



AutoClassifying the 3 R's: ROT, Records & Retention



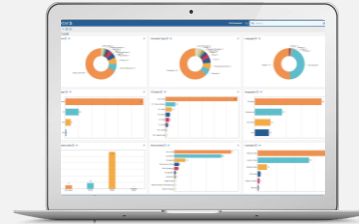
Sandra Serkes
Founder & CEO

Valora Technologies

Information Governance platform
leveraging proprietary
**AutoClassification technology &
machine learning (AI) algorithms.**

Automates the
**discovery, identification,
classification & defensible disposition**
of content across the enterprise.

Trusted by
**Data Governance, Records Managers,
Legal, Compliance & IT teams**
to defensibly govern their data.



360° View Single pane of glass

- Un/structured, on-prem/cloud, SaaS
- Identify, classify, defensibly disposition



Methodology Proven Process & Best Practices

- 20+ year pioneers in InfoGov
- In-house Professional Services team

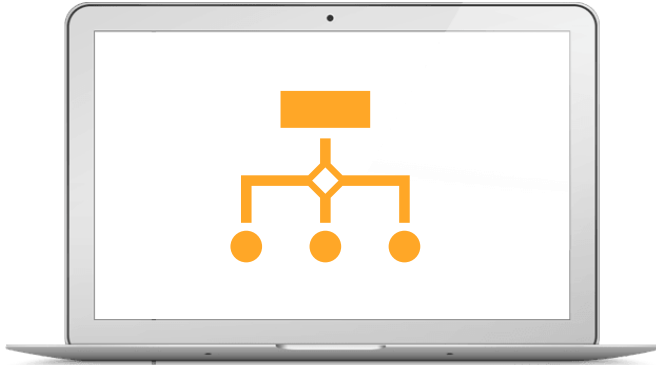


Technology Enterprise-grade SaaS platform

- Flexible, scalable, customizable
- True AutoClassification engine

Use Cases

Information Governance



Reduces data to records of business value, rich metadata attribution & rules

- ROT Processing
- Rich Metadata Attribution, AutoClassification by DocType
- Records Management
- Workflow Automation
- System Migrations

Compliance



ID files that contain personal data, mitigate risk, comply with regulations

- Privacy Regulations / DSAR
- Data Discovery
- Defensible Risk Mitigation
- Regulatory Compliance
- Reporting

Legal / Finance



Reduces effort, reduces turnaround, defensible decisioning and audit trail

- Early Data Assessment: data set reduction / eDiscovery
- Legal Hold
- Virtual Vault contracts management
- Mergers, Acquisitions & Divestitures: DD, Post-Merger Integration (PMI)

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?

Agenda – the 3 R's



- AutoClassification 101
- How to AutoClassify & Defensibly Remove **ROT**
- AutoClassifying **Records** the Right Way
- Putting it all together: **Retention** & lifecycle management with AutoClassification
- How to get started



AutoClassification 101

Everything you need to know
about this exciting technology
in 5 minutes or less!



What is Data Classification?

