Northwestern
INFORMATION TECHNOLOGY

Responsible Use of AI in Human Research

IRB Brown Bag

Christina Maimone
Northwestern IT Research Computing and Data Services

Daniel Schneider
NMEDW

May 2024

Research Computing and Data Services

https://bit.ly/rcdsinfo

- Platforms and services in support of all researchers
- Computing (Quest, Genomics Compute Cluster)
- Data management (storage, security, workflows)
- Data science, statistics, and visualization

Note: Services and guidance are changing over time. This information is from May 2024.

AI/Artificial Intelligence: Scope

- Shorthand for computational systems that use input to make or inform decisions or generate output related to that input
- Includes machine learning, predictive models, generative models
- Text, images, and other input and output formats
- Generative AI adds some complexity, but most considerations and issues apply to other tools as well

Generative AI output often seems reasonable while being very wrong

Perspectives

- General principles
- IT security issues
- Research considerations: IRB and otherwise

>> Some clarity around tools, lots of questions to consider for research design and practices

General Principles:
Blueprint for an AI Bill of Rights

Office of Science and Technology Policy
https://www.whitehouse.gov/ostp/ai-bill-of-rights/

AI Bill of Rights

You should be protected from unsafe or ineffective systems.

You should not face discrimination by algorithms, and systems should be used and designed in an equitable way.

You should have agency over how data about you is used.

AI Bill of Rights

You should know that an automated system is being used and understand how and why it contributes to outcomes that impact you.

You should be able to opt out, where appropriate, and have access to a person who can quickly consider and remedy problems you encounter.

What is IT worried about?

Data Classification Levels

Level 1: Public data: Is public or OK to be public

Level 2: Sensitive, not public: General research data, university business operations

Level 3: Contractual and legal restrictions: Personal sensitive information (SSNs), PHI, DUAs, restricted financial data, CUI

Level 4: Classified, Export Controls

https://policies.northwestern.edu/docs/data-classification-policy.pdf

Data Security

Data being stolen or publicly exposed through attacks on, mistakes by, or insufficient security practices of third-party companies

Ownership/transfer/retention of data: inputs, outputs, prompts, results

Maintaining terms of **data agreements**

Data Security

Private **data leaking** into the public by being used to train new AI models

Reidentification of data when combined with other information

Broad **access to file systems** may exacerbate problems with erroneous or overly liberal permissions settings

Rapidly **changing use agreements** and terms

IT Guidance on Generative AI

Public data (Level 1)

Interactive/chatbots: Public tools OK

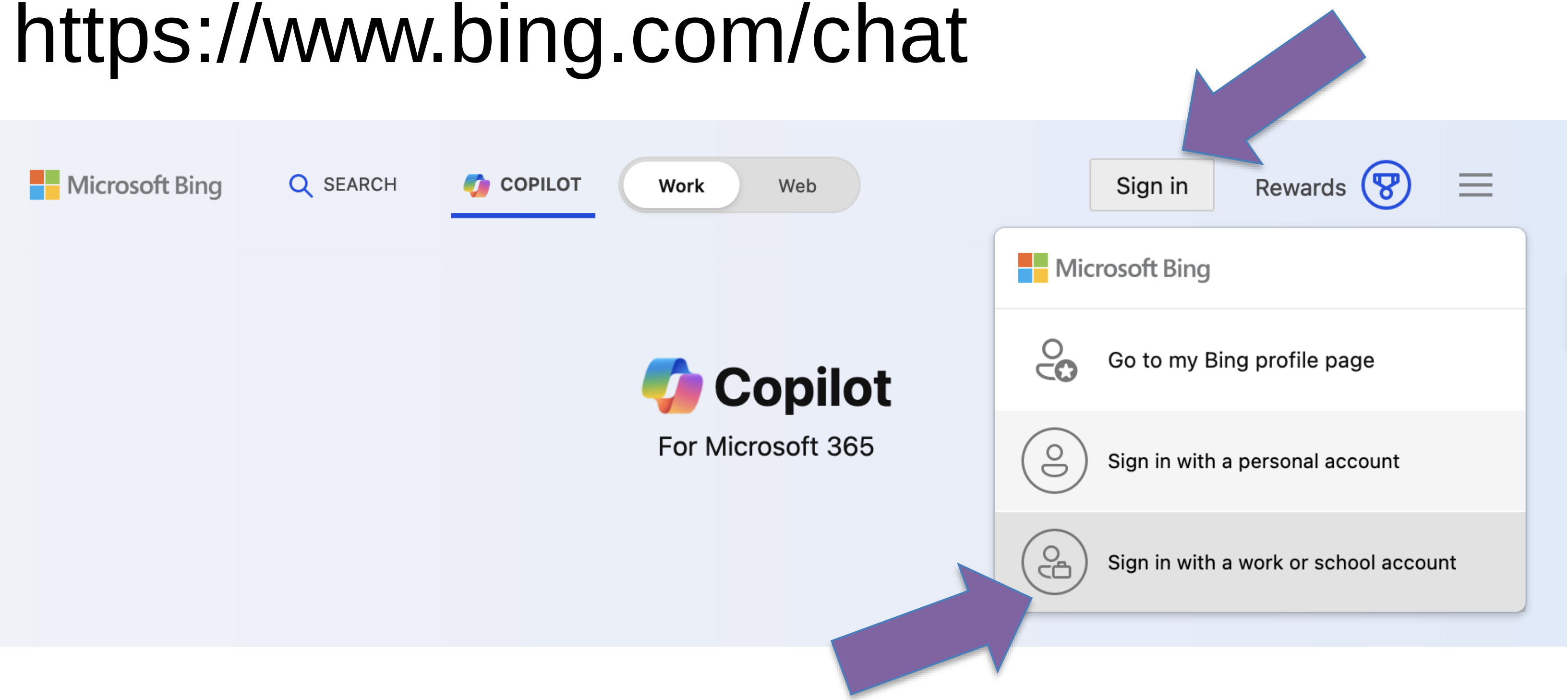
- Sign in with NU account with Microsoft Bing/Copilot for additional protections

