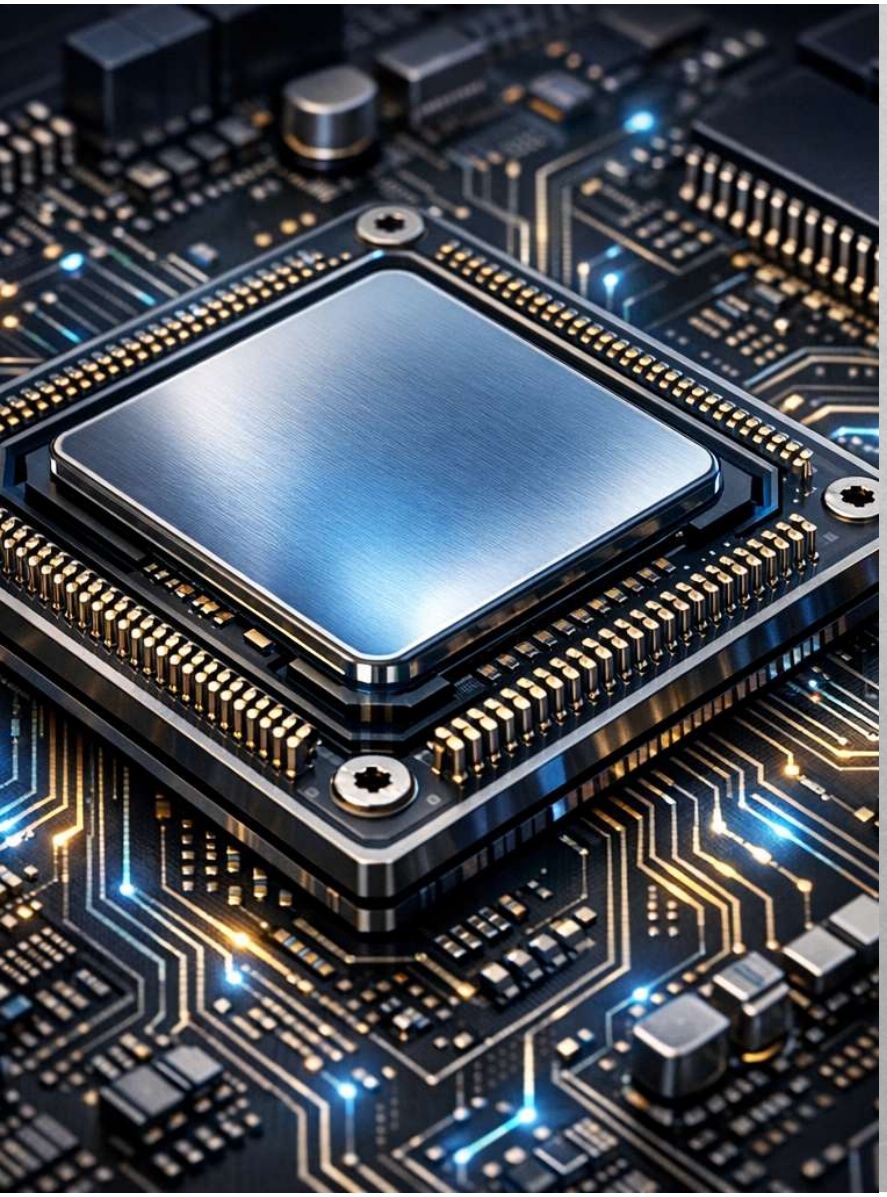


PREVENTING A REPEAT OF THE PAST: ETHICAL CONSIDERATIONS IN DATA COLLECTION & USE IN THE AGE OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

