

# AI & AutoClassification for Structured Data Management & Disposition





— WEBINAR SERIES —



Sandra Serkes  
Co-Founder & CEO



Jennifer Nelson  
VP Strategic Solutions



# Webinar Series



[www.ValoraTech.com](http://www.ValoraTech.com)



ROT Remediation:  
file share clean-up  
February 12



Information Governance &  
Enterprise AI Adoption  
June 11



Managing Data Privacy  
while Managing Records  
March 19



AutoClassifying the 3 Rs:  
ROT, Records & Retention  
September 17



The Role of InfoGov  
in AI Readiness  
April 16



Getting Ready for  
AutoClassification  
October 15

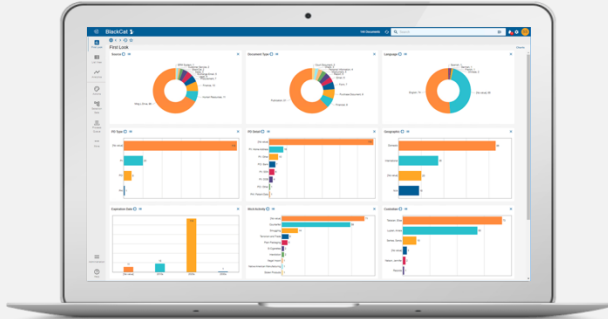


AI & AC for Structured  
Data Management  
May 14



Demo Day  
November 12

# About Valora



- Information Governance platform
- Automates **discovery**, **identification**, **classification** & defensible disposition
- Trusted by IG teams, Records & IM teams, Legal, Compliance, IT teams
- Brings multiple enterprise repositories into a single view

# About Valora



## Information Governance / Records Management



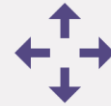
Classification



ROT



Retention



Disposition



Archive

# About Valora



## Enterprise-wide Data Governance



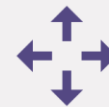
Classification



ROT



Retention



Disposition



Archive



eDiscovery



Legal Hold



Data Privacy



Data Security



Migration



Compliance



Minimization



Provenance



Lineage



AI Readiness

# Poll Results

---

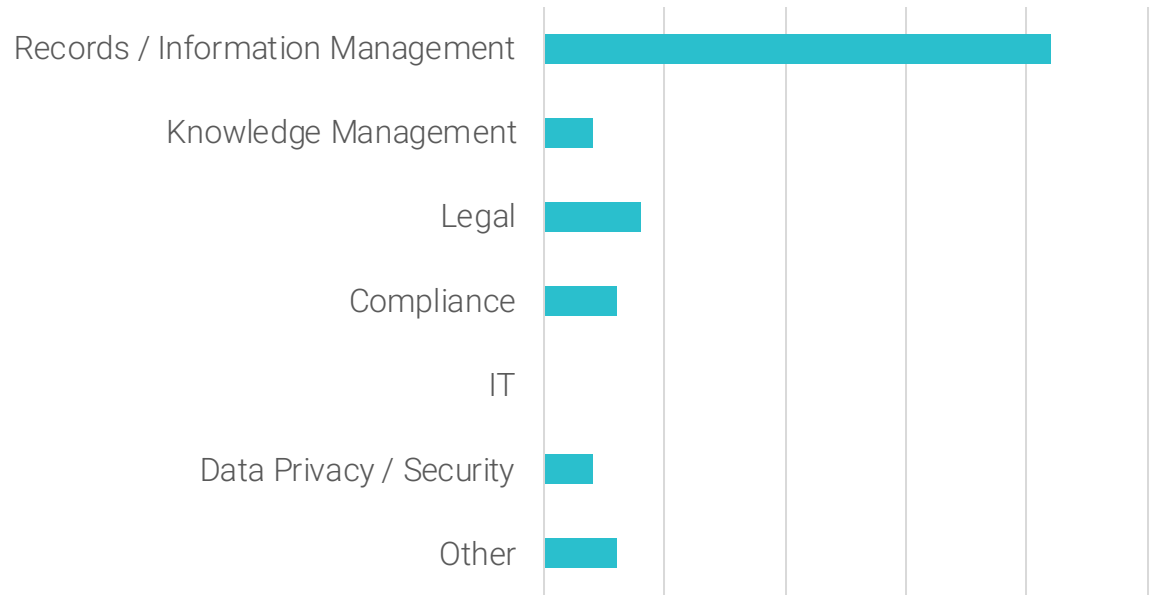
Who's here today  
& what are your interests?