APIs/code: through Microsoft Azure

- Multiple models available, including OpenAI models
- Why? Help guard against accidental data leaks; terms of use and policies change rapidly, especially for non-enterprise accounts

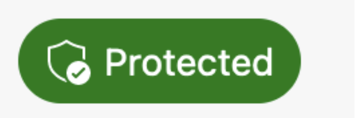
https://www.it.northwestern.edu/about/policies/guidance-on-the-use-of-generative-ai.html

https://www.bing.com/chat



Microsoft Bing
SEARCH
COPILOT
Work
Web
Sign in
Rewards
Copilot
For Microsoft 365
Microsoft Bing
Go to my Bing profile page
Sign in with a personal account
Sign in with a work or school account

Sign in with @northwestern.edu address through university SSO; Look for



IT Guidance on Generative AI

Sensitive/restricted data (Level 2 and 3*)

Interactive/chatbot

- Sign in with NU account to Microsoft Bing/Copilot; no history, conversation limit
- Do not enter restricted data in other tools

APIs/code

- Microsoft Azure OpenAI service with security controls
- Local models on secure, university-owned devices or systems

*Specific requirements vary; talk to IT

Zoom/Meeting Assistants

- Zoom is approved for use in most human research
  - Local recordings OK; avoid transcriptions and cloud recordings
- Increased security version of Zoom available for conversations with PHI or personal identifiers such as SSN
  - If you need transcriptions or cloud recordings, you may need this version

- Do not use external tools like Otter.ai
- Do not join meetings where personal assistant tools are an "attendee" if personal information or data will be discussed

Coding Assist Tools

Concerns

- Unknowingly incorporating code that violates license agreements or copyright into your files
- Your code ending up in others’ projects
- Information from data files and credentials leaking
- Low code quality and maintainability
- Security vulnerabilities and lack of security best practices

Writing Code

Practices

- Use a separate chat window instead of a plugin for your code editor
- GitHub Enterprise: Northwestern account, offers additional security controls and copyright precautions for repos and GitHub Copilot (still not OK for PII/PHI)
- Use synthetic or fake data files

Confidentiality of Research

Think of sharing data with a service/platform/ app/company like sharing with a person

- Should that entity have access to your research data?
- Do they know how, and have the right tools, to protect it?
- What is the recourse if something goes wrong?