Creating tags that tell us about a file



PDF

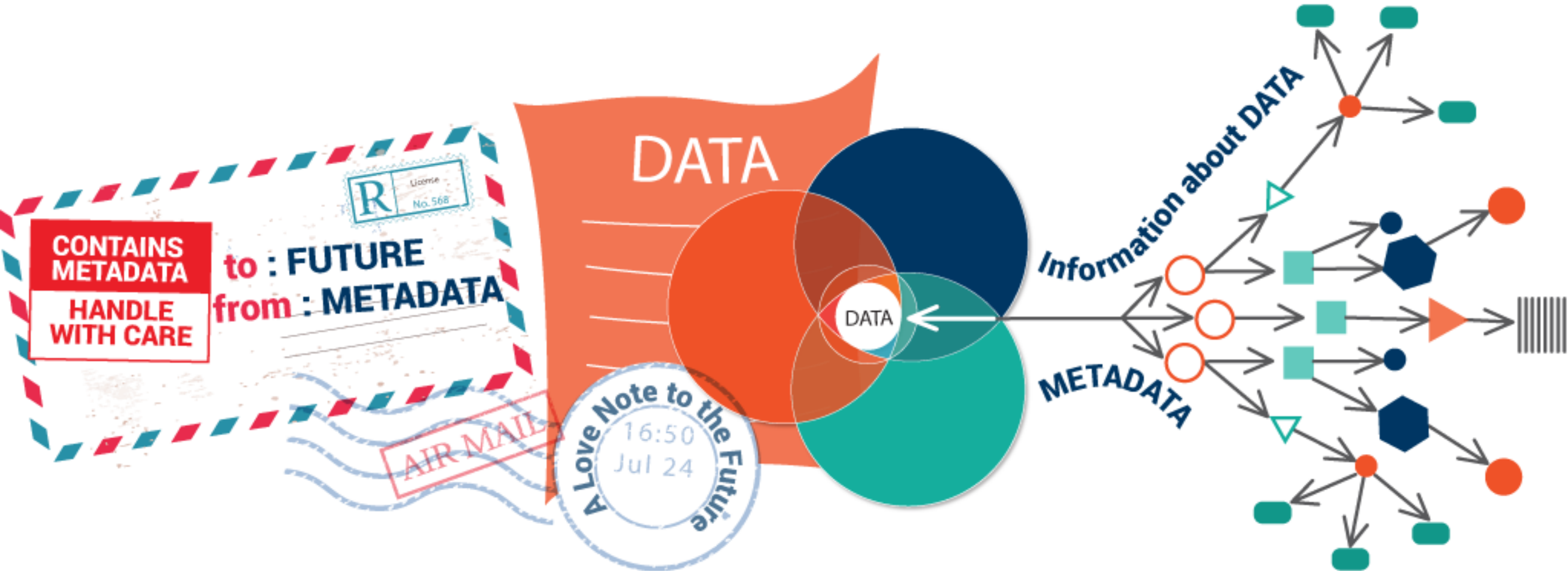
Created date

Contract

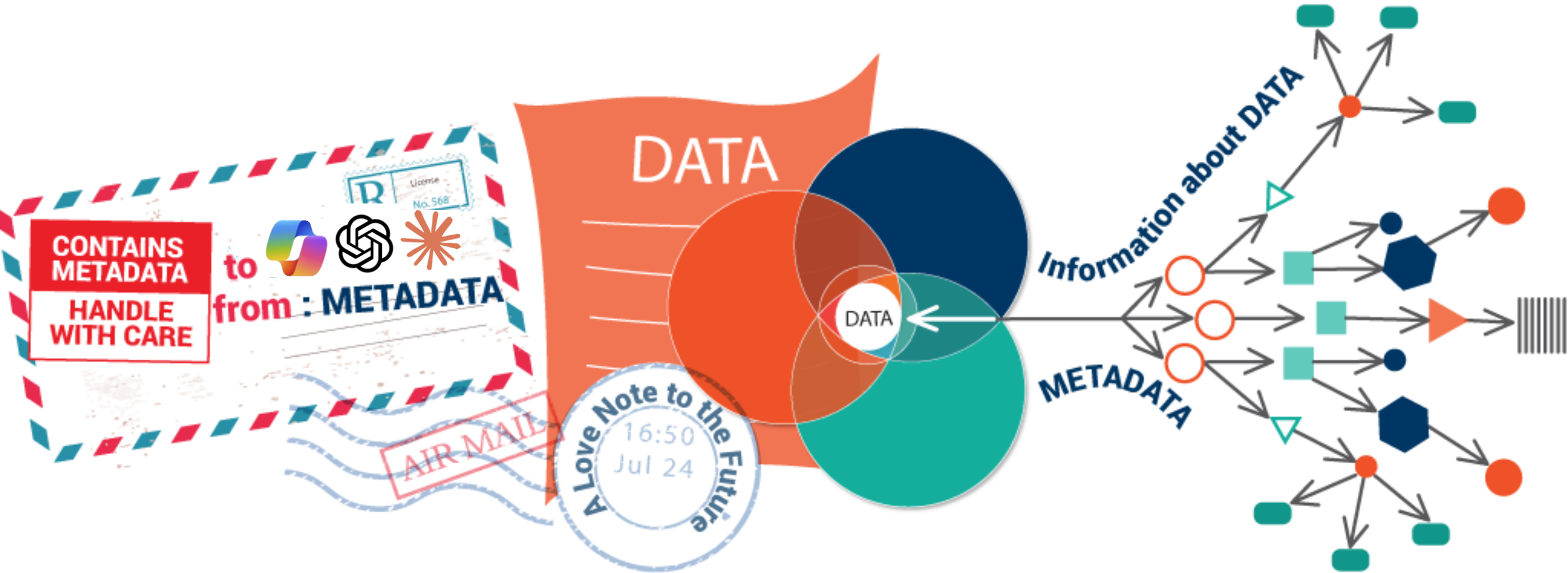
Expired

METADATA

Metadata is a “love note to the future”



Metadata is a “love note to AI”



You CAN classify manually...

