big Data & AI as a medical device; new work in virtual Big Data; evolving predictive analytic techniques using clinical data

III. Key regulatory and governance concerns and solutions

Big Data and AI in Digital Health: AI and Precision Medicine

RESEARCH AND AI - ENABLED PRECISION MEDICINE



Source: reference 115

Health data sources can be standard or expanded. Now, some private sector data, or even virtual data based on aggregate clinical records is also being brought in for use in AI methods in health.

WHO: The health data ecosystem

TEMPUS

Improving patient care through high quality testing, clinical trial matching, and deep research data that powers scientific discovery

50%+
of all Academic Medical Centers in the US are connected to Tempus

50%+
of oncologists in the US connected to Tempus through sequencing, clinical trial matching, and research-enabled partnerships

90%
of the top 20 pharma oncology companies partner with Tempus

200+
biopharma partnerships

~6,000,000
de-identified research records to power scientific discovery to improve patient outcomes

26,000+
patients have been identified for potential enrollment into clinical trials in our network

Research and AI- Enabled Precision Medicine, Two Key Areas:

- **Large libraries of clinical data and molecular data are being built using existing patient information.** The clinical data is **de-identified**, and the AI is used to create **advanced abstraction** from electronic health records. NLP methods have advanced, and are being used to get more information out of clinical records.
- Techniques that combine comparing large molecular data libraries with large clinical data repositories, are driving advances.
- **Tempus AI is a specific use case in this area.**
  - https://www.tempus.com **(Example)**
- KEY: This combination of Big Clinical Data and AI techniques in analyzing the data and molecules together is considered to be a **medical device,** as defined by the FDA. These devices are regulated and are listed on the FDA website.
- Part 3 of this presentation discusses how this regulation works. Note: this use case example was chosen because it is clear, and it has received FDA approval. Hundreds of these types of use cases now exist.

Research and AI- Enabled Precision Medicine, Citations:

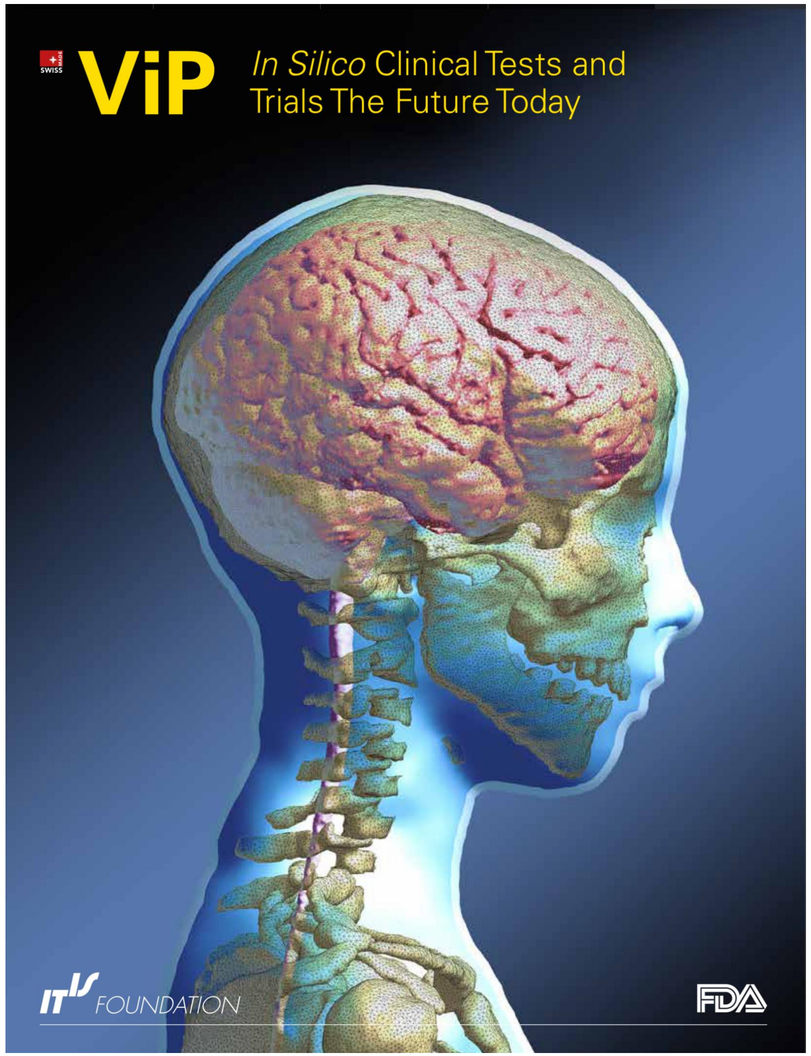
- Additional information on this use case:

- *Large Scale DNA Alteration Prediction from H&E Images Across Different Cancer Types Generalizing to External Images*, USCAP 2024

- Xiaoyong Fu, Jordan Kardos, Paola Correa, Min Wang, Li, Jackson Egen, Shahed Iqbal. *Molecular profiling of advanced non-small cell lung cancer in response to first-line immune checkpoint inhibitors and/or chemotherapy using multimodal real-world data [abstract]*. In: Proceedings of the American Association for Cancer Research Annual Meeting 2024; Part 1 (Regular Abstracts); 2024 Apr 5-10; San Diego, CA. Philadelphia (PA): AACR; Cancer Res 2024;84(6_Suppl):Abstract nr 6459.

Big Data and AI in Digital Health: Predictive Analytics

PREDICTIVE ANALYTICS USING HEALTH DATA AT SCALE: virtual data, clinical data, and more



Virtual Big Data and AI:

One major category of Big Data involves AI / Predictive Analytics Performed on De-identified Clinical Data.

There is also a new category of Big Data: **virtualized data sets extracted from human research**. These are started with detailed MRI scans and reconstructed as computer-aided design objects.

This allows for new and very flexible and accurate health datasets that can be explanded to be used for population health research.

← Resources for You (Medical Devices)

The Virtual Family: A Set of Anatomically Correct Whole-body Computational Models

Share | Post | Linkedin | Email | Print

Catalog of Regulatory Science Tools to Help Assess New Medical Devices

Technical Description

The Virtual Family (VF) is a set of four highly detailed, anatomically correct whole-body models of an adult male, an adult female, and two children [1].

