T305: Navigating Compliance Challenges in the Age of Advanced Information Technology

Presenter: Michael Jones, DMgt
Presentation Date: May 8th, 2025

Introduction

Presenter's Background

Artificial Intelligence (AI)

Cybersecurity

Big Data

Presenter Background – Michael Jones

Administrative Director, Mitochondrial Medicine Frontier Program, Children's Hospital of Philadelphia

Associate Director, Farber Institute for Neurosciences, Thomas Jefferson University

Associate Director, Research Administration Center of Excellence, Thomas Jefferson University

Financial Analyst, Farber Institute for Neurosciences, Thomas Jefferson University

Artificial Intelligence

AI vs ML vs DL vs Data Science

Artificial Intelligence (AI)

- Broad concept of developing machines that can simulate human thinking, reasoning, and behavior

Machine Learning (ML)

- A subset of AI wherein computer systems learn from the environment and, in turn, use those learnings to improve experience and processes.
- All machine learning is AI, but not all AI is machine learning

Deep Learning (DL)

- Part of a broader family of machine learning methods based on artificial neural networks
- DL uses multiple layers to progressively extract higher – level features from raw input

Data Science

- The art of processing, analysis, and extraction of relevant assumptions from data
- Data Scientists make use of machine learning to predict future events

What is AI?

"A brand of computer science devoted to developing data processing systems that performs functions normally associated with human intelligence, such as reasoning, learning, and self environment" (Virginia Tech)
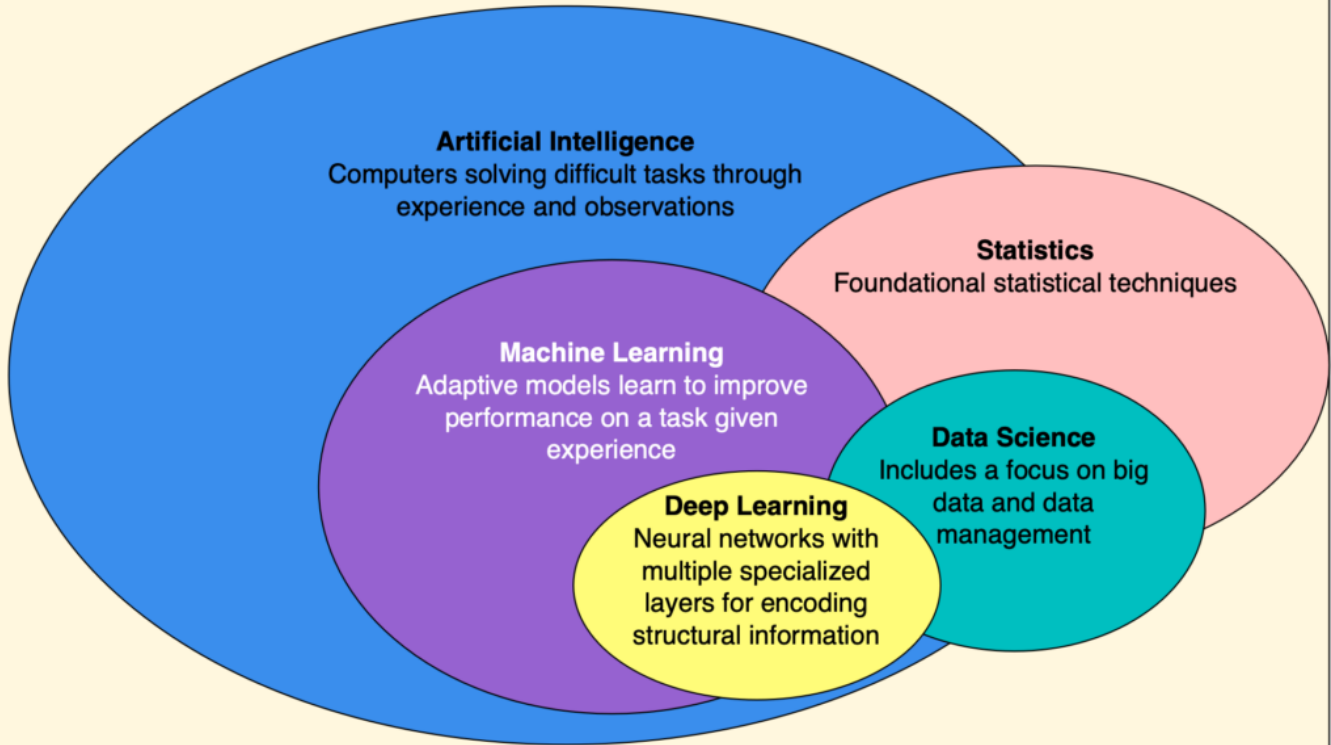
⬇

"The capability of a device to perform Functions that are normally associated with human intelligence such as reasoning, learning, and self improvement" (Virginia Tech)

⬇

"The simulation of human intelligence processes by computer systems, leveraging machine learning and deep learning using artificial neural networks. AI is typically trained on large volumes of data to perform a task at a level that approximates or exceeds human performance" (CHOP)