ASHLEY C. SMITH, PH.D.*

&

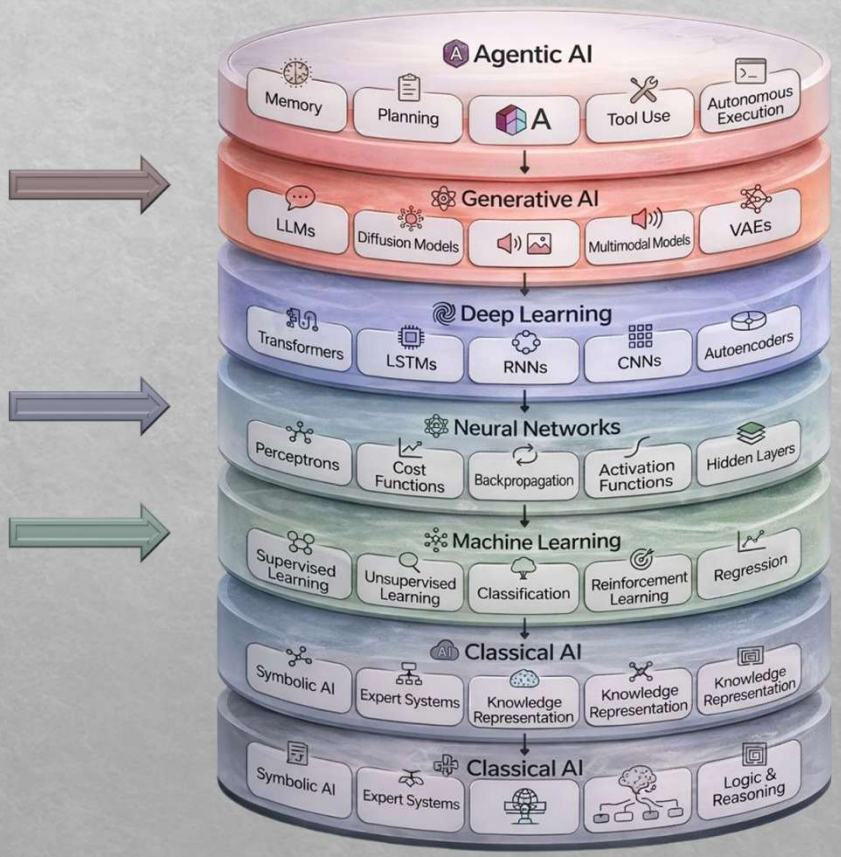
EMAN FAISAL, M.SC.

AMERICAN ACADEMY OF FORENSIC SCIENCES
FEBRUARY 13TH, 2026, NEW ORLEANS, LA

UNIVERSITY OF TORONTO

Definitions

- Artificial Intelligence
  - Computational model & structure designed to perform tasks that typically require human intelligence
    - Learning, reasoning, perception, problem-solving
    - Perform output
- Machine Learning
  - Uses algorithms & data sets to generate predictions, classifications, & trends
- Neural Networks
  - Model training based on human brain w/ layered, interconnected structure
- Diffusion Language Models
  - Generative AI model that produce results by iteratively denoising a sequence
- Large Language Models
  - Deep learning model that utilizes pre-trained data-sets to analyze billions of parameters
  - Enables the model to understand, summarize, translate, and (importantly) generate human-like language



Ethical Models

American Association of Biological Anthropologists

- American Association for Biological Anthropology
  - Ad hoc committee, developing white paper on skeletal collection curation & use
- Canadian Association for Biological Anthropologists
  - Use of human remains listed in Code of Ethics (Article A of Constitution)
  - Respects individuality of remains
- British Association of Biological Anthropology & Osteoarchaeology
  - Statement on use of human remains
  - Excavation when absolutely necessary
  - Encouraged reburial after study
- Australian Anthropological Society
  - Codified in CoE, respects individuality
- **None have positions on ethics related digital data collection & curation**