The four VF models are based on high-resolution magnetic resonance imaging (MRI) data of healthy volunteers. Organs and tissues of the VF 2.0 models are represented by three-dimensional, highly detailed computer-aided design (CAD) objects without self-intersections and gaps. The CAD objects allow the models to be meshed at arbitrary resolutions without loss of small features.

The following two VF model versions are available:

- VF 1.0: The VF 1.0 models include segmentation of approximately 80 high-resolution organs and tissues. The Virtual Family Tool can be used to discretize and export the CAD objects in a generic voxel-based format. All four VF 1.0 models and the Virtual Family Tool are provided free of charge (except for handling fees) to the scientific community for academic purposes only.
- VF 2.0: The VF 2.0 models consist of simplified CAD files optimized for finite-element modeling in any third-party platform. These models are based on a new, high-end generation of VF models that have been re-segmented at finer resolution to afford a higher degree of precision and anatomical refinement, as well as improved structural continuity of approximately 300 organs and tissues. For the purpose of simplification, these structures are combined into 22 high-resolution tissues. The VF 2.0 models are available free of charge (except for handling fees) to everyone.

*Virtual Family models (Age, Sex, Height, Weight, BMI)

The Virtual Family and Virtual Population Big Data Sets and AI research

The **Virtual Population (ViP)** consists of 15 high-resolution, full-body anatomical human models and three pregnant woman models. The models were developed from high-resolution magnetic resonance imaging (MRI) data of healthy volunteers and reconstructed as three-dimensional computer-aided-design (CAD) objects. The CAD format allows the models to be meshed at arbitrary resolutions without any loss of detail or small features. This allows extensive AI research using new kinds of synthetic data at the population level.

The ViP project began in 2005 with the development of **the Virtual Family**, a joint project between the IT'IS Foundation and the US Food and Drug Administration. Additional models were gradually generated to broaden the population coverage, forming ViP v1.. The v2.0 models, consisting of 22 simplified CAD files, were developed to support finite-element modeling in third-party commercially available platforms.

The newest generation VIP3.0 models, available since mid-2015**, elevate computational simulations in 3D anatomies to an unprecedented level of detail and accuracy, with more than 300 tissues and organs per model**, a resolution of 0.5 mm' throughout the entire body, and specific physical, physiological, and biological properties for all segmented tissues. (From US FDA and IT'IS Foundation.)

Predictive Analytics Performed on Clinical Data at Scale: The use of NLP Software and advanced Virtual Datasets

- Natural Language Programming software has changed clinical health data analysis, especially with the addition of Large Language Models.

- Example: Today's NLP software can pull **100 percent of data from clinical charts.** In the past, medical charts have been hard to analyze at scale because data is often hidden in the chart for various structuring and technical reasons. In the US use cases, NLP may be used to ensure quality control as well as for determining insurance risk.

- A landmark **Nature Digital Medicine** article describes how NLP combined with an AI Transformer (Large Language Model) took 90 billion words from de-identified clinical texts and used them for analysis. The article has been made available publicly free of charge See: https://www.nature.com/articles/s41746-022-00742-2.pdf

npj | digital medicine

www.nature.com/npjdigitalmed

ARTICLE OPEN

Check for updates

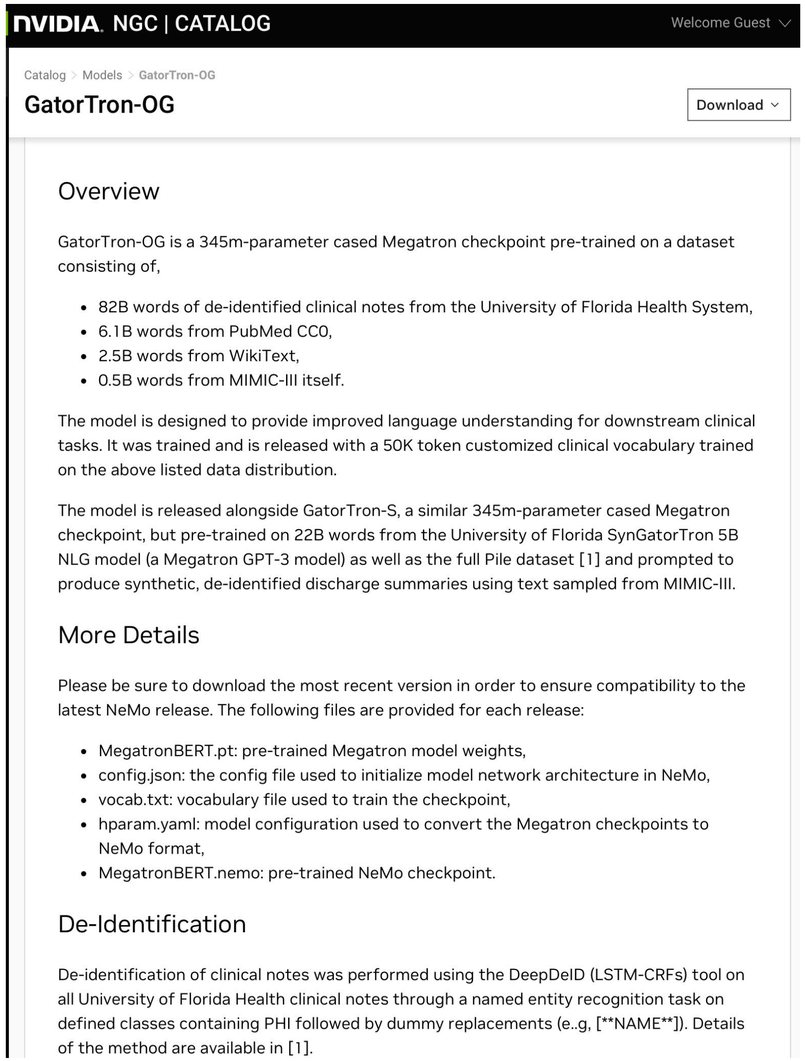
A large language model for electronic health records