# Who you are & what your challenges are



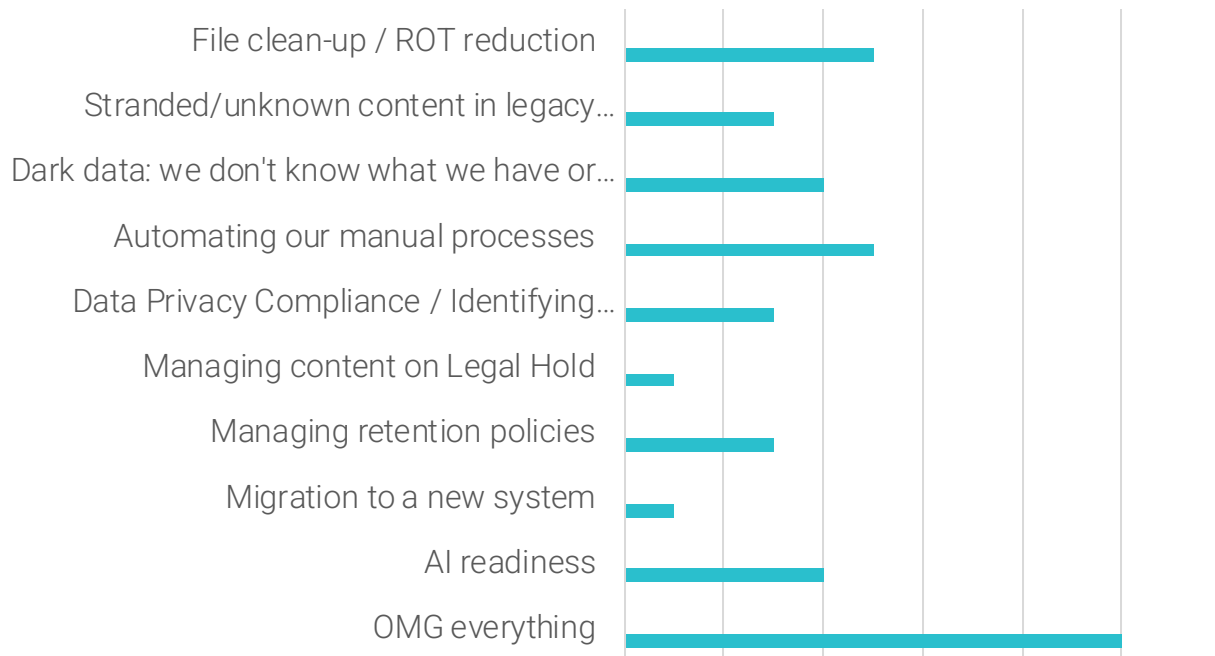
Who You Are



# Who you are & what your challenges are



## Your RIM & IG Challenges



# Housekeeping



Q&A



Recording



Slides



Feedback

# Agenda



- Responsibilities & challenges of managing data stored in siloed databases and semi-structured formats
- What are AI & AutoClassification and their roles in data analysis
- Transforming structured data values into user-friendly and actionable “Mockuments”
- How AI & AutoClassification can help IG/IM teams improve efficiency, compliance and decision making
- Use Case: data requests from structured data

# Structured Data Systems



IDC's annual study (2023)

“Structured Databases/Data Management workloads drive the largest share of enterprise IT infrastructure spending, at **\$6.4 billion of compute and storage hardware infrastructure** to support this workload category.”