Some Issues with Past Skeletal Collections

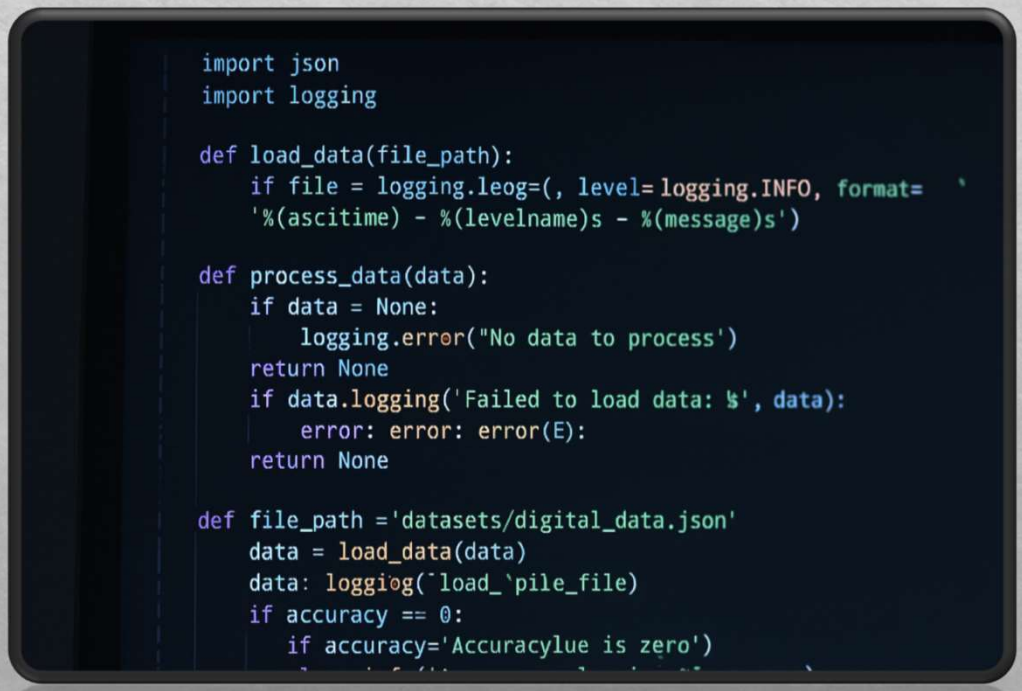
- Curated using remains from ethically questionable assemblages
  - Limited or lacking documentation of provenance
  - Developed from indigenous, enslaved, and impoverished communities
  - Questions can be raised regarding the legality of the remains in some collections
- Lacked respect, input, or guidance from descendent communities with regards to exhumation and curation
- Often lacked informed consent from decedents or family regarding use of remains
- Lacks/ed proper documentation regarding the contents and history of the collection from accession to present
- **Most of these collections are no longer in use for research or teaching**
  - **Being returned to descendent communities**

Positives with More Modern Collections

- Documented informed consent from decedents regarding donation & use
- Documented life history of the decedent
- Documented history of the remains from accession to final disposition/present
- Multiple forms of documentation and verification to prevent ethical & legal failures
- **Allows for wide-ranging research conducted in a more ethical manner**

Issues with Digital Data Collections

- Follows legal requirements, but often are ethically malleable
- Created from legally sourced data
  - May lack informed consent of decedents, families, & communities
- Curated from casework data
  - Can lack same permissions
- Many, less reputable, datasets lack proper documentation regarding the data within
  - Accuracy, history, use-permissions (outside of informed consent), provenance
- **Improved technology, but falling on old errors in curation**