Xi Yang[1,2], Aokun Chen[1,2], Nima PourNejatian[3], Hoo Chang Shin[3], Kaleb E. Smith[3], Christopher Parisien[3], Colin Compas[3], Cheryl Martin[3], Anthony B. Costa[3], Mona G. Flores[3], Ying Zhang[4], Tanja Magoc[5], Christopher A. Harle[1,5], Gloria Lipori[5,6], Duane A. Mitchell[6], William R. Hogan[1], Elizabeth A. Shenkman[1], Jiang Bian[1,2] and Yonghui Wu[1,2]✉

There is an increasing interest in developing artificial intelligence (AI) systems to process and interpret electronic health records (EHRs). Natural language processing (NLP) powered by pretrained language models is the key technology for medical AI systems utilizing clinical narratives. However, there are few clinical language models, the largest of which trained in the clinical domain is comparatively small at 110 million parameters (compared with billions of parameters in the general domain). It is not clear how large clinical language models with billions of parameters can help medical AI systems utilize unstructured EHRs. In this study, we develop from scratch a large clinical language model—GatorTron—using >90 billion words of text (including >82 billion words of de-identified clinical text) and systematically evaluate it on five clinical NLP tasks including clinical concept extraction, medical relation extraction, semantic textual similarity, natural language inference (NLI), and medical question answering (MQA). We examine how (1) scaling up the number of parameters and (2) scaling up the size of the training data could benefit these NLP tasks. GatorTron models scale up the clinical language model from 110 million to 8.9 billion parameters and improve five clinical NLP tasks (e.g., 9.6% and 9.5% improvement in accuracy for NLI and MQA), which can be applied to medical AI systems to improve healthcare delivery. The GatorTron models are publicly available at: https://catalog.ngc.nvidia.com/orgs/nvidia/teams/clara/models/gatortron_og.

npj Digital Medicine (2022)5:194 ; https://doi.org/10.1038/s41746-022-00742-2

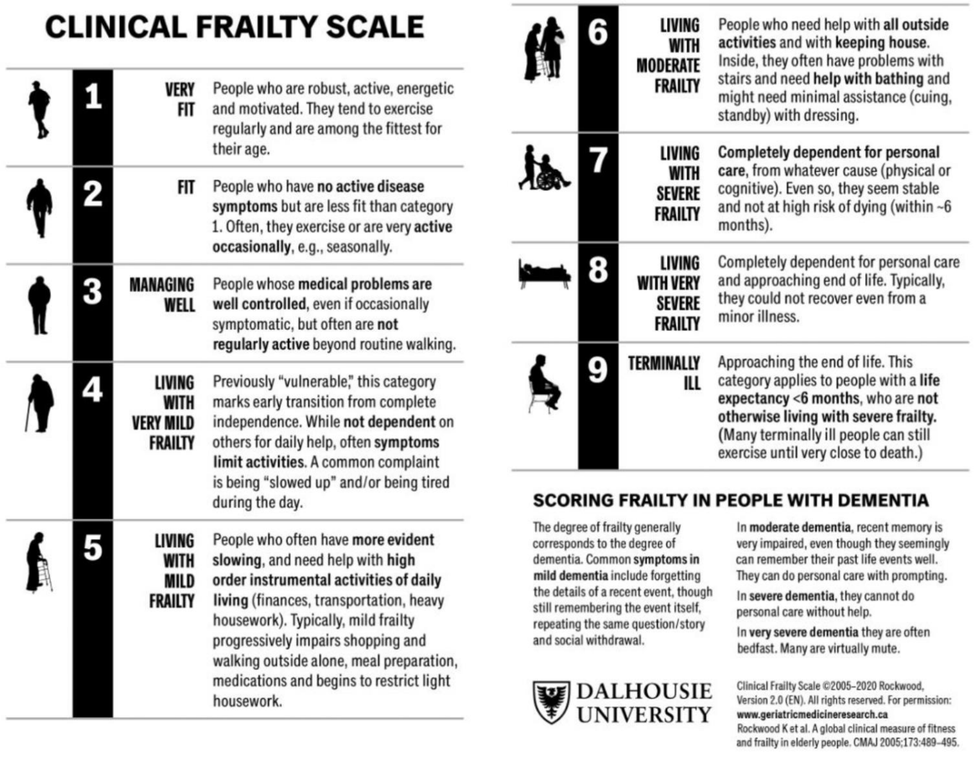
https://www.nature.com/articles/s41746-022-00742-2.pdf



NLP/ Large Language Model Dataset:

- https://catalog.ngc.nvidia.com/orgs/nvidia/teams/clara/models/gatortron_og .
- This model is trained on billions of de-identified clinical record data points.
- This model is available free of charge.

AI Use Case with clinical data: Patient Scoring or Health Scoring using de-identified and/or identified health data



The **Clinical Frailty Scale (CFS)** was introduced in the second clinical examination of the Canadian Study of Health and Aging (CSHA) as a way to summarize the overall level of fitness or frailty of an older adult after they had been evaluated by an experienced clinician (Rockwood *et al.*, 2005).

- This is not a new use case, in fact, this is classical Big Data analytics in health care. What is different is that techniques have improved.

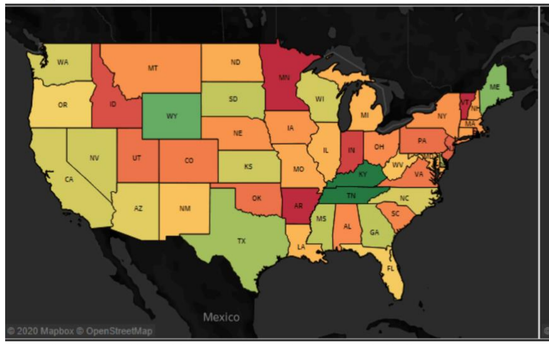
- Key example: **Frailty Score, sometimes called the Frailty Scale** This score predicts how sick a patient is, and can predict the approximate death risk of a patient. There are now many versions of this score. Usually it uses hospital-wide clinical data as a base, some scales use larger data sets

- See: Priya Mendiratta et al, *Clinical Frailty Scale, National Library of Medicine*, 2023. https://www.ncbi.nlm.nih.gov/books/NBK559009/