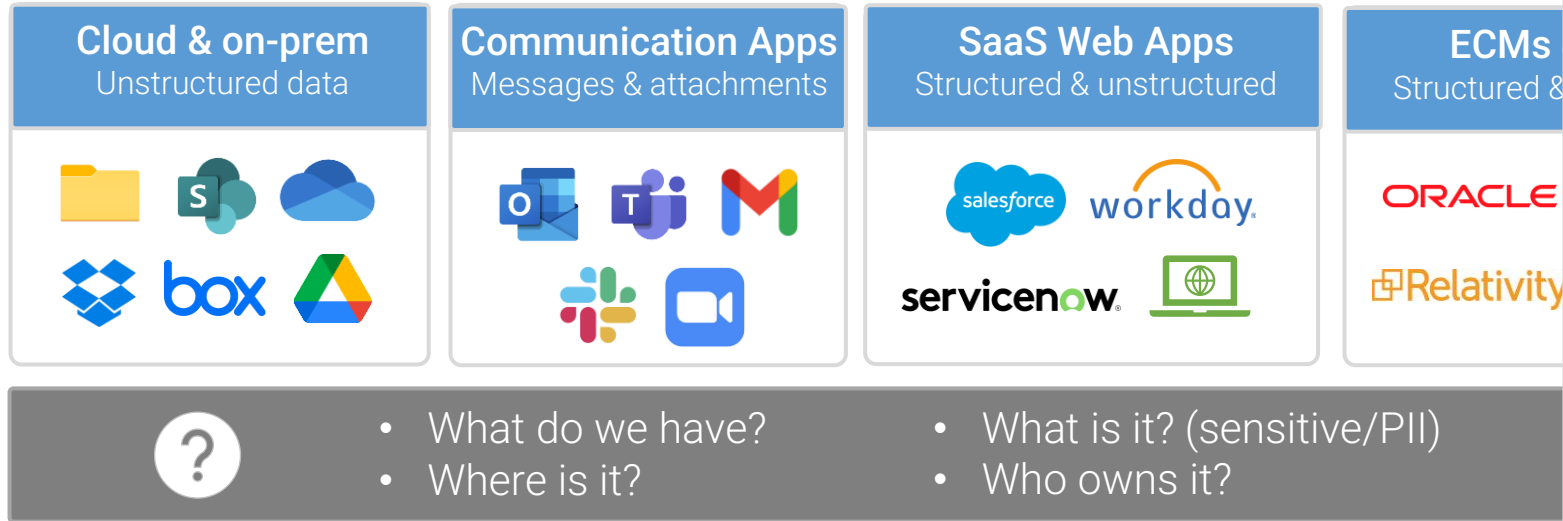
But how will you ever keep up?

Ultimately, is that the best use of your organization's time and resources?

(And how will the quality be?)



Humans cannot keep up



Human Error

Inconsistency, interpretations/
reluctance, "not my job"