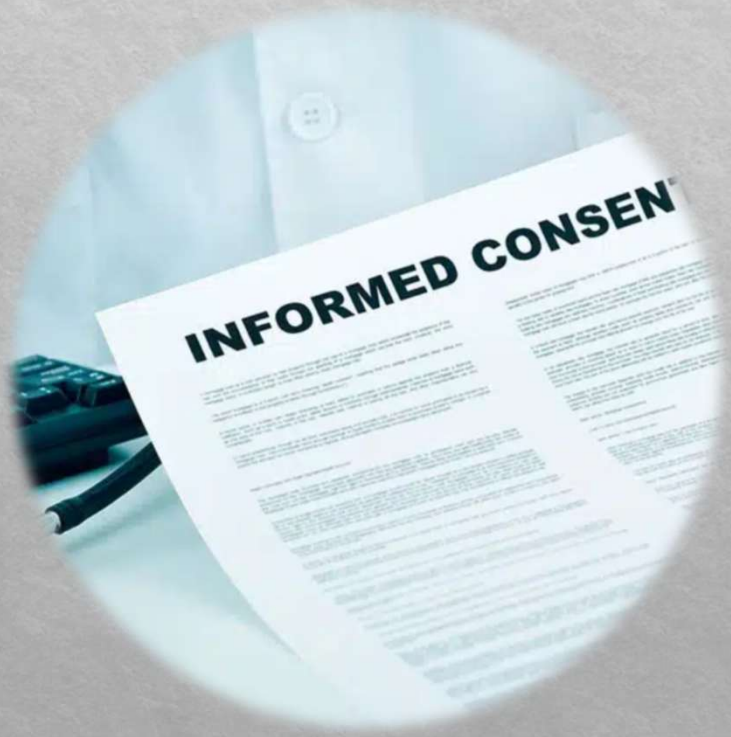


Proposed Ethical Principles

Overarching Principles

- Recognize the humanity of the individual
  - Remains
  - Data
- Recognize the ownership of the individual
- Balance with the needs of society
- Balance with the needs of the nature of research
- Move from an archaeological/museum framework to healthcare/biomedical framework
- Curators and Researchers must know
  - Data
  - Models
  - Outcomes
  - Avoid over-reliance
- Situational Ethics preferred model
  - Each situation is unique
  - Ethical decisions based on individual needs & situations
  - No “one size fits all”

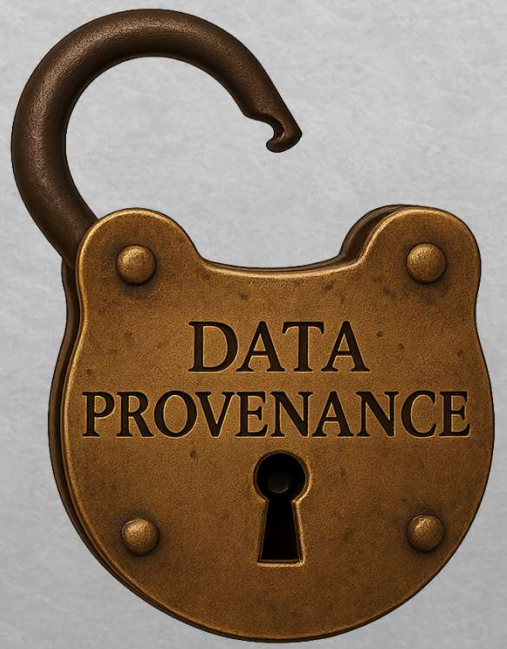
Informed Consent



- Primary Ethical Principle
- Decedents & families should:
  - 1˚: Opt-in to data being used for forensic research
    - Know & note what research they are consenting to
    - The time frame data will be used for (limited or infinite)
  - 2˚: Allowed to opt-out data being used for research
- Researchers continue use, but recognize the lack of expressed consent if not present/known
- Individual universities/programs may have different requirements for need of IC
  - Development of protocols for familial acknowledgement for use of data

Open Access to Data Provenance

- Provenance of all data within dataset should be available to researchers
- This includes:
  - Data origins
  - Informed consent releases
  - Data histories/modifications
  - Curators & managers of data
    - Ideally data input as well
  - Goes beyond whole datasets to individual data within the set