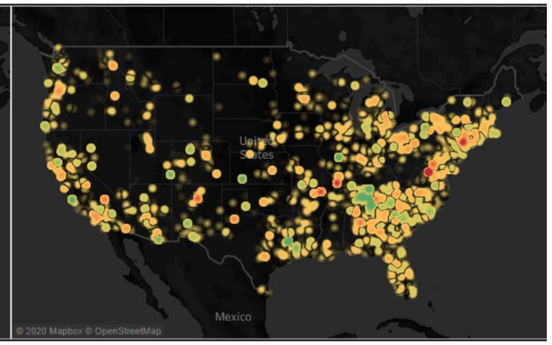
verizon Date Range Covid-19 Impact on Call Volume (Voice, Data Connectivity & Device Drivers) contact analytics

	Metric 1	Metric 2	Metric 3	Metric 4	Metric 5	Metric 6
Post Covid Volumes	###	###	###	###	###	###
Pre Covid Median	###	###	###	###	###	###
Vol Difference from Median	###	###	###	###	###	###
% Difference from baseline	###	###	###	###	###	###

State geo heatmap by WoW call presurre. (Click on State to filter for city & zip)



Zip code density map by WoW call pressure (factoring zip code population)



Top cities experiencing call pressure (factoring city population)

		City Population	Post Covid Volumes	Pre Covid Median	Vol Difference from Median	City Call Pressure
City 1	State A	###	###	###	###	###
City 2	State B	###	###	###	###	###
City 3	State C	###	###	###	###	###
City 4	State D	###	###	###	###	###
City 5	State D	###	###	###	###	###
City 6	State E	###	###	###	###	###
City 7	State F	###	###	###	###	###
City 8	State G	###	###	###	###	###
City 9	State H	###	###	###	###	###
City 10	State I	###	###	###	###	###

Top zip codes experiencing call pressure (factoring zip code population)

	Zip Population	Post Covid Volumes	Pre Covid Median	Vol Difference from Median	Zip Call Pressure
Zip 1	###	###	###	###	###
Zip 2	###	###	###	###	###
Zip 3	###	###	###	###	###
Zip 4	###	###	###	###	###
Zip 5	###	###	###	###	###
Zip 6	###	###	###	###	###
Zip 7	###	###	###	###	###
Zip 8	###	###	###	###	###
Zip 9	###	###	###	###	###
Zip 10	###	###	###	###	###

Use Case: Predictive Analytics for health and the use of Private Sector Data or Commercial Data

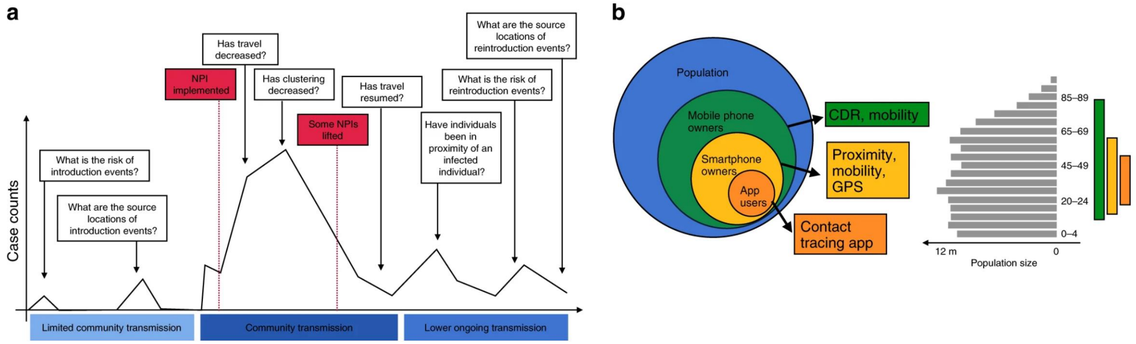
- Big Data Analytics for health purposes took a **historic leap forward during the pandemic when private sector data was used for health data research.**

- Example: Verizon and other telecommunication companies released **aggregate mobile phone location and/or mobile usage data** to track movement during the pandemic. This was revolutionary then, now it has become a more common practice. Before the pandemic, this was not an allowable use of private sector data.

- See Artur Strzelecki, *Apple Mobility Trends Data in Human Mobility Patterns during Restictions and Prediction of COVID-19,* National Library of Medicine, 2022. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9778143/ .

Fig. 1: The uses of mobile phone data to inform COVID-19 public health response and their possible biases.

From: The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology



a Over the course of the epidemic, mobile phone data and applications may be relevant to help answer a number of important epidemiological questions needed to guide the implementation and evaluation of various interventions. **b** However, these data should be considered in light of ownership and use biases that may or may not limit generalizability to the overall population. Mobile phone owners and users only represent a subset of the population and may have additional age (shown here for a synthetic population for illustrative purposes), socio-demographic, or geographic biases. Applications that require the use of a smartphone or application may further limit the generalizability of these data since they represent smaller subsets of the user population.

Predictive Analytics performed on private sector Big Data can present nuanced challenges for health uses of the data

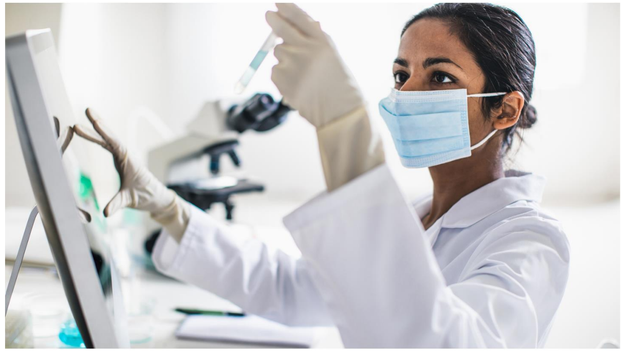
- Big data is very complex in all contexts, but especially health.
- The bias and errors that can happen in AI systems are often present in health contexts and can impact quality of data for use in clinical decision-making.
- See: Nature research publication detailing bias and other problems in the pandemic mobile geolocation data sets. Private sector Big Data provides many nuanced challenges which this article details.
- See: Ryra H. Grantz et al. *The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology*, Nature Communications, 2020. https://www.nature.com/articles/s41467-020-18190-5 Open Access.