# Structured Data Systems



## Isolated Systems

Structured & semi-structured data are often isolated system, making retrieval difficult



## Volume & Complexity

Sheer volume and complexity of data overwhelms traditional governance methods



## Lack of Context

Standalone fielded data can lack context without other data and/or supplementary content

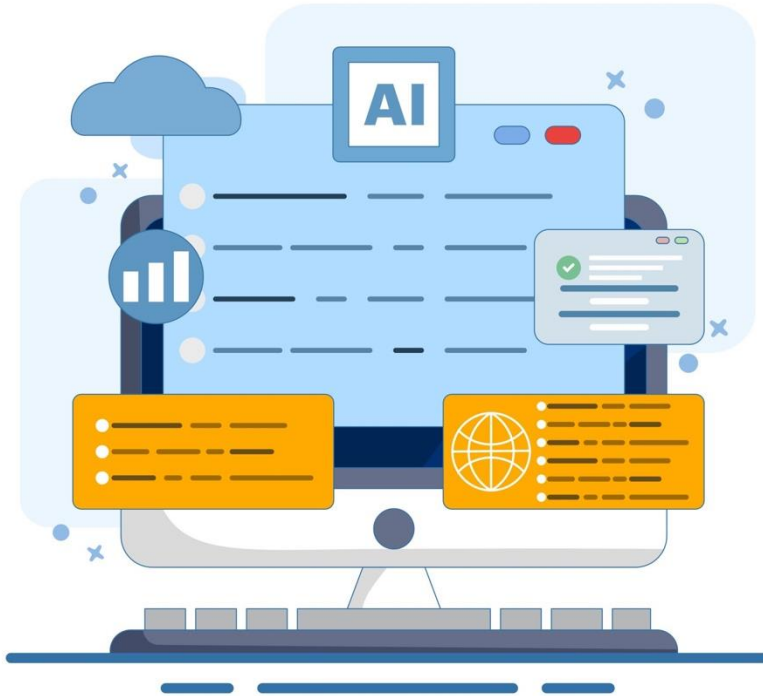


## Compliance Challenges

When responding to DSARs, audits or legal holds, retrieving accurate data quickly is a challenge



# The Role of AutoClassification & AI



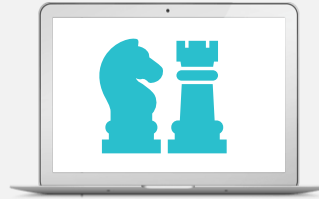
- Use AI to **automatically analyze and categorize** data based on content, context, and metadata
- Create a "single pane of glass," allowing users to search, view, and **manage structured and unstructured data in one place**.
- AI **enhances searchability** and ensures relevant data is surfaced, regardless of its original format or source.

# What is AI?

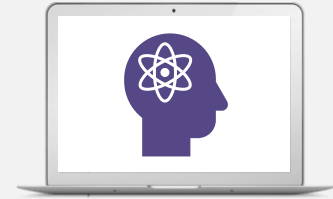


## Definition:

Artificial intelligence refers to the simulation of human intelligence by machines, including learning, reasoning, and self-correction.



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



**Neural Networks**  
Interactive  
search & retrieval

1956

2025

# Types of AI



## Machine Learning

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



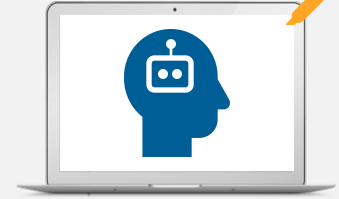
## Extractive AI

Extracts specific information from large datasets or text.



## Generative AI

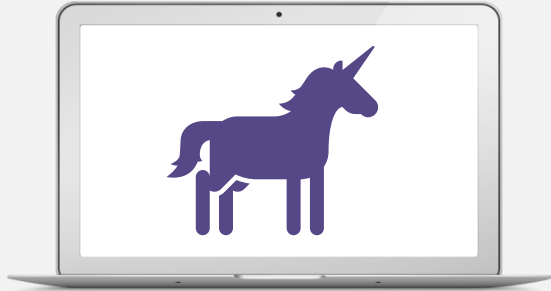
Creates new content on the data it has been trained on.



