

---

# RIM/IG's Essential Role in Enterprise AI Adoption



# Agenda



- Nuts and bolts discussion of what AI is, hype vs reality, and what is needed for implementation success
- The role(s) that RIM/IG will play in the preparation, adoption and oversight of AI
- The emergence of a new sub-discipline to RIM/IG – AI readiness

# What even is AI?

A silver laptop with a teal icon of a chess knight and a chess rook on the screen, representing early AI applications like chess playing.

**Early Use Cases**  
OCR, chess playing,  
speech recognition, robotics

A silver laptop with a purple icon of a human head containing a white atomic symbol, representing neural networks.

**Neural Networks**  
Interactive  
search & retrieval

1956

2024

# What even is AI?



**Artificial Intelligence**  
Development of systems to perform tasks that typically require human intelligence.



**Machine Learning**  
Algorithms that enable computers to learn from and make predictions or decisions based on data.

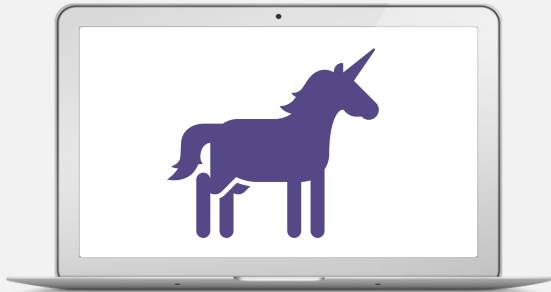


**Extractive AI**  
Extracts specific information from large datasets or text.



**Generative AI**  
Creates new content on the data it has been trained on.

# Gen AI reality vs. fantasy



## Trained on datasets

Learns the data to  
generate new content



## All-Knowing

Magical search engine,  
retrieving info from the internet



## Hallucinations

Takes a best guess, wants to  
give you what you're looking for



## Not up to date

Only as accurate as its  
last run of source data

# 5 Ways Orgs Will Want to Implement GenAI



## Accelerating Product Discovery & Development

Reduces time & cost, generate potential products and upgrades



## Enhancing Marketing and Quality Testing

Analyzing market data, predicting verticals, simulate buying scenarios



## Streamlining Regulatory Submissions & Compliance

Automate documentation, report generation, and data analysis



## Enhancing Products Support

Analyze customer-specific data, creating tailored solutions



## Improving Internal Knowledge Management & Data Utilization

Organize massive amounts of unstructured data, generate summaries

# What GenAI needs to Operate Properly



## Large bodies of data Content to train on

- Trusted Unstructured & structured data
- Respecting content access/permissions
- A way to get at the data in multiple systems
- Content organized by authenticity & relevance
- Clean data (data hygiene)



## Content that is unique Specialized/proprietary

- Unique content has the highest value
- Creates the most competitive advantage
- Example: Google Maps

# Deeper Dive on Access & Permissions

## Understanding and respect of content access/permissions:

- ✓ Who is allowed to query (or use) this content?
- ✓ Under what conditions?
- ✓ How is its use being tracked and/or audited?
- ✓ Who or what is the gatekeeper?
- ✓ Restrictions based on the content itself vs. on the user identity/role.

GenAI can be easy to trick without restrictions.



## Data Authenticity

Makes the data trustworthy

- Proper/appropriate disclosure (access) builds authenticity
- What should be included? (doc types)
- What should be excluded? (trade secrets)
- Identify redundant data (old versions)
- Identify personal data (patient/employee)
- Identifying unblinding data



## Identity & Permission

Who is asking? What can they see?

- What is the sensitivity of the data?
- Should sensitive docs be tagged?
- Event-based trigger access (ex. DSAR = employee records)

# Deeper Dive on Content Analysis

For GenAI to work,  
the starter/training content must be:


- ✓ Indexed (tagged)
- ✓ Scored/sorted (relevance)
- ✓ Served (in a manner the GenAI expects), and
- ✓ Gated (permissioned)



Repeat for every repository  
Repeat for every file

- Every file in every data repository needs to undergo these steps, or analysis
- Industry estimated most large organization have 15-20 currently unconnected systems
- Baseline indexing vs. delta indexing
- Consistency is key

# AutoClassifying Documents

Name	Date modified	Type	Size
 LGPIF Final.pdf	8/13/2016 9:26 AM	Adobe Acrobat Document	171 KB

# AutoClassifying a document

Informed Consent Form

Title of Study: Phase II Randomized Controlled Trial NeuroCure-X in Mild to Moderate Alzheimer's Disease

Principal Investigator: Dr. Jane Researcher

**Introduction:**

You are being invited to participate in a research study to evaluate the safety and efficacy of the investigational drug NeuroCure-X in treating mild to moderate Alzheimer's disease. Before you decide whether to participate, it is important that you understand the purpose, procedures, potential risks, and benefits of the study.