Big Data and AI Governance in Digital Health

BIG DATA and AI GOVERNANCE in the HEALTH CONTEXT: EMERGING METHODS AND GUIDANCE

Catalog of Regulatory Science Tools to Help Assess New Medical Devices

Share | Post | Linkedin | Email | Print



Update: January 12, 2024: The FDA is providing an updated Regulatory Science Tool (RST) Catalog that will provide additional tool search capability allow for additional capacity as the catalog continues to grow.

About the catalog of regulatory science tools

The Catalog of Regulatory Science Tools provides a peer-reviewed resource for medical device companies to use where standards and qualified Medical Device Development Tools (MDDTs) do not yet exist. These tools do not replace FDA-recognized standards or MDDTs. This catalog collates a variety of regulatory science tools that the FDA's Center for Devices and Radiological Health's (CDRH) Office of Science and Engineering Labs (OSEL) developed. These tools use the most innovative science to support medical device development and patient access to safe and effective medical devices. If you are considering using a tool from this catalog in your marketing submissions, note that these tools have not been qualified as Medical Device Development Tools and the FDA has not evaluated the suitability of these tools within any specific context of use. You may request feedback or meetings for medical device submissions as part of the Q-Submission Program.

For more information about the Catalog of Regulatory Science Tools, email OSEL_CDRH@fda.hhs.gov.

Health Big Data and AI is Regulated as a “Breakthrough Medical Device” by the US Food and Drug Administration, FDA.

- A database of all Breakthrough Medical Devices is kept and made available to the public with documentation. Additional information on the FDA database as AI and Big Data as “Medical Devices”:

- FDA **Breakthrough Medical Devices Database (AI)**: https://www.fda.gov/medical-devices/how-study-and-market-your-device/breakthrough-devices-program#s1

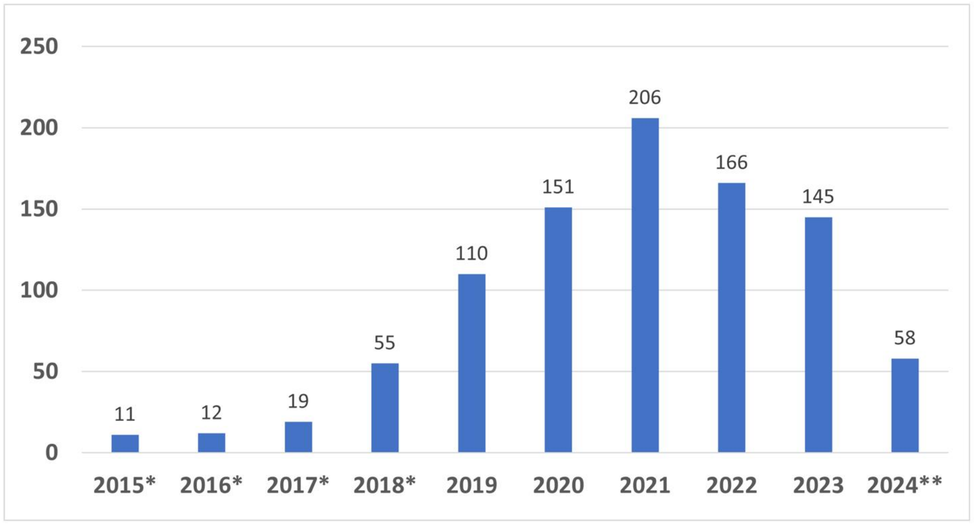
FDA Regulation of Big Data and AI as a Breakthrough Medical Device, continued:

Criteria	Description
First Criterion	The device provides for more effective treatment or diagnosis of lif
Second Criterion	The device also meets **at least one** of the following:
	a. Represents Breakthrough Technology
	b. No Approved or Cleared Alternatives Exist
	c. Offers Significant Advantages over Existing Approved or Cle
	d. Device Availability is in the Best Interest of Patients

- Criteria for being a Breakthrough Medical Device (AI + Big Data):
  - The device must provide **effective treatment**; and must also meet at least one of four additional criteria (see image to left)
  - AI tools in healthcare are “breakthrough technology” and are now part of this program and have formal guidance applied to them.

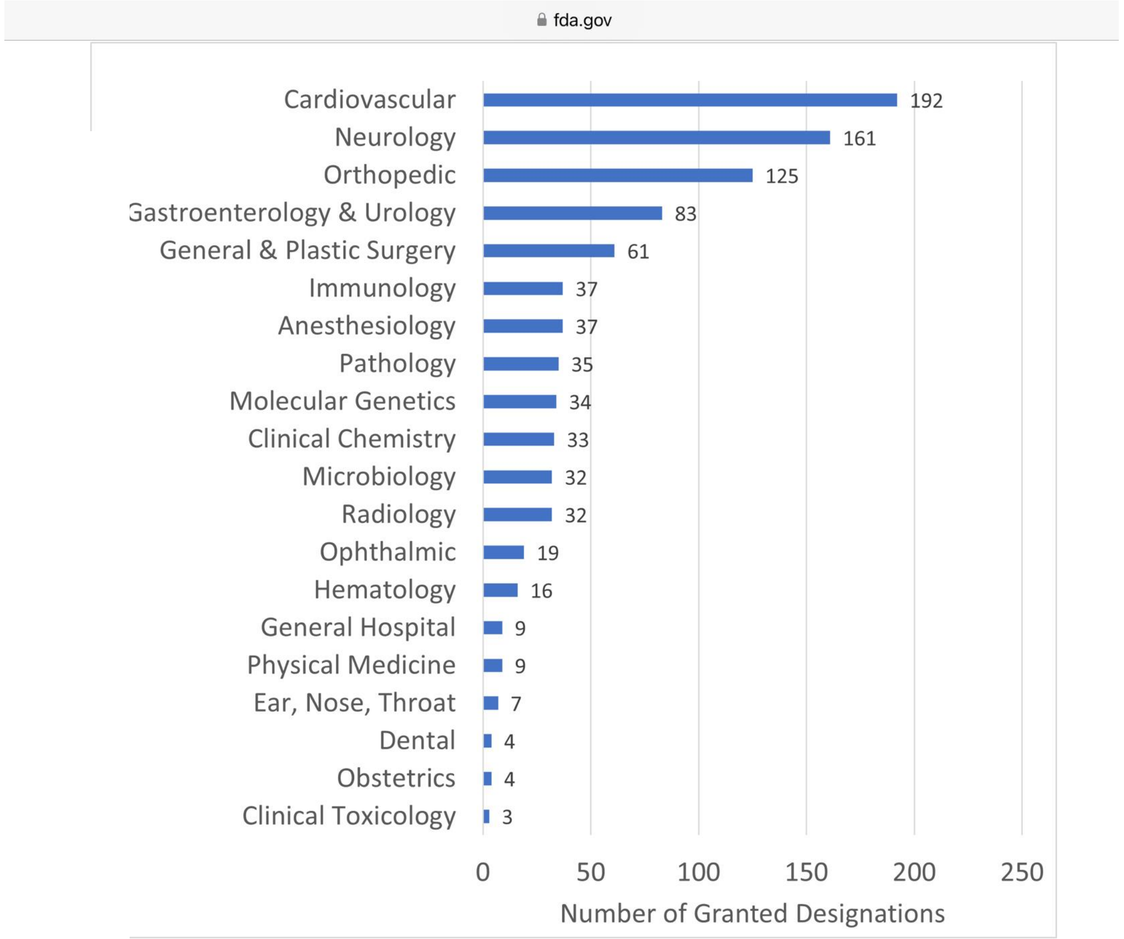
How many Breakthrough tools, and in what areas?