## Agentic AI

Operates independently, making decisions & performing tasks without human intervention.

# Gen AI reality vs. fantasy



## Trained on datasets

Learns the data to  
generate new content



## Not magical search engine

Creates new content based on  
learned patterns



## Hallucinations

Takes a best guess based on  
training data, does not fact check



## Not up to date

Only as accurate as its  
last run of source data

# AutoClassification is a type of AI



## Machine Learning

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



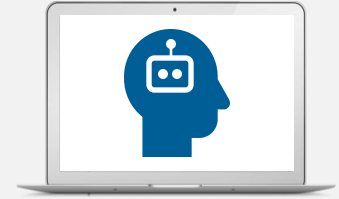
## Extractive AI

Extracts specific information from large datasets or text.



## Generative AI

Creates new content on the data it has been trained on.



## Agentic AI

Operates independently, making decisions & performing tasks without human intervention.

# What is AutoClassification?



File & Rich Metadata  
Determine Attributes

+



Recognition Algorithms  
Determine DocType

+



Disposition Algorithms  
Rules & Actions

AutoClassification = Rich Metadata + Rules

Software that performs automated analysis & disposition of file/document content



Answers the question:

What is this thing & what do I do with it?

...and then does it.

# File Attributes become Rich Metadata



File & Rich Metadata  
Determine Attributes

+

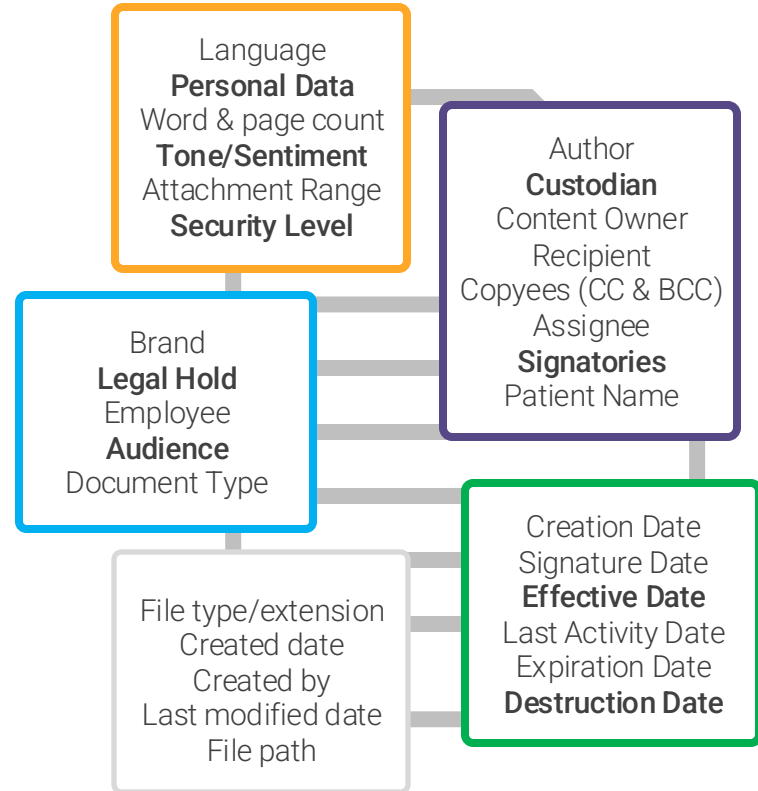


Recognition Algorithms  