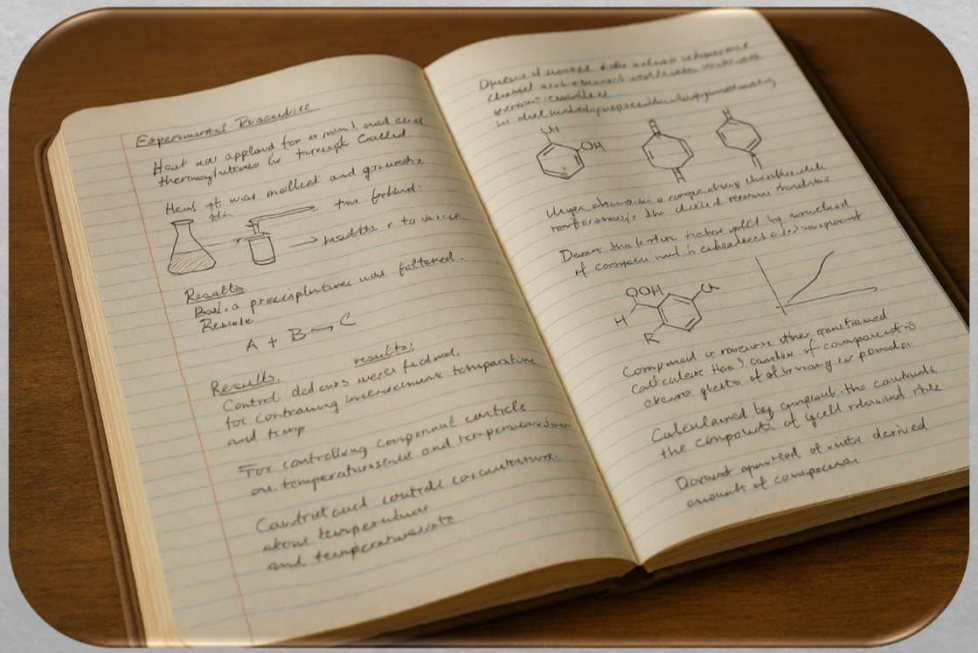


Open Access to Data

- All data should be open & free for use to qualified researchers
  - Data managers should maintain right to determine requirements for "qualified"
- Selling access to data limits scientific viability & reliability
  - Limits access to select researchers & institutions
- Includes algorithmic models
  - Researchers should have understanding of algorithm in use
- This principle is largely standard practice in field
  - Rising reliance & commodification of technology could raise this as an issue in future

Researchers Detailing Use of All Data

- Researchers developing or utilizing any of the AI models should detail use of all data
- Includes:
  - What specific data was included
  - How it was used
  - What data was excluded
  - Parameters for inclusion/exclusion
  - Modifications to data
  - Algorithmic developments and modifications
  - Recording & securing of consent data (if available)
- Like *Open Access to Data*, this is largely done in field today, but may change in future as technology and commodification changes

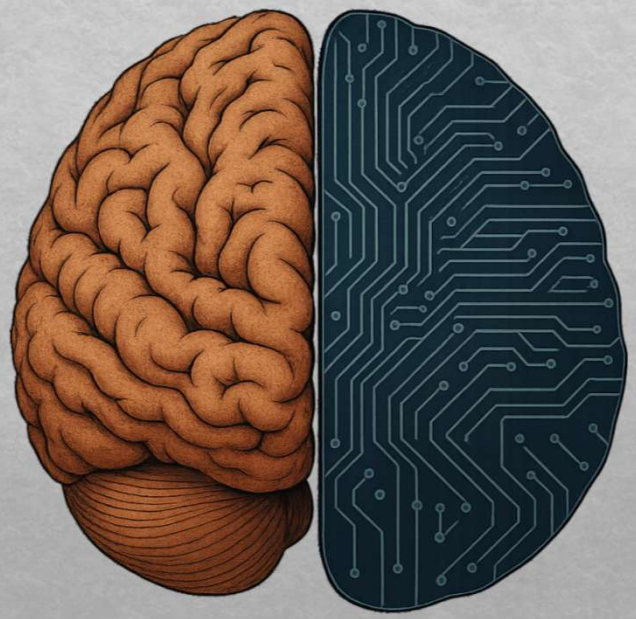


Better Record Keeping in Digital Datasets

- Database histories including:
  - Management
  - Technical participants
  - Software & hardware used
- Record of any data modifications
  - Segmented data from larger sets
- Life history of decedent & data
  - Important to return the individual to anonymized data
- Publication record from those who used datasets
  - Honors individuals within data
  - Ensures accountability & validation of datasets & researchers