Graph 1: Number of Granted Breakthrough Device Designations by Fiscal Year



*Data includes devices that were designated under the precursor Expedited Access Pathway (EAP). Since the vision and designation criteria between the precursor EAP Program and the Breakthrough Devices Program are consistent, the FDA considers devices granted designation under the EAP to be a part of the Breakthrough Devices Program.

**Indicates the 2024 data are from October 1, 2023 through December 31, 2023.



ow many Breakthrough Devices have received arketing authorization?

Examples:

Breakthrough Devices program: FDA, https://www.fda.gov/medical-devices/how-study-and-market-your-device/breakthrough-devices-program#s1

1:06 PM Sun Apr 7 — fda.gov — 65%

CDRH and CBER Breakthrough Device Marketing Authorizations

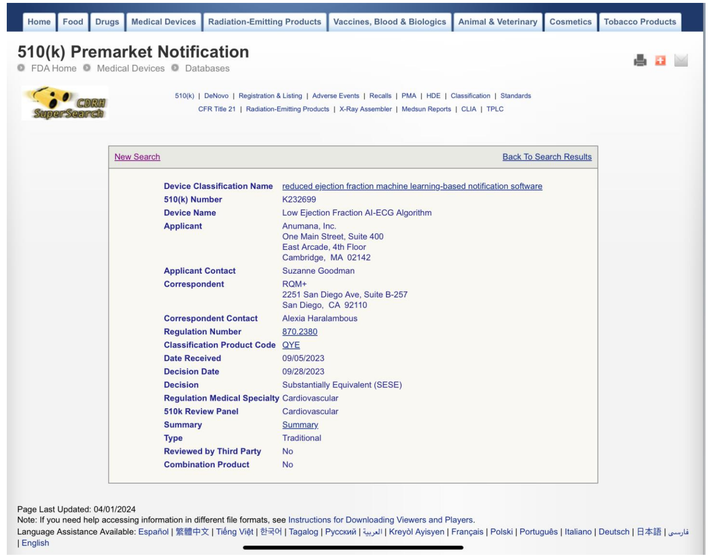
Data as of December 31, 2023

Total of 95 Marketing Authorizations, including 91 CDRH devices and 4 CBER devices

Search: Show entries

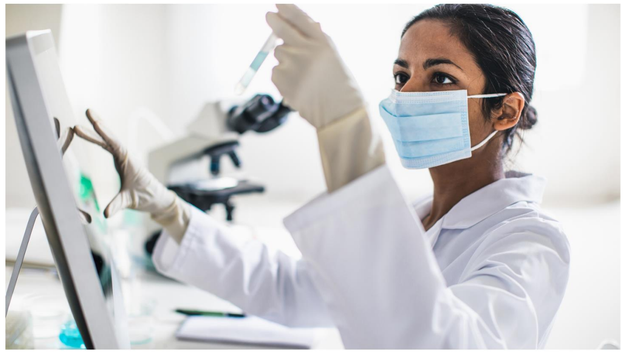
Manufacturer	Trade Name	Marketing Submission Number	Marketing Submission Decision Date
ENDOSOUND, INC.	ENDOSOUND VISION SYSTEM	K232518	12/27/2023
AUTOGENOMICS, INC.	AVERTD and AVERTD BUCCAL SAMPLE COLLECTION KIT	P230032	12/19/2023
MEDTRONIC, INC.	PULSESELECT PULSED FIELD ABLATION (PFA) SYSTEM	P230017	12/13/2023
BIOPORTO DIAGNOSTIC, INC.	PRONEPHRO AKI (NGAL)	K232761	12/7/2023
MEDTRONIC, INC.	SYMPLICITY SPYRAL RENAL DENERVATION SYSTEM	P220026	11/17/2023
RECOR MEDICAL, INC.	PARADISE ULTRASOUND RENAL DENERVATION SYSTEM	P220023	11/7/2023
PERFUZE, LTD.	MILLIPEDE 070 ASPIRATION CATHETER, PERFUZE ASPIRATION TUBE SET	K232524	10/18/2023
HISTOSONICS, INC.	EDISON SYSTEM	DEN220087	10/6/2023
ANUMANA, INC.	LOW EJECTION FRACTION AI-ECG ALGORITHM	K232699	9/28/2023
LAMINATE MEDICAL [redacted]		DEN220026	9/26/2023

Top



Catalog of Regulatory Science Tools to Help Assess New Medical Devices

Share | Post | Linkedin | Email | Print



Update: January 12, 2024: The FDA is providing an updated Regulatory Science Tool (RST) Catalog that will provide additional tool search capability allow for additional capacity as the catalog continues to grow.