Third-party Tools

Requires **Service Provider Security Assessment** through IT

Examples: Transcription or image processing services, automated data analysis platforms

Unlike with lab supplies, there are not existing preapproved vendors for use with PII/PHI

fsmhelp@northwestern.edu; researchdata@northwestern.edu

Who To Talk To

Medical Records Data: EDW
https://www.feinberg.northwestern.edu/research/cores/units/edw.html

Other Feinberg Research: FSM IT fsmhelp@northwestern.edu

Everyone Else: Northwestern IT Research Computing and Data Services researchdata@northwestern.edu

Use University-Approved Resources

Do not upload research data into any tool, service, or website (AI-related or not) that is not approved by the university for use with human research data

Research Considerations

Consent

People are challenging the use of their information and creations in building LLMs and other generative AI models

Many populations have been harmed by biased AI or predictive models

>> *Consent to interaction with, or data use for, AI tools should not be assumed*

Identity: Data Collection and Use

Is it public?

Expectations of privacy based on

- Restricted access groups or sites (logins, subscriptions, friend networks)
- Terms of service of sites/platforms/apps

Identity: Data Collection and Use

Is it identifiable?

Removing names (real or otherwise) alone does not sufficiently deidentify data when text or images can be searched and easily reidentified

Identity: Data Collection and Use

Is it deidentified?

Reidentification of data through combination with other information

- By researchers (ex. by linking across social media platforms)
- By combining data from multiple sources in creating LLMs or other models

Privacy

Possibility of ML/AI models revealing private information about participants (or others) that was not part of the study

>> Model that predicts sexual orientation from photographs, with higher accuracy than humans

- Data collected from dating sites (no consent)
- Would it be OK to use this model to infer sexual orientation of participants in other studies? To fill in missing data values?

Is it research?

“It is the nature of research that diagnostic tools must be developed with data from a subset of the full population, hence…development of an AI/ML tool is not different from the development of an in vitro **diagnostic device** and SACHRP takes the position that it should have the same degree of regulatory oversight.”

IRB Considerations on the Use of Artificial Intelligence in Human Subjects Research

Risk of Harm

- From reidentification, inference of private data
- From inaccurate output
  - What would happen if the model or tool you’re using gave the wrong answer or output?
- From participant exposure to AI-generated content
  - What safeguards are in place?
  - What if a tool misgenders them? Misrepresents their race or ethnicity? Outputs something offensive?
- From biased models/output

Bias Across Population Groups

Bias from the data used to create the models

- Historical: models seek to make future outcomes look like past outcomes
- LLMs built from limited text (English, public, western, electronic)
- Using proxies for variables of interest: studies using health care expenditures/costs to proxy for need for care

Bias from model evaluation and decision criteria

- Algorithms can optimize performance metrics by excluding minority groups

Assessing Bias

Have a plan!

- Document approaches that will be used to assess *and mitigate* bias in study protocols
- New models: assess throughout the development process
- Using existing models: assess bias in outcomes for your use case, especially for models used in decision-making processes
- Consider independent review

Assessing Bias

Assessment checklists can help, but may not be enough

Example: AI models can predict patient self-reported race from medical images (study)

- Humans cannot, and generally did not consider it possible
- Humans did not know to look for the model potentially capturing features of race because they did not think such features could be detected

NMEDW

NMEDW Overview

- The Northwestern Medicine® Enterprise Data Warehouse (NMEDW) is a joint initiative across the Northwestern University Feinberg School of Medicine (FSM) and Northwestern Memorial HealthCare (NMHC)