Scalability Issues

Impossible to keep up with
volume of data

Compliance Risk

Regulatory non-compliance,
retention & deletion errors

Limited Auditability

Lack of transparency, why
things were categorized

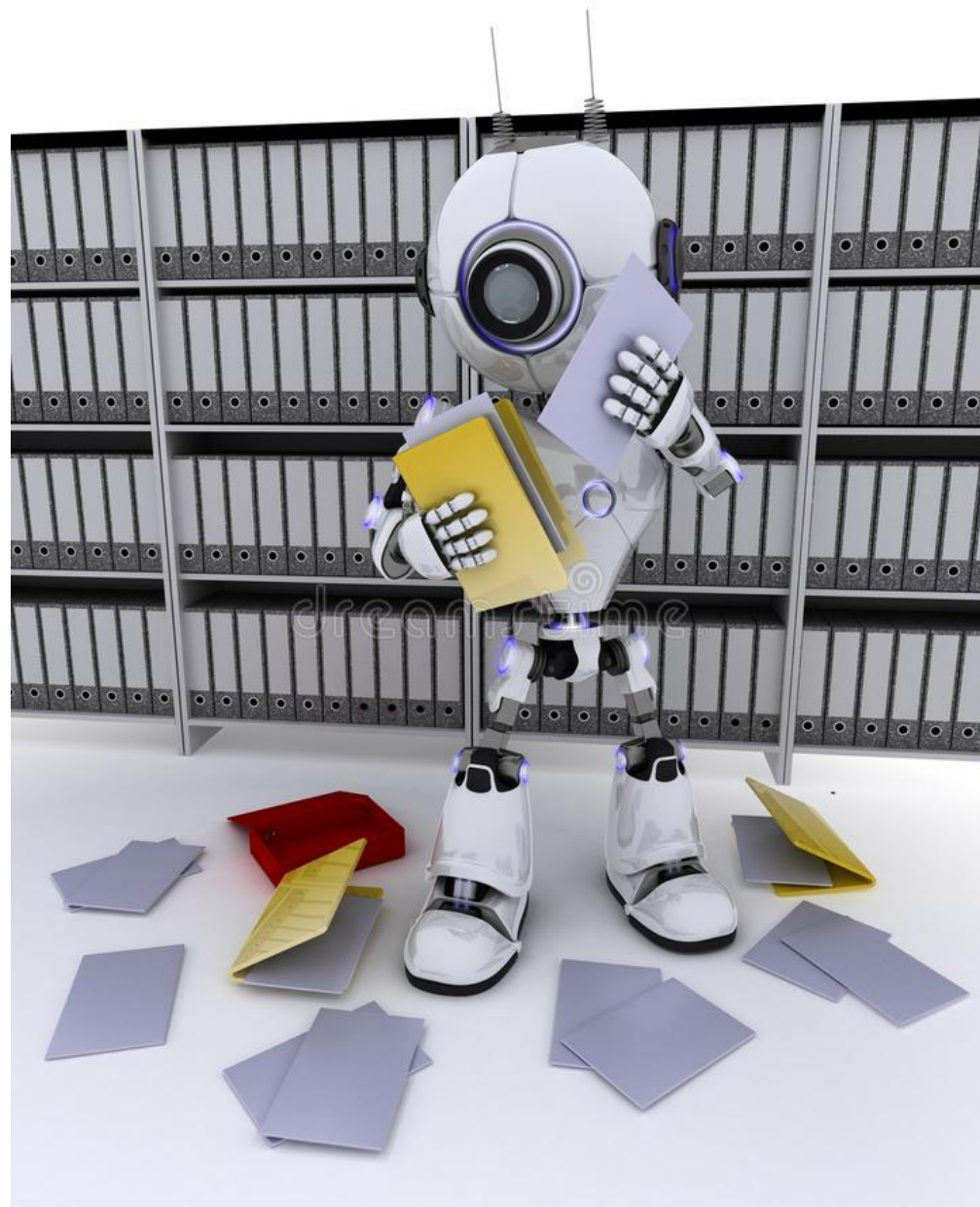
Time Consuming & Costly

- 30 seconds = employee to manually classify 1 doc
- 1M docs
 - = 30M seconds / 8000 hrs ÷ 2,080 work hrs/year
 - = 3.8 employee-years
 - = 1 employee manually classify 264k docs/year
- Average employee in a business:
 - = sends/receives 100 emails, works on 10 docs/day
- 1000-person organization
 - = 110,000 doc/day
 - = 1.32M docs/year
- Been in business 20 years
 - = backlog of 438M docs
 - = 1,664 employee-years to get through the backlog
- Get it done over 5 years = 87M docs/year
 - = 264k docs/year/person
 - = require a team of 329 humans
- US @ \$65k/year average salary
 - = \$21.3M in wages per year
 - = \$106.9M in wages over 5 years
- Overseas @ \$5k/year average salary
 - = \$1.65M in wages per year
 - = \$8.25M over 5 years
- That's just to get through the backlog

AutoClassification

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

A note about **true**
AutoClassification – if you
(or your colleagues) are
doing the work, it isn't
AUTO!