About the catalog of regulatory science tools

The Catalog of Regulatory Science Tools provides a peer-reviewed resource for medical device companies to use where standards and qualified Medical Device Development Tools (MDDTs) do not yet exist. These tools do not replace FDA-recognized standards or MDDTs. This catalog collates a variety of regulatory science tools that the FDA's Center for Devices and Radiological Health's (CDRH) Office of Science and Engineering Labs (OSEL) developed. These tools use the most innovative science to support medical device development and patient access to safe and effective medical devices. If you are considering using a tool from this catalog in your marketing submissions, note that these tools have not been qualified as Medical Device Development Tools and the FDA has not evaluated the suitability of these tools within any specific context of use. You may request feedback or meetings for medical device submissions as part of the Q-Submission Program.

For more information about the Catalog of Regulatory Science Tools, email OSEL_CDRH@fda.hhs.gov.

The FDA also maintains a new (2024) catalogue of emerging Big Data and AI tools “medical devices”

- This area of AI in health is moving so quickly the FDA is now tracking **pre-approval** information.
- This is a new program with rich scholarly literature that is very recent.
- AI tools for health are classified as “medical devices” under the program.

GUIDANCE DOCUMENT

Breakthrough Devices Program

Guidance for Industry and Food and Drug Administration Staff

SEPTEMBER 2023

Download the Final Guidance Document | Read the Federal Register Notice

Final

Share | Post | Linkedin | Email | Print

Docket Number: FDA-2017-D-5966

Issued by: Center for Devices and Radiological Health
Center for Biologics Evaluation and Research

The full FDA guidance for AI as a medical device is here:

https://www.fda.gov/regulatory-information/search-fda-guidance-documents/breakthrough-devices-program

Example request for Breakthrough Device Designation:

https://www.fda.gov/regulatory-information/search-fda-guidance-documents/breakthrough-devices-program

Appendix 1: Illustrative Example: Breakthrough Device Designation Request

This appendix provides an example of information that may be helpful to include in a request for designation into the Breakthrough Devices Program.

Background Information

Device Description: This section provides an overview of the product, including principles of operation (including device and/or drug components) and properties relevant to clinical function, if known. Images or engineering schematics are also encouraged for inclusion, as appropriate.

Indications for Use: This section presents indications for use for which you are requesting designation. The indications for use should clearly outline a patient population that meets the designation criteria.[61]

Regulatory History: This section details the history of previous FDA interactions and submissions, including feedback received and resolution of that feedback, as applicable. All relevant IDE, 513(g),[62] and Q-Submission numbers should be included.

Designation Criteria

Criterion 1: Device "provides for more effective treatment or diagnosis of life-threatening or irreversibly debilitating human disease or conditions."[63]

This section provides a discussion regarding how the first designation criterion is met by the proposed device and indications for use.

Criterion 2: Device meets one of the components of the criterion, listed below:

(A) Device "represent[s] breakthrough technolog[y];"[64]

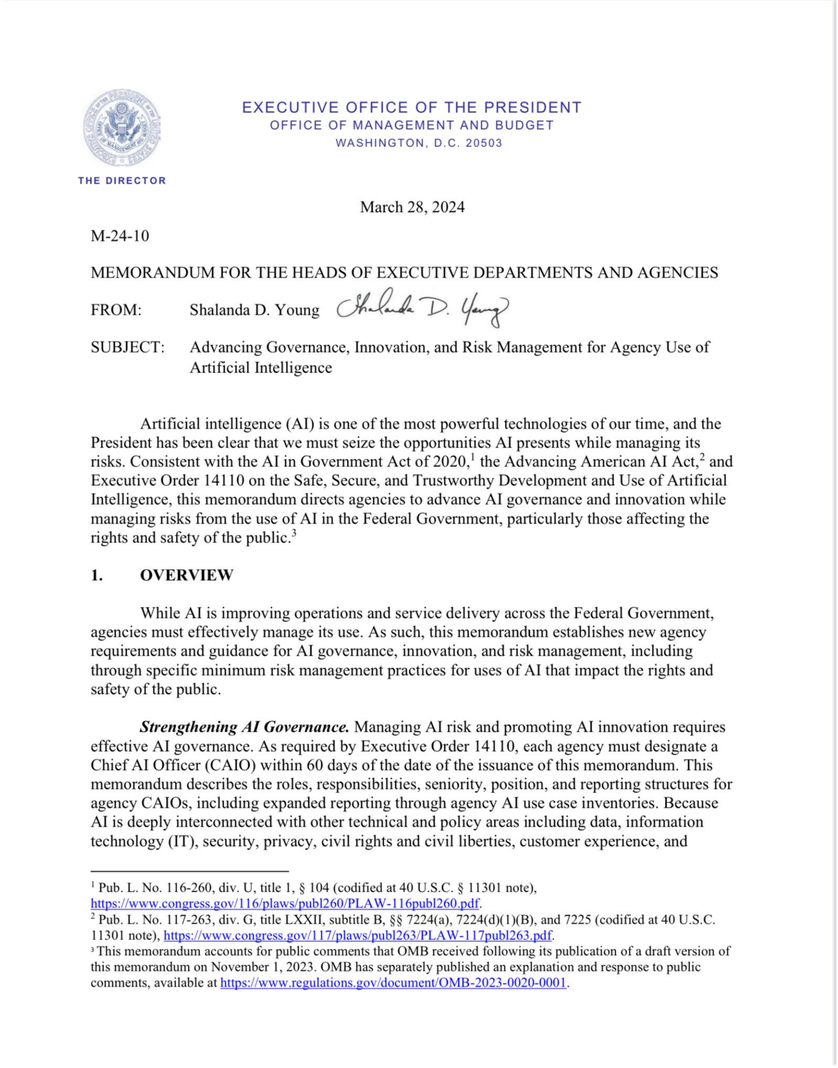
(B) "[N]o approved or cleared alternatives exist;"[65]