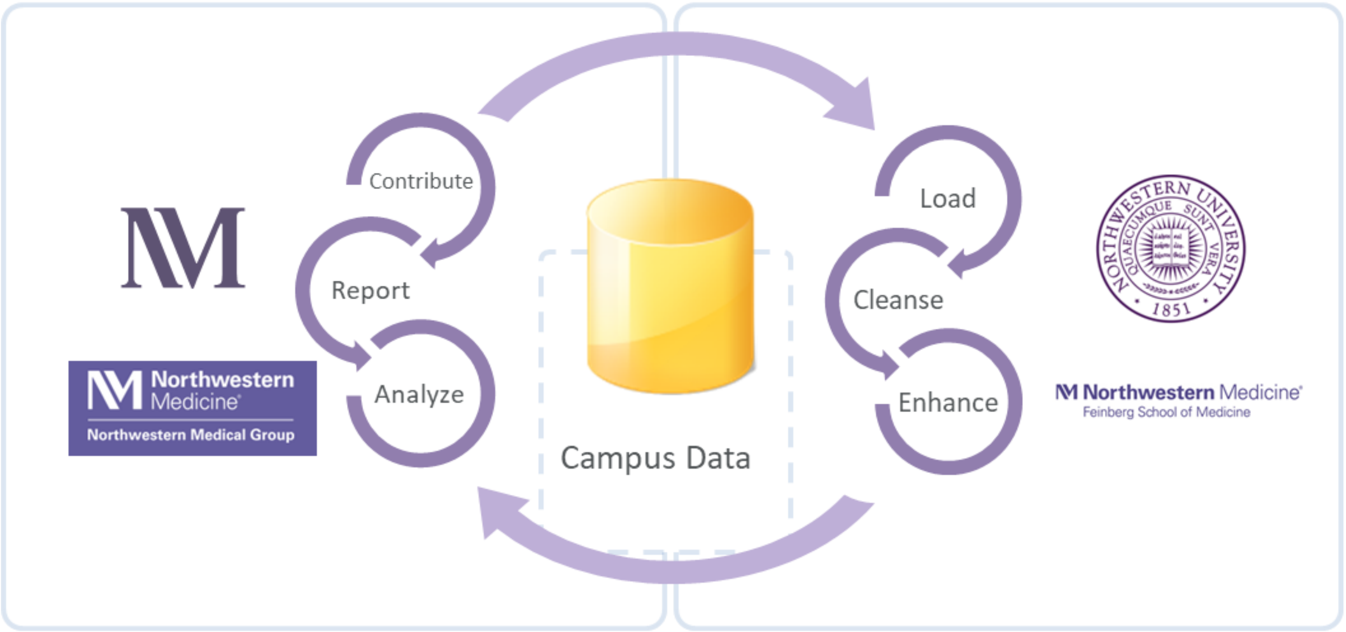
- Mission Statement:
  - To create a single, comprehensive, and integrated repository of all clinical and research data sources on the campus to facilitate research, clinical quality, healthcare operations, and medical education

- The EDW contains data from 150+ source systems, and is continuously loading data from new source systems. The EDW also manages all Clarity and Caboodle environments at NMHC