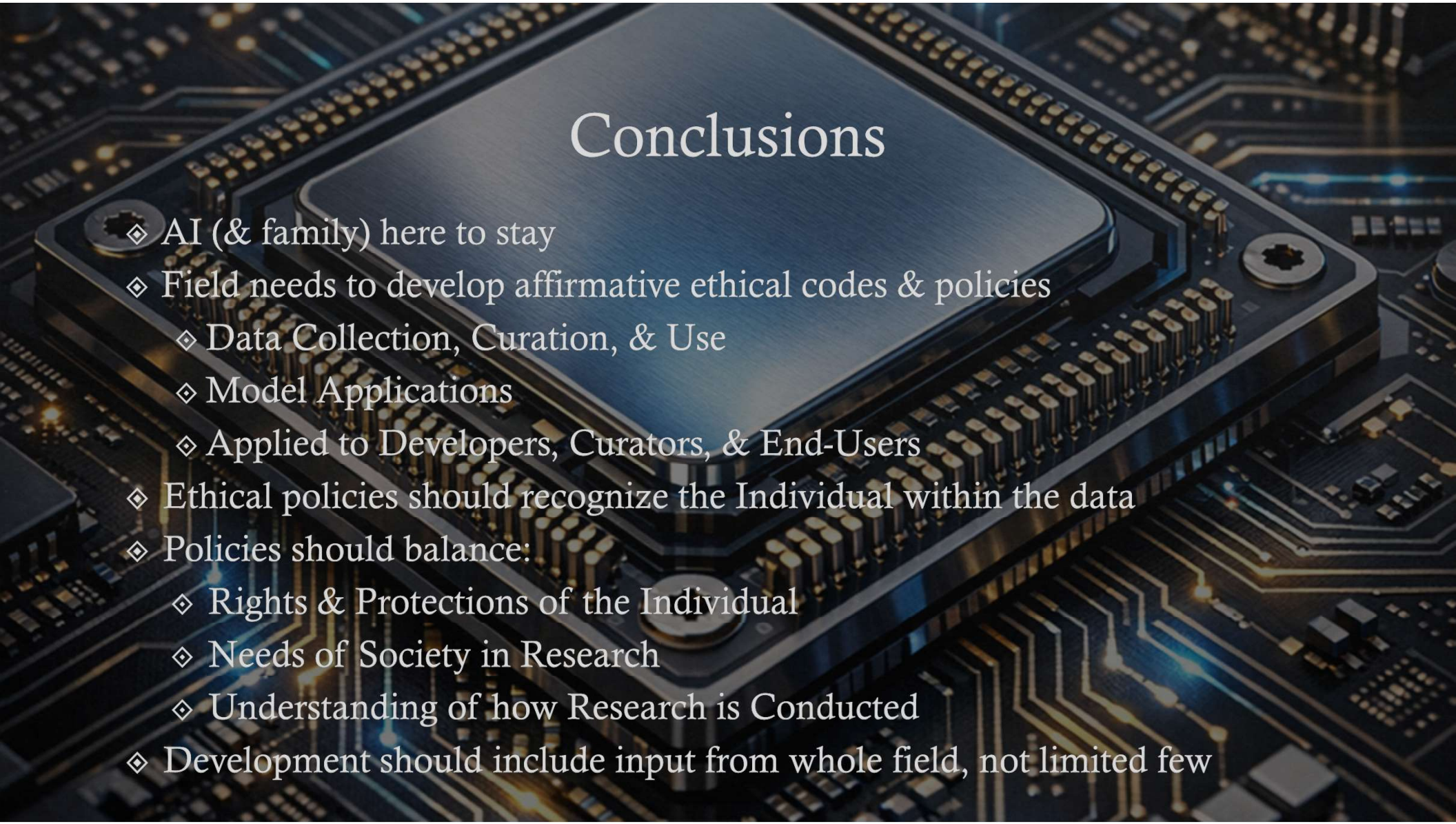
Better Training & Understanding of Technology

- Researchers & Analysts need a firm understanding of AI models and technology
  - Possibilities
  - Limitations
  - Errors & Hallucinations
- Rely on technology as an assistive tool
- Rely on experience as primary assessment
- Analysts & researchers are more than prompt engineers
  - Must have a core concept of the biology and science behind the analyses & technology *beyond* AI use
- **If Informed Consent is primary for data input, Training & Understanding must be primary for Data Output**



“*Be careful that your use of AI does not limit your true human growth. Use it in such a way that, if it disappeared tomorrow, you would still know how to think, create, and act on your own.*”

— Leo XIV, NCYC 2025



Conclusions

- AI (& family) here to stay
- Field needs to develop affirmative ethical codes & policies
  - Data Collection, Curation, & Use
  - Model Applications
  - Applied to Developers, Curators, & End-Users
- Ethical policies should recognize the Individual within the data
- Policies should balance:
  - Rights & Protections of the Individual
  - Needs of Society in Research
  - Understanding of how Research is Conducted
- Development should include input from whole field, not limited few