**Purpose of the Study:**

The purpose of this study is to determine whether NeuroCure-X is effective in improving cognitive function in individuals with mild to moderate Alzheimer's disease.

**Procedures:**

If you agree to participate, you will be randomly assigned to either receive NeuroCure-X or a placebo. You will be required to take the assigned treatment daily for a duration of 12 months. During this time, you will undergo regular cognitive assessments, medical evaluations, and safety monitoring.

**Risks and Benefits:**

Potential benefits may include improved cognitive function. Risks may include side effects associated with the study drug. Common side effects are whatever.

**Confidentiality:**

Your personal information will be kept confidential, and only authorized individuals involved in the study will have access to your records.

**Voluntary Participation:**

Participation in this study is entirely voluntary. You may withdraw at any time without penalty or loss of benefits.

**Questions:**

You are encouraged to ask any questions you may have before deciding to participate. Dr. Jane Researcher can be contacted at 736-831-9833 for any further information.

**Consent:**

I have read and understand the information provided in this Informed Consent Form. I have had the opportunity to ask questions, and my questions have been answered to my satisfaction. I voluntarily agree to participate in the study.

Participant's Name: Jennifer Nelson

Participant's Signature:  Date: Dec 14, 2023

Investigator's Signature:  Date: Jan 5, 2024

DocType = Informed Consent Form

Drug Name = NeuroCure-X

Trial = Phase 2 Randomized Controlled Trial

Investigator = Dr. Jane Researcher

Participant = Jennifer Nelson

Date Executed = Dec 14, 2023  
[Date format → 12/14/2023]

RecordType = Consent Form  
RecordClass = Clinical Trial  
Retention = Executed + 8Y  
Sensitivity = Confidential  
Handling = Delete after authorization

Retain.  
Delete on Dec 14, 2031

# Reducing Risk & Exposure

- ✓ Uncover dark data
- ✓ Inform & action content
- ✓ Ensure compliance



## Data Governance Risk

Where is it? What is it?  
Is it being handled properly?



## Privacy Risk

Does it contain personal  
or sensitive data?



## Compliance Risk

Onside with your own policies &  
regulatory obligations?



## AI Risk

Source data clean & trusted for  
Generative & Extractive AI?

# Key Things the IG Team Can Do to Promote Good GenAI Outcomes

What is the impact of training the algorithms on wrong or old data?



## Stress the need for good data hygiene

Remove ROT = remove low-value content that will screw up your LLM training



## Make sure the GenAI algorithm matches your RIM policies

First instance of a dupe, last/most recent version, most trusted source?



## Establish protocols for how data should be tagged & served to the GenAI engines

Taxonomies, field tables, RRS's are all good



## Establish what GenAI results auditing must cover

What will you, as RIM/IG professionals, need to see to be convinced the "right things are happening"?

# Key Things the IG Team Can Do to Promote Good GenAI Outcomes



## Establish Robust Data Governance Frameworks

Clear data ownership, classification, and lifecycle management policies in place: access, retention, storage



## Develop Clear Policies for Ethical AI Use

Create AI governance frameworks that address bias mitigation, transparency, and accountability



## Address Data Privacy & Security Concerns

Update policies, ensure data used for AI training is anonymized, properly consented, track data usage.



## Create Guidelines for Model Transparency and Explainability

Establish guidelines for explainability: how decisions or outputs are generated, data lineage, human-validated

# Key Things the IG Team Can Do to Promote Good GenAI Outcomes



## Prepare for Intellectual Property and Legal Challenges

Policies for clarifying ownership and use of outputs created by GenAI, licensing & copyright compliance



## Train Employees on AI Literacy and Governance

Training on AI literacy, focusing GenAI use within organization, data governance policies, impact on roles



## Monitor and Audit AI Outputs for Compliance and Quality

Update policies, ensure data used for AI training is anonymized, properly consented, track data usage.

# Roundtable Discussion Topics



- ✓ What is your org's timeline for implementing AI?  
What details of the rollout plan can you share?  
(Timing, business units, sequence, use cases, etc.)
- ✓ What threats or concerns do you have?
- ✓ Is your organization including you/voice of RIM/IG in its plans?
- ✓ What can we produce as joint/cross-industry best practices and/or documentation?  
What standards/recommendations can we publish?  
(Any interest in meeting on this topic regularly and offline from today?)