Determine DocType

+



Disposition Algorithms  
Rules & Actions



# DocType as defined by RRS



File & Rich Metadata  
Determine Attributes

+

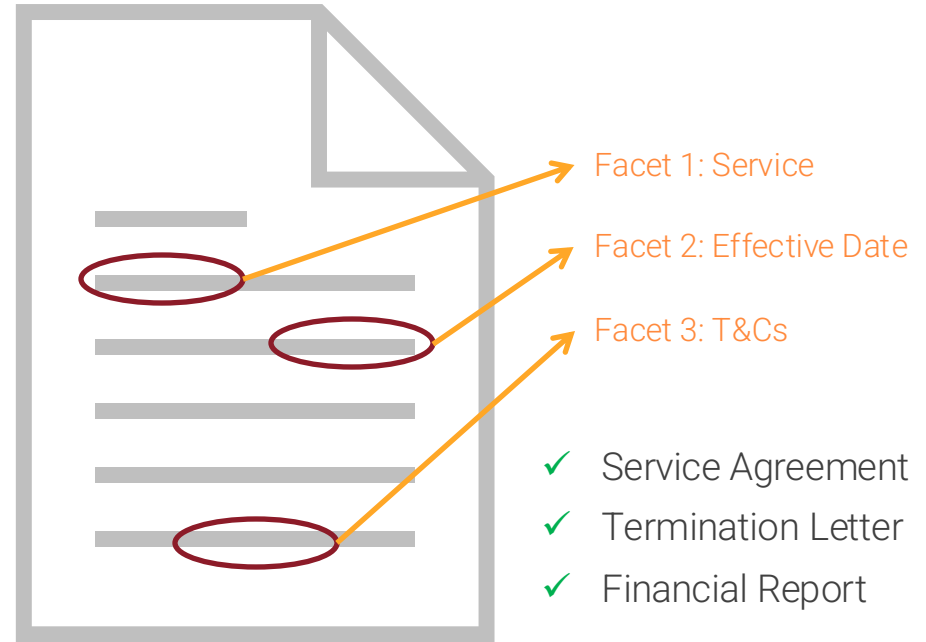


Recognition Algorithms  
Determine DocType

+



Disposition Algorithms  
Rules & Actions



# Defensible Disposition



+



+



- Determine the disposition & handling of content
- Typically follow an IF-THEN format, often nested
- Ex: Employment Contract:
  - ✓ Mark as Confidential
  - ✓ Move it to SharePoint
  - ✓ Retain for 8 years after termination
  - ✓ HR approval required before deletion

## Automated, Manual and/or Hybrid Disposition



Tag



Delete



Report



Move



Archive



Hold




Migrate



Anonymize

# AutoClassifying Documents

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

# AutoClassifying a document

PHARMACEUTICAL

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



# Visualizing Structured Data with “Mockuments”



## What is a “mockument”?

- Solution for making structured data more interpretable by transforming it into document-like formats
- Makes data accessible to non-technical stakeholders, improving visibility, collaboration and decision-making.

# Visualizing Structured Data with “Mockuments”

Export data or  
process-in-place  
from database  
or data lake



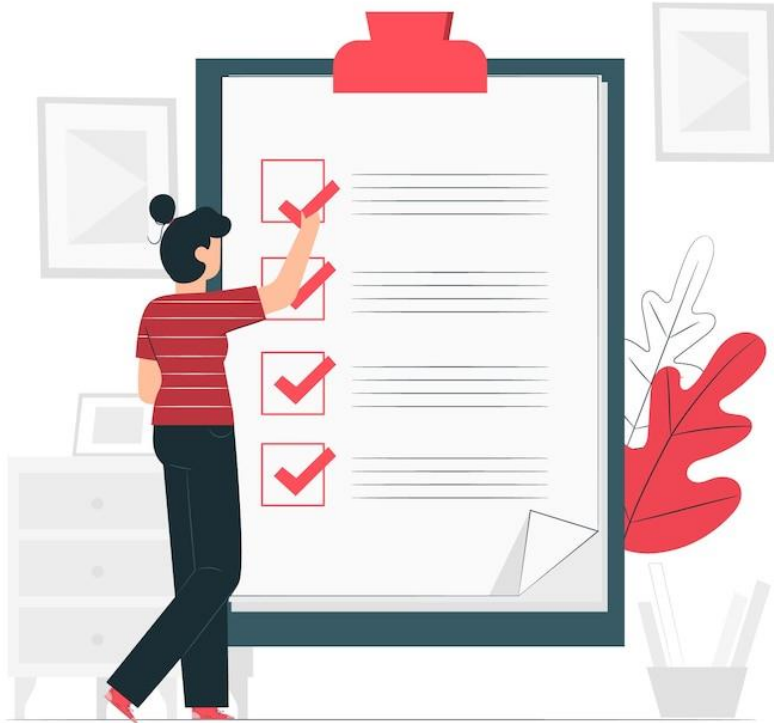
Database exports to  
JSON, CSV, XML, etc.