SRA
INTERNATIONAL

Behavioral science, psychology, social science, mathematics, many other disciplines

Artificial Intelligence
Computers solving difficult tasks through experience and observations

Statistics
Foundational statistical techniques

Machine Learning
Adaptive models learn to improve performance on a task given experience

Data Science
Includes a focus on big data and data management

Deep Learning
Neural networks with multiple specialized layers for encoding structural information

Practical uses of AI

Recommendation Systems	Large Language Models	Generative pretrained transformers	Bidirectional transformers
Natural language processing systems	Computer vision	Image processing	Chatbots
	Meeting transcription services	Language translational services	Virtual personal assistants

SRA
INTERNATIONAL

AI Technology Risks and Concerns

Safety

Welfare and safety of individuals and communities by the use of AI

Data Exposure

Ensure that organizational confidential information is not exposed and inappropriately used

Security

AI tools may be vulnerable to malware and compromised if not properly secured

Brand and Reputational

If misused, can subject the organization to brand and reputational risks

System Bias

Based on the information AI is told, the outputs may introduce systematic bias

AI Technology Risks and Concerns

Privacy and Compliance

Organizations to ensure AI tools comply with Federal, State, and organizational policies

Ethics and Rules of Conduct

AI tools should be responsible to promote trust, transparency, and equitable

Accuracy of Code and Outputs

A human should be involved to review and be fully accountable for any code created by AI technology

Research Integrity

Researchers should be aware of and follow policies and guidance of journal editors and funding agencies or can be subject to potential plagiarism or misconduct when using AI technologies when conducting research

Impact of AI on Research Administration

Data Ownership

- Understanding who owns and has rights to the data provided as input as well as the research data generated as output

Data Privacy and Protection

- Understanding how the information provided to or collected by these technologies is protected and who has access to the data
- Classify data as low-risk, medium- risk, and high-risk
  - Low Risk
    - intended for public disclosure
  - Medium Risk/ High Risk
    - not intended for immediate public disclosure, such as, data protected by HIPAA, GDRP, Export Control, IP, contract, or other agreements.
    - When working with collaborators, additional terms and conditions might need to be included in data use agreements to ensure responsible and ethical us of AI tools by collaborating researchers and organizations

Protection of Human Subjects

- Ensure research participants are fully informed that AI is being used
- Potential Risks and Limitations
  - Biased Data
    - Can lead to invalid conclusions resulting in harm to historically marginalized and vulnerable populations by perpetuating negative stereotypes, discrimination, racism, classism, sexism, and other problematic assumptions
  - Inaccurate Data
  - Secondary use of existing data
  - Privacy and confidentially of human subjects

Authorship, Plagiarism, and Reproducibility

Cybersecurity

Types of Cybersecurity

Malware	• Software designed to harm or exploit computer systems
Phishing	• Social engineering attacks that trick users into revealing sensitive information
Ransomware	• Type of malware that demands payment in exchange for restoring access to encrypted data
Denial of Service	• Attacks that overwhelm computer systems with traffic
Social Engineering	• Manipulating individuals into revealing confidential information and valuable data that may compromise like identify theft

SRA
INTERNATIONAL

Why is Cybersecurity Important?

Protecting
Sensitive Data

Preventing
Financial Loss

Ensuring
Business
Continuity

Maintaining
Customer Trust

Compliance with
Regulations

Protecting
Intellectual
Property

Reducing
Cybercrime

Cybersecurity Risks and Concerns

Increase of technology escalates cyber threats

Funding shortages and not having the resources to invest in advanced security tools

Fragmented IT systems such as decentralized IT systems

Managing a diverse user base and staff turnover to maintain strong protocols

Lack of Cybersecurity defenses

Awareness and training
Multi-Factor Authentication
Security Audits and Data Encryption

Research Compliance Requirements

Federal Acquisition Regulation (FAR)
- FAR 52.204-21: Basice Safeguarding of Covered Contractor Information Systems

Defense Federal Acquisition Regulation Supplement (DFARS)

National Security Presidential Memorandum 33 (NSPM-33)

Criminal Justice Information Services (CJIS)

Cybersecurity Maturity Model Certification (CMMC) Levels 1-5
- Level 1: Basic Cyber Hygiene
- Level 2: Good Cyber Hygiene

NIST
- NIST SP 800-171: Protecting Conrolled Unclassified Information in NonFederal Systems and Organizations

Cybersecurity Practices

Institution Owned Devices
- Patching and updating operating systems
- Effective antivirus, antimalware and antispyware software

Federal prohibitions on the acquisition and use of devices, software, and services

Research data should be collected, stored, processed, and archived on institutional devices, systems, and services. See your Data Protection Policies

Data Backup, Cloud Services, and Data Centers

Use organization email instead of personal email

Remote Workers and Cybersecurity Polices and Practices

Reassess data access permissions

New hires and Change of Roles/Responsibilities
Team Projects

SRA
INTERNATIONAL

Big Data

What is Big Data?