(C) Device "offer[s] significant advantages over existing approved or cleared alternatives, including the potential, compared to existing approved alternatives, to reduce or eliminate the need for hospitalization, improve patient quality of life, facilitate patients' ability to manage their own care (such as through self-directed personal assistance), or establish long term clinical efficiencies;"[66] *or*

(D) Device availability "is in the best interest of patients."[67]

---

[61] See section 515B(b) of the FD&C Act (21 U.S.C. 360e-3(b)).
[62] See FDA Guidance "FDA and Industry Procedures for Section 513(g) Requests for Information under the Federal Food, Drug, and Cosmetic Act" available at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/fda-and-industry-procedures-section-513g-requests-information-under-federal-food-drug-and-cosmetic.
[63] See section 515B(b)(1) of the FD&C Act (21 U.S.C. 360e-3(b)(1)).
[64] See section 515B(b)(2)(A) of the FD&C Act (21 U.S.C. 360e-3(b)(2)(A)).
[65] See section 515B(b)(2)(B) of the FD&C Act (21 U.S.C. 360e-3(b)(2)(B)).
[66] See section 515B(b)(2)(C) of the FD&C Act (21 U.S.C. 360e-3(b)(2)(C)).
[67] See section 515B(b)(2)(D) of the FD&C Act (21 U.S.C. 360e-3(b)(2)(D)).



EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF MANAGEMENT AND BUDGET
WASHINGTON, D.C. 20503

THE DIRECTOR

March 28, 2024

M-24-10

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: Shalanda D. Young

SUBJECT: Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence

Artificial intelligence (AI) is one of the most powerful technologies of our time, and the President has been clear that we must seize the opportunities AI presents while managing its risks. Consistent with the AI in Government Act of 2020,[1] the Advancing American AI Act,[2] and Executive Order 14110 on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, this memorandum directs agencies to advance AI governance and innovation while managing risks from the use of AI in the Federal Government, particularly those affecting the rights and safety of the public.[3]

1. OVERVIEW

While AI is improving operations and service delivery across the Federal Government, agencies must effectively manage its use. As such, this memorandum establishes new agency requirements and guidance for AI governance, innovation, and risk management, including through specific minimum risk management practices for uses of AI that impact the rights and safety of the public.

Strengthening AI Governance. Managing AI risk and promoting AI innovation requires effective AI governance. As required by Executive Order 14110, each agency must designate a Chief AI Officer (CAIO) within 60 days of the date of the issuance of this memorandum. This memorandum describes the roles, responsibilities, seniority, position, and reporting structures for agency CAIOs, including expanded reporting through agency AI use case inventories. Because AI is deeply interconnected with other technical and policy areas including data, information technology (IT), security, privacy, civil rights and civil liberties, customer experience, and

[1] Pub. L. No. 116-260, div. U, title 1, § 104 (codified at 40 U.S.C. § 11301 note), https://www.congress.gov/116/plaws/publ260/PLAW-116publ260.pdf.
[2] Pub. L. No. 117-263, div. G, title LXXII, subtitle B, §§ 7224(a), 7224(d)(1)(B), and 7225 (codified at 40 U.S.C. 11301 note), https://www.congress.gov/117/plaws/publ263/PLAW-117publ263.pdf.
[3] This memorandum accounts for public comments that OMB received following its publication of a draft version of this memorandum on November 1, 2023. OMB has separately published an explanation and response to public comments, available at https://www.regulations.gov/document/OMB-2023-0020-0001.

White House Regulation of Big Data and AI and Health; new guidance

https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf

Agencies must manage and mitigate the risks from the use of AI. This includes:

- Designating Chief AI Officers
- Convening an Agency AI Governance Body
- Making a compliance plan,
- Keeping use case inventories
- Reporting on AI Use Cases not subject to inventory, among other tasks.

AI that is used for health purposes has been designated as “safety impacting AI” and is required to undergo additional regulatory protections. For example, agencies working with Health and AI must

- Complete an AI Impact Assessment,
- Test the AI for performance in a real - world context,
- Independently evaluate the AI, and
- Conduct ongoing monitoring,
- Regularly evaluate risks from the use of AI
- Mitigate emerging risks to rights and safety,
- Ensure adequate human training and assessment
- Provide additional human oversight, intervention, and accountability as part of decisions or actions that coul result in a signficant impact on rights or safety.
- Provide public notice and plain language documentation.
- Among additional requirements.

Pam Dixon Executive Director
World Privacy Forum | www.worldprivacyforum.org
pdixon@worldprivacyforum.org
@privacyforum

Resources and additional thoughts:

- Principles for Digital Development: https://digitalprinciples.org
- World Bank Principles on identification for sustainable development https://id4d.worldbank.org/principles
- OECD AI Recommendation and implementation manual http://www.oecd.org/going-digital/ai/principles/
- ICO Big Data Report, with sample PIA and DPIA for big data: https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf

Updated considerations regarding health data flows and health data flow infrastructures:

Multi-jurisdictional: Health data flows are important at the local, national, regional, and international levels. (Post-pandemic changes).

High volume: The generation of public health data is so large and complex that it can no longer be saved or analyzed using conventional data processing methods. COVID-19 was a “Big Data” pandemic. It created changes in how health data is handled.

High velocity: The speed with which health data is generated, processed, and analyzed has increased dramatically. Ideally, the data processing will occur within fractions of a second, also known as “real time.” This speed also requires greater automation of data and privacy governance.

High complexity: The diversity of data types and sources in health data is a challenge that still needs to be solved fully.

High sensitivity to design choices: there are many trade-offs in design choices (for example, in databases and in contact tracing), and it is crucial to work through all of the possibilities through the lifecycle of the data flows, uses, and storage.

Core considerations when thinking about Big Data:

The importance of detailed planning in order to do no harm and to provide benefit is central:

- Depending on how it is designed, implemented, and used, Big Data (and the AI and ML analysis it typically undergoes) can produce a wide range of consequences for individuals, groups of people, and institutions.

- Beneficial consequences are the goal, but they require considerable work and planning. Negative consequences can result from improper planning, and can sometimes even happen where there has been good planning in place.

- Big data and AI does not exist in a vacuum: it is almost always a part of a much larger technology and policy ecosystem

- Big data (and AI) implementations express values and ethics