References

(2015). Canadian association for biological anthropology constitution & code of ethics. Canadian Association for Biological Anthropology/l'Association Canadienne d'anthropologie biologique. Retrieved from https://caba-acab.ca/constitution/

(2025). Joint statement on the retention, display, and provision of access to human remains [Press release]. Retrieved from https://babao.org.uk/joint-statement-on-the-retention-display-and-provision-of-access-to-human-remains/

(2025). Australian anthropological society code of ethics. Curtin, AU: Australian Anthropological Society. Retrieved from https://www.aas.asn.au/code-of-ethics

Agarwal, S. C. (2024). The bioethics of skeletal anatomy collections from India. *15* (1). doi:10.1038/s41467-024-45738-6

Auerbach, B. M. (2024). AABA task force for the ethical study of human remains. *Governance - Ad Hoc Committees.* Herndon, VA: American Association of Biological Anthropologists. Retrieved from https://bioanth.org/about/aaba-task-force-for-the-ethical-study-of-human-remains/

Begum, N. F., Ramalingam, K., & Ramani, P. (2024). Storage, retention, and use of leftover pathology specimens: The underestimated treasures. *16* (1), e53025. doi:10.7759/cureus.53025

Boggio, A. (2008). Ownership of samples and data and territorial restrictions concerning data and samples beyond national boundaries. In B. Elger, N. Biller-Andorno, A. Mauron, & A. M. Capron (Eds.), *Ethical issues governing biobanks: Global perspectives* (1st ed., pp. 197–205). New York: Routledge.

References

Elger, B. S. (2008a). Consent and use of samples. In B. S. Elger, N. Biller-Andorno, A. Mauron, & A. M. Capron (Eds.), *Ethical issues in governing biobanks: Global perspectives* (1st ed., pp. 57–88). New York: Routledge.

Elger, B. S. (2008b). Consent to research involving human biological samples obtained during medical care. In B. S. Elger, N. Biller-Andorno, A. Mauron, & A. M. Capron (Eds.), *Ethical issues in governing biobanks: Global perspectives* (1st ed., pp. 89–120). New York: Routledge.

Elger, B. S. (2008c). Withdrawl of consent and destruction of samples. In B. S. Elger, N. Biller-Andorno, A. Mauron, & A. M. Capron (Eds.), *Ethical issues governing biobanks: Global perspectives* (1st ed.). New York: Routledge.

Elger, B. S., Biller-Andorno, N., Mauron, A., & Capron, A. M. (Eds.). (2008). *Ethical issues in governing biobanks: Global perspectives*. New York: Routledge.

Fletcher, J. F. (1966). *Situation ethics: The new morality*. Philadelphia: Westminster Press.

Ganguli-Mitra, A. (2008). Collective consent. In B. S. Elger, N. Biller-Andorno, A. Mauron, & A. M. Capron (Eds.), *Ethical issues in governing biobanks: Global perspectives* (1st ed., pp. 121–130). New York: Routledge.

Kakaliouras, A. M. (2014). When remains are "lost": Thoughts on collections, repatriation, and research in American physical anthropology. *57* (2), 213–223. doi:10.1111/cura.12062

References

Nichols, C. A. (2014). Lost in museums: The ethical dimensions of historical practices of anthropological specimen exchange. *57*(2), 225–236. doi:10.1111/cura.12063

Radin, J. (2014). Collecting human subjects: Ethics and the archive in the history of science and the historical life sciences. *57*(2), 249–258. doi:10.1111/cura.12065

Turner, T. R. (2014). Large scale collections of biological material and ethical first principles. *57*(2), 259–267. doi:10.1111/cura.12066

Watkins, D. (2025, November 21). Pope Leo to young people: Technology can help us live our Christian faith. *Vatican News*. Retrieved from https://www.vaticannews.va/en/pope/news/2025-11/pope-leo-xiv-usa-national-catholic-youth-conference.html

*Most images generated using Microsoft™ CoPilot® Generative Artificial Intelligence System (2026)