	A	B	C	D
1	Name	Company	Email	Phone Number
2	Caleb West	Diam Inc.	sapient.molastie@acmattisomnare.edu	101.180.8138
3	Harper Hines	Libero Mauris LLC	sed.libero.Proin@Aliquaminciduntunc.net	102.322.8325
4	Chandler Wood	Aliquet Inc.	at.libero@orcim.co.uk	786.867.8273
5	James McKinney	Nisi Quisque Ltd	vehicula.Pellentesque@auctorMaurisvel.com	286.843.7071
6	Deborah Chandler	Sit Amet Diam PC	at.pede.Cras@torquentperconubia.com	411.105.8745
7	Xena Joseph	Ac Turpis Egestas Industries	Quisque.varius.Nam@nec.edu	307.983.8017
8	Patricia Fitzgerald	Nulla Integer Associates	vestibulum@maurisque.org	987.637.7059
9				
10	Len Barry	Phasellus Limited	sapient.gravidia@sollicitudinmalesuada.ca	262.563.3418
11	Jorden Jacobson	Sodales LLC	Proin.dolor.Nulla@Donec.ca	871.304.7789
12	Cleo Camacho	Tristique Ac PC	Phasellus.vitae.mauris@vitaeinibhDonec.edu	486.250.9502
13	Morgan Pitts	Eget Industries	torror@tempor.net	441.873.9538
14	Sigourney Sutton	Et Company	Nunc.mauris@arcuCurabitur.com	415.878.9886
15	Norman Bauer	Dictum Eu Industries	mus@torquent.ca	350.369.2914
16	Kameko Joyner	Et LLP	quis.pede.Suspendisse@lobortisrius.edu	676.439.7761
17	Elvis Kelley	Vulputate Posuere Associates	sapient@parturientmontes.edu	700.497.1614
18	Oleg Meyer	Donec Fela Orci Inc.	avitor.nunc.Nulla@parturientmontes.com	212.813.5628
19	Xander Cochran	Morbi Tristique LLC	nibh.dolor@Nuncacsem.org	591.327.8919
20	Kato Small	Eu Nibh Ltd	erat@erosnec.co.uk	784.335.9208
21	Emerson Hendricks	Ultrices Sit Company	nisi.Cum@liberoest.ca	121.557.7052
22	Gary Frost	Mauris LLP	turpis.egestas.Aliquam@liberoMorbi.ca	747.393.2416
23	Quinn Leon	Neque Et Incorporated	Mauris.eu@urnaNullam.co.uk	428.502.4396
24	Cameron Snow	Eu Placerat Eget Industries	senectus@massaQuisque.edu	124.734.5169
25	Acton Avila	Dis Parturient Montes Limited	Nulla.magna.malesuada@Donechonjusto.ca	915.455.8712
26	Charles Conley	Commodo PC	massa.Mauris@turpis.co.uk	735.764.8318
27	Raphael Mays	Mauris Corp.	per.inceptos.hymenaeos@vivamusrhoncusDonec.edu	925.191.0656
28	Tanner Harrell	Diam Duis Associates	Mauris@sagittisNullamvitae.edu	103.647.7574
29	Henry Villarreal	Lectus Inc.	fermentum.fermentum@ipsum.edu	879.841.4865
30	Kim Robinson	Orci Associates	diam.am@nibhQuisque.edu	418.499.1787
31	Evan Hull	Morbi Industries	lobortis.ultrices@etipsummaurus.ca	194.265.9732
32	Winter Parrish	Augue PC	suscipit.nonummy.Fusce@loremut.co.uk	647.125.2722