The NMEDW Provides a Campus Data Platform for Your Approved Projects

Participating Entities

NMEDW



Contribute

Report

Analyze

Campus Data

Load

Cleanse

Enhance

Northwestern Medicine

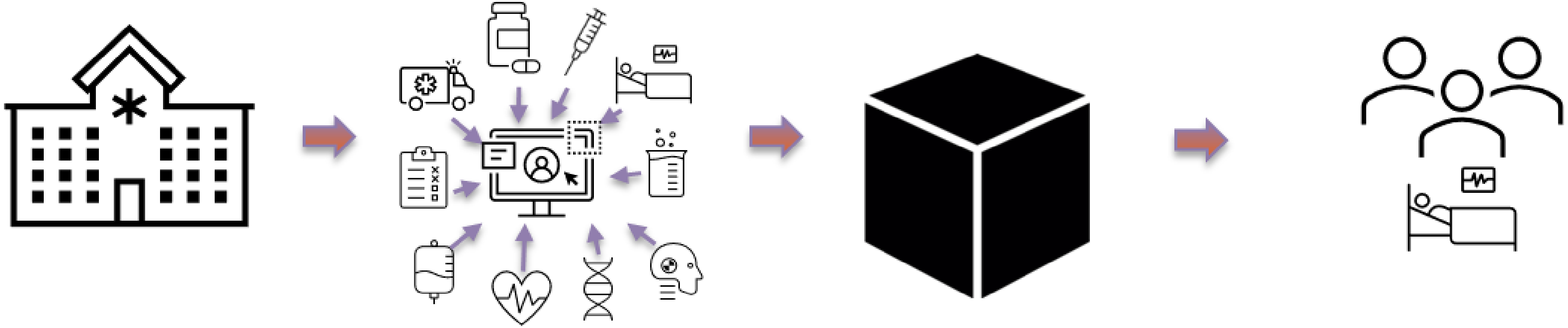
Northwestern Medical Group

Northwestern Medicine Feinberg School of Medicine

Analyze Data and Create Reports

Provide Campus Data Platform

Healthcare Data is Complex

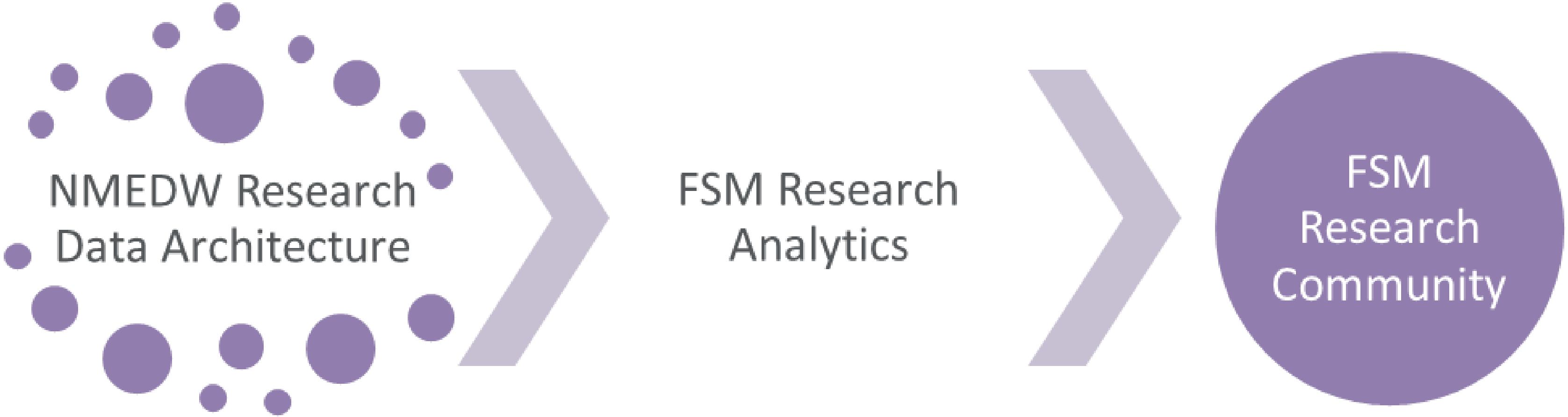


Input
(11 NM hospitals; 150+ source systems)

NMEDW Infrastructure
(integrated repository of clinical and research data sources)

Output
(translational research, publications, reporting dashboards, improved patient care)

Research Analytics Team Mission



NMEDW Research Data Architecture

FSM Research Analytics

FSM Research Community

Develops Research Architecture and Data Modeling to support cross-institutional research initiatives

Supports analytic and reporting needs via unique data objects/solutions for the FSM research community

Consumes data and analytics to secure grant funding, top tier publications, innovative translational research, and to further medical education

What is a Data Steward?

- The data in the NMEDW is still owned by the contributing organizations
- Data Stewards are representatives from these organizations and approve data and access requests
- For example, when requesting clinical data from NMHC, there needs to be a documented approval by their designated data steward

Resources

Working with AI

There are data scientists and research analysts with experience working with AI models and tools available to help

- Northwestern Medicine Enterprise Data Warehouse
  - https://www.feinberg.northwestern.edu/research/cores/units/edw.html
  - Health and clinical research data – not just medical records
- Research Computing and Data Services
  - https://bit.ly/rcs_data
  - All types of data, all fields

External Resources

- IRB Review of Research Involving AI, Benjamin Silverman
- Artificial Intelligence Human Subjects Research IRB Reviewer Checklist (with AI HSR and Exempt Decision Tree) by Tamiko Eto
- HHS Human Research Protections: IRB Considerations on the Use of Artificial Intelligence in Human Subjects Research
- FDA Software as a Medical Device: multiple resources in this section
- FDA: Good Machine Learning Practice for Medical Device Development: Guiding Principles

External Resources

- University of Tennessee, Research Integrity and Assurance
- Michigan Institute for Data Science, Using Generative AI for Scientific Research
- NIST, Towards a Standard for Identifying and Managing Bias in Artificial Intelligence
- AI Reporting Guidelines: How to Select the Best One for Your Research
- AI Fairness, by Trisha Mahoney, Kush R. Varshney, Michael Hind (available through the library); IBM AI Fairness 360 (one example; there are others)

External Resources

- Association of Internet Researchers, Ethical Guidelines
- Algorithmic Justice League, Equitable AI
- WHO: Ethics and governance of artificial intelligence for health
- Northwestern Center for Advancing Safety of Machine Intelligence
- More resources on bias in data science algorithms and models from Research Computing and Data Services

Use University-Approved Resources

Do not upload research data into any tool, service, or website (AI-related or not) that is not approved by the university for use with human data

Questions?

fsmhelp@northwestern.edu
researchdata@northwestern.edu