True AutoClassification



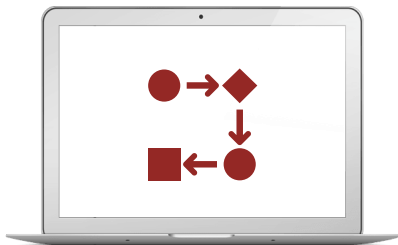
File & Rich Metadata
Attributes

+



Recognition Algorithms
Determine DocType

+



Disposition Algorithms
Rules & Actions

AutoClassification = Rich Metadata + Rules

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

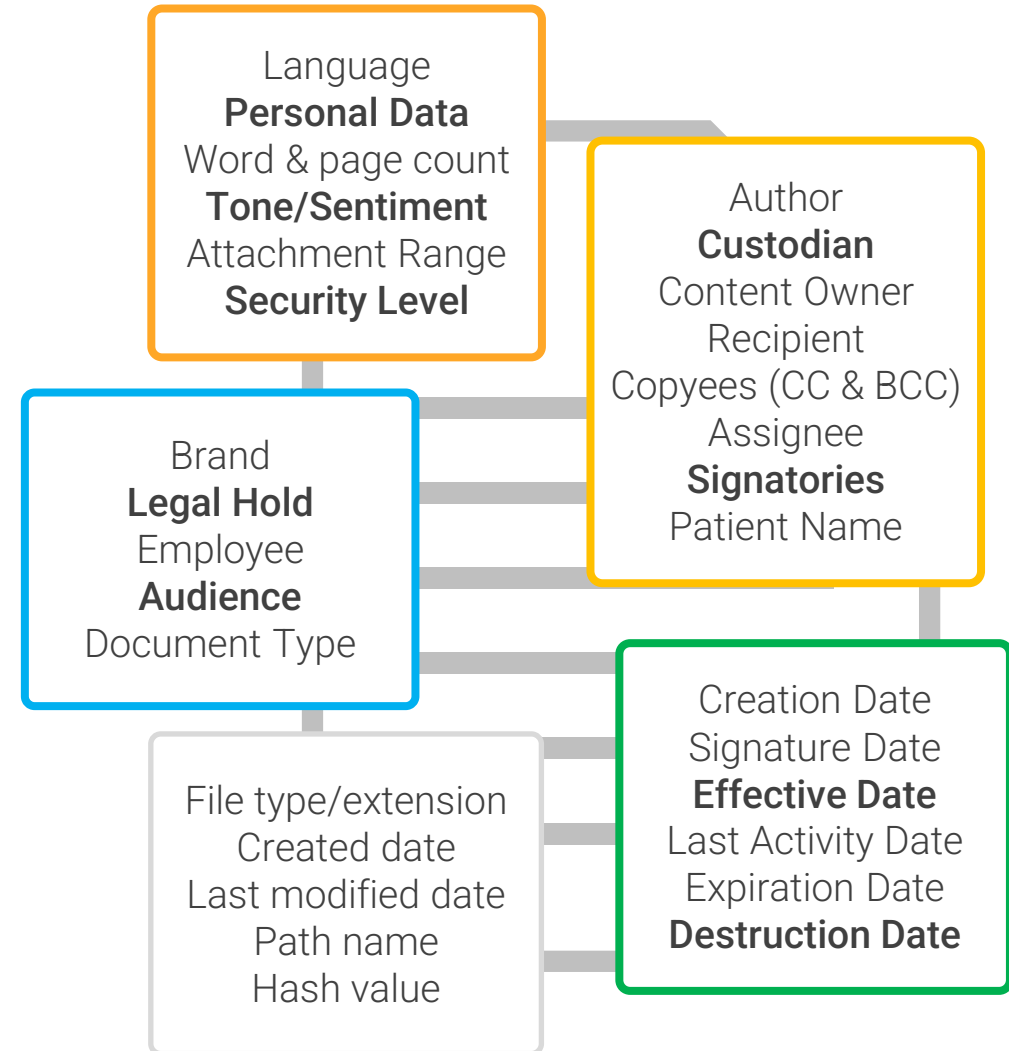
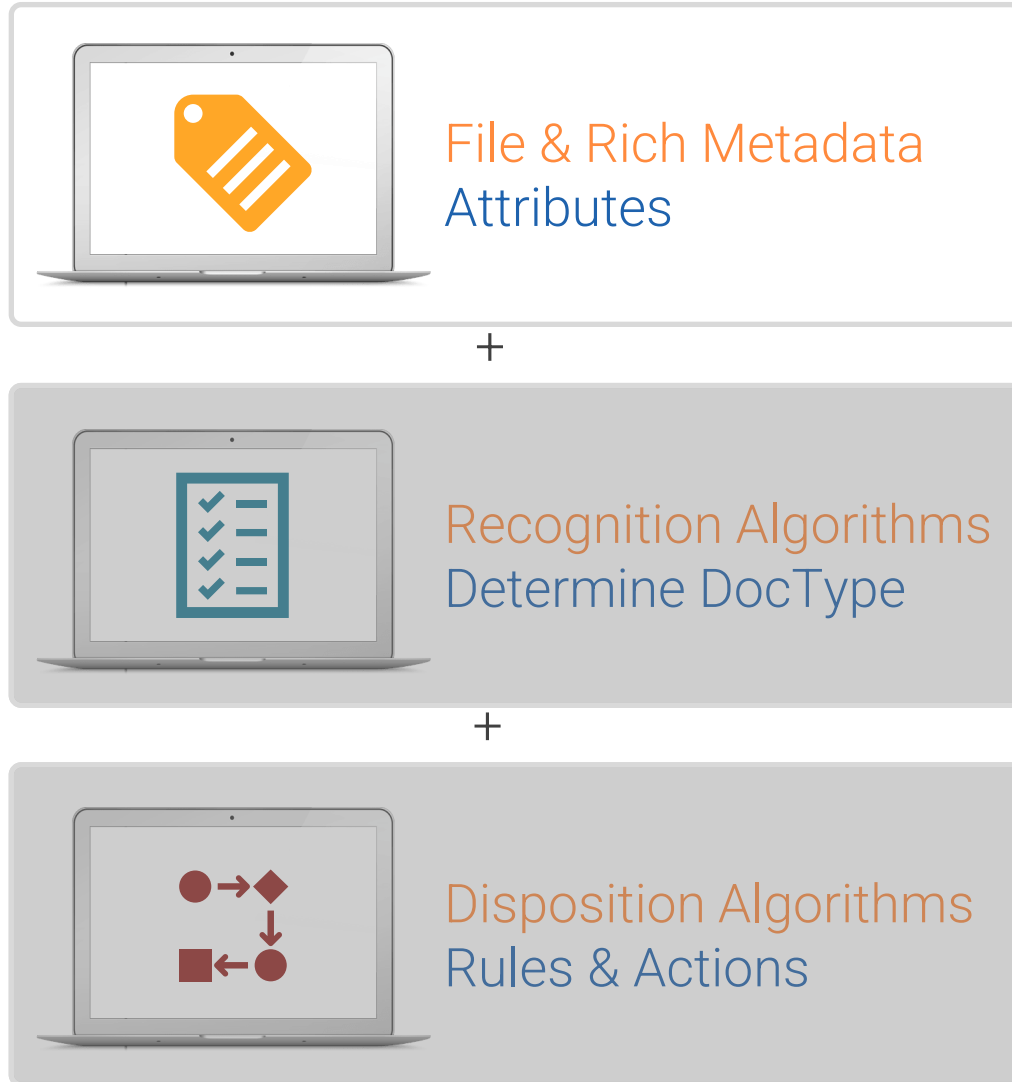
Answers the questions:

What is this thing?

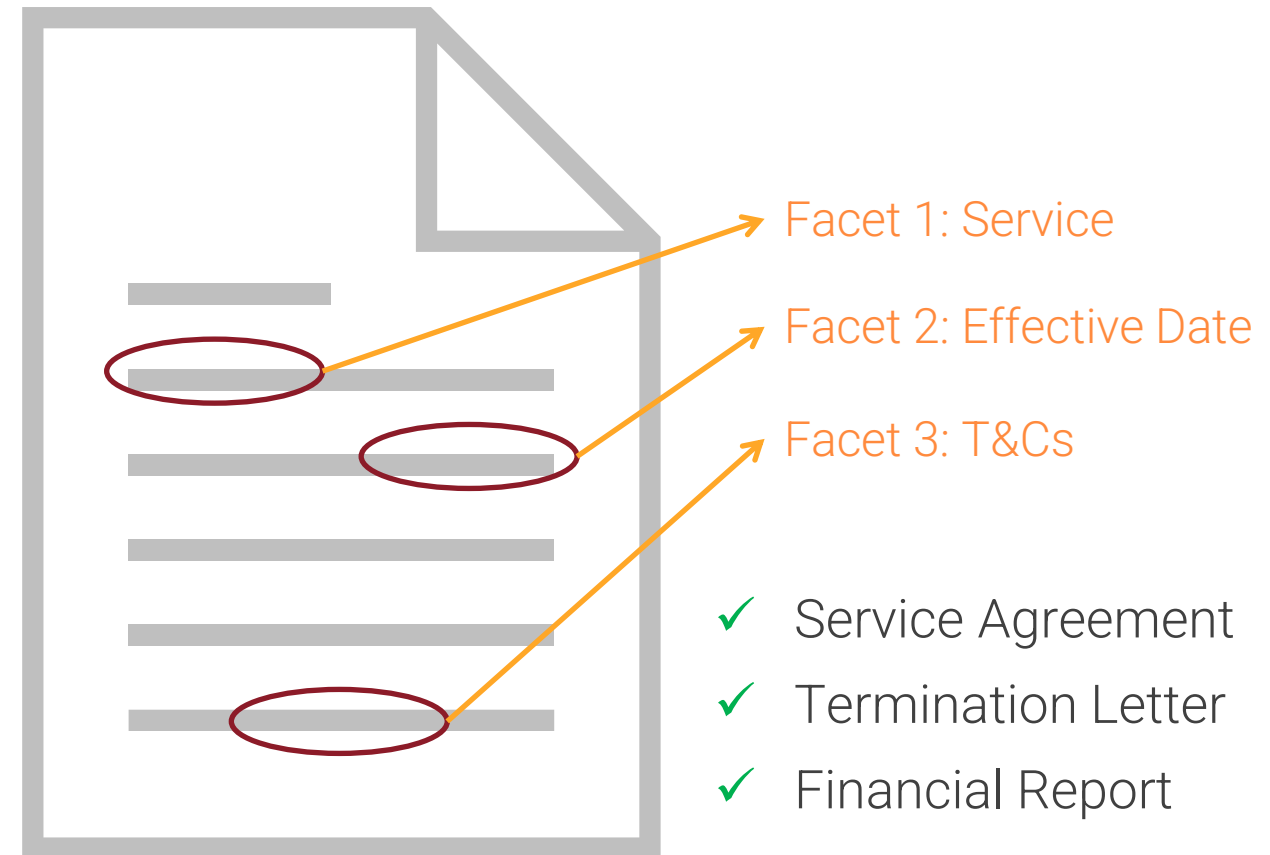
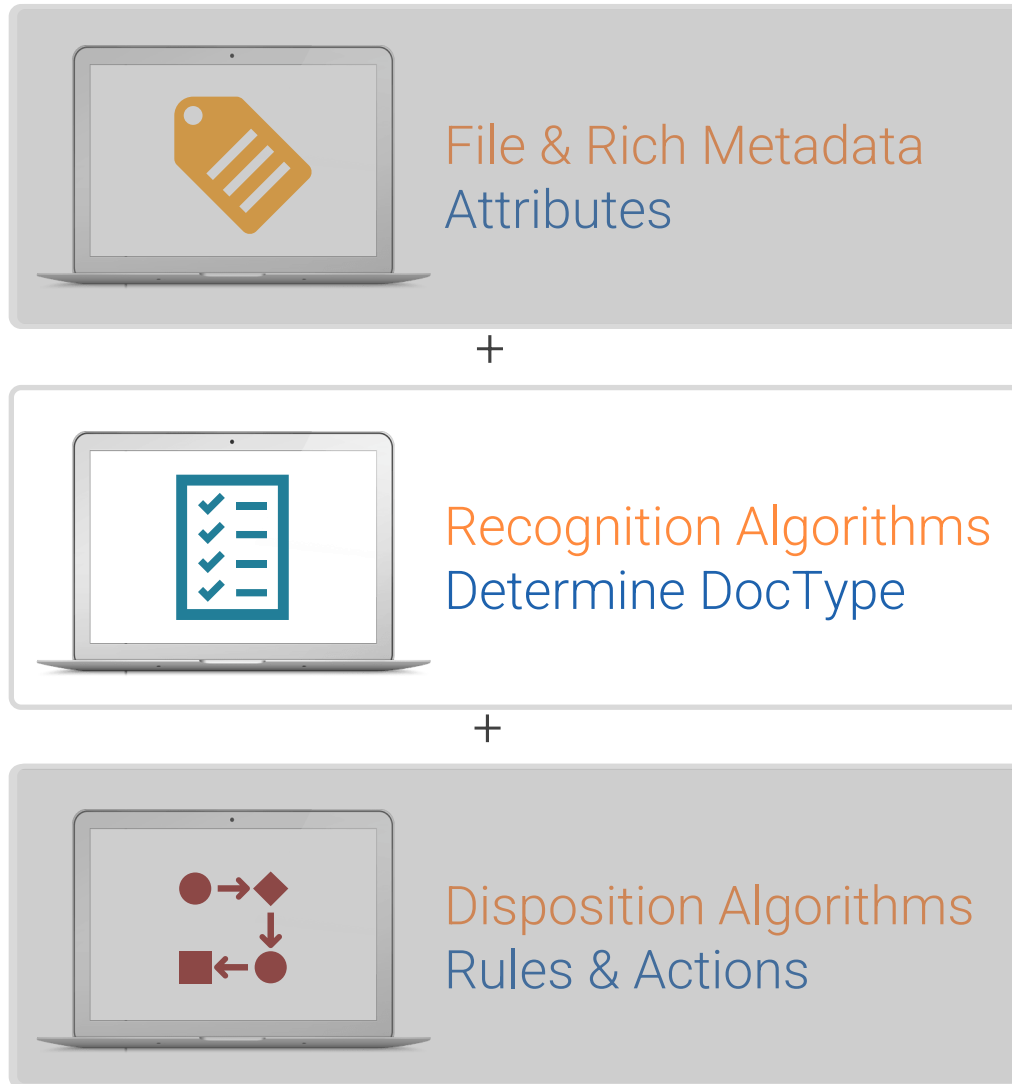
What do I do with it?