Create a “Mockument”  
User-friendly view of a single  
database record

<b>First Name</b>	Patient
<b>Last Name</b>	Sally
<b>MI</b>	G
<b>Patient ID</b>	12345
<b>DOB</b>	1980-01-04
<b>SSN</b>	123-45-6789
<b>Street Address 1</b>	101 Fake Street
<b>Street Address 2</b>	Apt 2
<b>City</b>	Springfield
<b>State</b>	MA
<b>Zip Code</b>	11014
<b>Phone</b>	123-456-7890
<b>Mobile</b>	234-567-8910
<b>Email</b>	spatient@gmail.com
<b>PCP</b>	Smith, Robert
<b>Patient Since</b>	2004-05-11
<b>Last Visit</b>	2021-07-23

# Real-World Applications



Q: What do mockuments allow you to do?

A: Visualize structured data from multiple structured sources in “dossier format”

- Respond to data requests:
  - past/present employee,
  - legal hold requests
  - DSAR requests(provide the data, not the source doc)

# Benefits of Unified Data View & Management



Inventory & Classification



Data quality



Standardization



Data retrieval

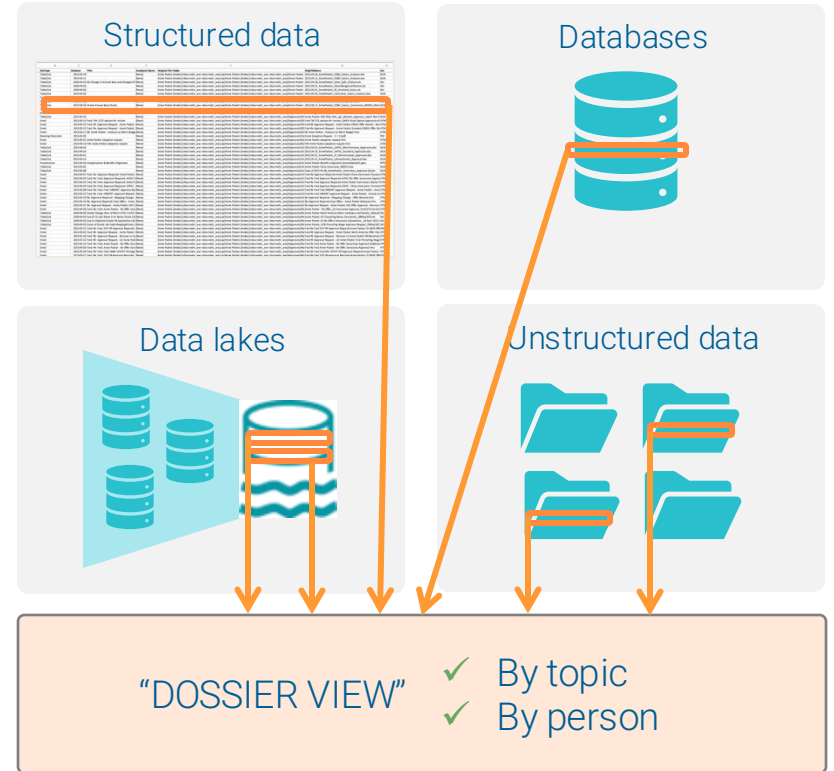


Transparency



Regulatory Compliance

# Benefits of Unified Data View & Management



# Benefits of AI-driven Approaches



Improved accuracy  
by minimizing  
human error



Saves time  
automate tagging &  
metadata management



Enhanced security  
automatically detect  
sensitive data



Scalability  
keep up with  
rapid data growth

Not replacing humans, enhancing their capabilities

# Case study: “Mockuments” in Action



## Global Organization

- 25+ years active
- 150 acquisitions
- Millions of inherited files
- 250,000+ employee records
- 20TB+ data



Data relevancy:  
what to keep, what to delete



People who know the data  
are no longer around



Out of compliance with our  
own data retention policies

## How to:

Execute on formal  
records retention  
policies?