Extremely large and complex data sets that cannot be easily managed or analyzed with traditional data processing tools, like spreadsheets.

Examples of Big Data in Science
- Genomic Sequencing
- Medical Records
- Climate Simulations
- Human and Consumer Behavior

Structured Data
- Inventory Database
- Financial Transactions

Unstructured Data
- Social Posts
- Videos
- Mixed data sets

Five "Vs" of Big Data

Volume
- Refers to the immense quantity of data being generated from a range of sources such as sensors, satellites, and simulations

Velocity
- Captures the rapid rate at which this data is produced and needs to be analyzed in real time

Variety
- Encompasses the different formats of data

Veracity
- The truthfulness of the data such as quality and integrity

Value
- What is the benefit of the data and how can it be used to improve efficiencies or engagement

Big Data Benefits and Challenges

Benefits

- Better insights
- Decision Making
- Personalized Customer Experience
- Improved operational efficiency

Challenges

- Keeping pace with the data and ability to store it
- Curation of Data
- Data Security and Privacy
- Creating a data-driven culture
- Big Data Technology is changing at a rapid pace

Big Data Best Practices

Align big data with specific business goals

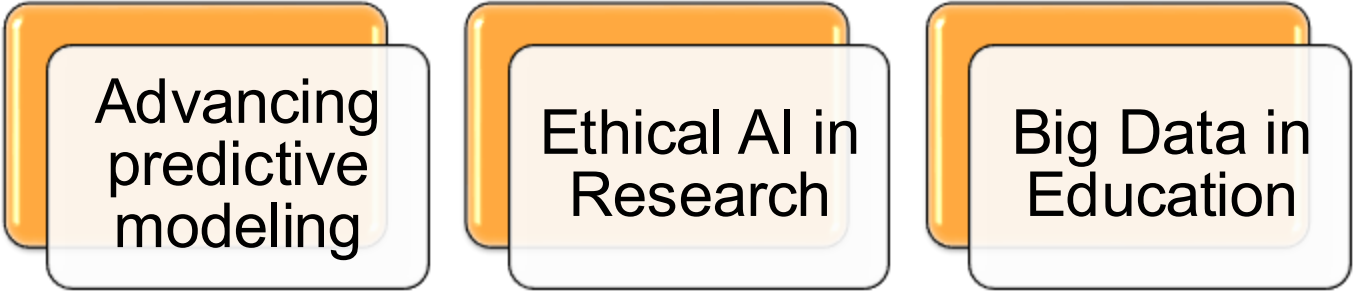Ease skills shortages with standards and governance

Optimize knowledge transfer with a center of excellence

Aligning unstructured with structured data

Plan a discovery lab for performance

Align with cloud operating model

Future of Big Data

Advancing predictive modeling

Ethical AI in Research

Big Data in Education

References

Chen, Michael. "What Is Big Data?" *Www.oracle.com*, 11 Mar. 2024, www.oracle.com/big-data/what-is-big-data/.

"Guidance: Using Artificial Intelligence during Research Activities." *Research.vt.edu*, www.research.vt.edu/research-support/forms-guidance/sirc/guidance-using-artificial-intelligence-during-research-activities.html.

Horvath, Karl. "Cybersecurity in Higher Education: Understanding the Current Landscape - Campus Consortium - a 501c3 R&D Foundation." *Campus Consortium - a 501c3 R&D Foundation - a 501c3 R&D Institution*, 25 Oct. 2024, campusconsortiumfoundation.org/blog/enhancing-cybersecurity-in-higher-education-key-strategies-for-defending-against-cyberattacks-copy/.

"Research Cybersecurity Baseline | Research Security." *Virginia.edu*, 2025, security.research.virginia.edu/research-data-security-compliance/research-cybersecurity-baseline. Accessed 18 Apr. 2025.

SRA
INTERNATIONAL

References

"Research Cybersecurity Compliance | UCI Information Security." *Uci.edu*, 2025, www.security.uci.edu/how-to/research-compliance/.

Saxena, Ayush. "Importance of Cyber Security: Benefits and Disadvantages." *Sprinto*, 19 Apr. 2023, sprinto.com/blog/importance-of-cyber-security/.

"The Impact of Big Data on Scientific Research | Penn LPS Online." *Upenn.edu*, 3 Mar. 2025, lpsonline.sas.upenn.edu/features/impact-big-data-scientific-research.

"What Does Cybersecurity Do? Understanding Its Role and Importance." *Apus.edu*, American Public University, 22 Jan. 2025, www.apu.apus.edu/area-of-study/information-technology/resources/what-does-cybersecurity-do/.
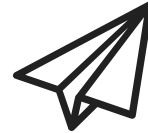
SRA
INTERNATIONAL

Questions, Comments, and Thank you!

Thank you for your participation!!!

Please reach out if you have any questions or feedback

Email:
Jonesm30@chop.edu