Integrate with new  
systems of record?

Maintain & respect  
Legal Holds?

# Case study: “Mockuments” in Action

## Database data

<b>First Name</b>	Jeffrey
<b>Last Name</b>	Roberts
<b>MI</b>	K
<b>Employee ID</b>	12345
<b>DOB</b>	1980-01-04
<b>SSN</b>	123-45-6789
<b>Street Address 1</b>	101 Fake Street
<b>Street Address 2</b>	Apt 2
<b>City</b>	Springfield
<b>State</b>	MA
<b>Zip Code</b>	11014
<b>Phone</b>	123-456-7890
<b>Mobile</b>	234-567-8910
<b>Email</b>	jroberts@acmeacom
<b>Title</b>	Quality Assurance Engineer II
<b>Supervisor</b>	Taylor, Martin
<b>Salary</b>	\$125,000
<b>Hire Date</b>	1998-11-15
<b>Pre- ACME Termination Date</b>	--
<b>ACME Acquisition Date</b>	2005-04-01
<b>ACME Termination Date</b>	2012-08-17

Database-stored  
information about  
Employee

## Files & Attachments

Filename	Document Type	Source	Date
JKR CV.pdf	Personnel	GSI	1998-10-01
Promo-Roberts.doc	Correspondence	GSI	2002-06-01
NDA.doc	Agreement	GSI	1998-11-15

Stored files  
about Employee

## Employee Dossier: Roberts, Jeffrey K



- Create an **Employee Dossier**, a single report of all known employee data, with links to all known employee files
- A “single pane of glass” UI to view, manage, and action all data & information

# Case study: “Mockument” becomes a record

<b>First Name</b>	Patient
<b>Last Name</b>	Sally
<b>MI</b>	G
<b>Patient ID</b>	12345
<b>DOB</b>	1980-01-04
<b>SSN</b>	123-45-6789
<b>Street Address 1</b>	101 Fake Street
<b>Street Address 2</b>	Apt 2
<b>City</b>	Springfield
<b>State</b>	MA
<b>Zip Code</b>	11014
<b>Phone</b>	123-456-7890
<b>Mobile</b>	234-567-8910
<b>Email</b>	spatient@gmail.com
<b>PCP</b>	Smith, Robert
<b>Patient Since</b>	2004-05-11
<b>Last Visit</b>	2021-07-23

## Mockument becomes a record

- ✓ Tied to applicable Record Class & corresponding retention rule
- ✓ Caveat: data aggregated into each Mockument must belong to the same Class
- ✓ We simplified and consolidated a few Classes to make this more straightforward, then defaulted retention to the longer rule.

# Challenges & Considerations



- **Data Quality:** AI is only as effective as the data it processes—clean data is critical.
- **Integration Complexity:** Seamless integration with legacy systems can be challenging.
- **User Adoption:** Ensuring teams understand and trust AI-driven outputs is vital.
- **Regulatory Compliance:** AI solutions must comply with data privacy laws (e.g., GDPR, HIPAA).

# Key Takeaways



- AutoClassification addresses challenges in managing structured, semi-structured and unstructured data
- Unified views improve efficiency & transparency
- Tools like mockuments make data accessible to diverse teams
- AI solutions drive scalability, accuracy, and better decision-making



[www.valoratech.com](http://www.valoratech.com)



Book a Demo



Webinar Series



[/valora-technologies](https://www.linkedin.com/company/valora-technologies)

## Q&A

---

Thank you



**Sandy Serkes**  
President & CEO  
[sserkes@valoratech.com](mailto:sserkes@valoratech.com)



**Jennifer Nelson**  
VP Strategic Solutions  
[jnelson@valoratech.com](mailto:jnelson@valoratech.com)