On what basis? (Why?)

Sophisticated Classification



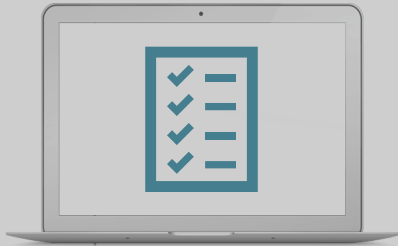
Highly customized



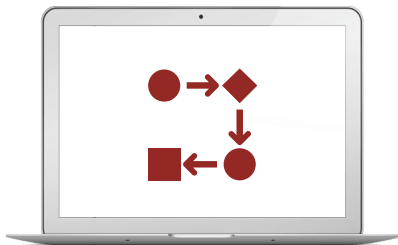
Comprehensive results



File & Rich Metadata
Attributes



Recognition Algorithms
Determine DocType



Disposition Algorithms
Rules & Actions

- Determine the disposition & handling of content
- Typically follow an IF-THEN format, often nested



Disposition Rule Example:

IF the *DocumentType* = Contract AND
the *Date* < 1/1/2025, THEN
Mark File as ACTIVE (Retain CY +7)

Disposition Rules dictate:

- Retention
- Access Security Level (down to content level)
- Legal Hold
- Organization & Location of Content (also Identification)
- ROT: Duplicates, Obsolete, Junk
- PII, PHI & Sensitive Info
- Compliance
- Migration, Promotion of Content
- Litigation Collection & Production
- Optimized Search Results (other apps also)
- Trends & Forecasts
- Routing & Removal of inbound (or outbound) content
- Transformation Needs (ex: language translation, offensive materials removal, routing removal, re-ordering, text creation, etc.)

AutoClassification turns this...

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

Into this



Local Government Property Insurance Fund Actuarial Services Contract

THIS AGREEMENT effective this 1st day of February 2005 (the "Effective Date") by and between the Office of the Commissioner of Insurance (the "Agency" or "State of Wisconsin") and AMI Risk Consultants, Inc. hereinafter referred to as "Contractor"

The purpose of this Agreement for actuarial services is to assist the Agency to accurately assess rate and premium levels and establish loss reserves (incurred but not reported, IBNR) for the Local Government Property Insurance Fund (the "Fund") on an annual basis to ensure the financial stability of the Fund.

I. Term. The term of this Agreement shall be for one (1) year, from the 1st day of February 2005, and expiring on the 31st day of January 2006. The contract may be renewed for two (2) successive one-year periods upon mutual written agreement of the parties. Contractor will notify the Agency six months prior to the annual expiration if it wishes to renew this Agreement and shall specify any amendments the Contractor wishes to propose.

II. Contact Persons. For purposes of administering this Agreement, the following representatives of each party are hereby designated as appropriate contact persons:

(a) For the Agency:

Danford Bubolz, Insurance Program Officer
Local Government Property Insurance Fund
125 South Webster Street, Madison, Wisconsin 53702

(b) For the Contractor:

Aguedo M. Ingco, President
AMI Risk Consultants, Inc.
11410 North Kendall Drive, Suite 208
Miami, Florida 33176-1031

III. Actuarial Services Required

A. Assist the Fund in annually establishing rates and adequate incurred but not reported claims reserves for the Fund.

DocType = Contract

Effective Date = 2/1/2005

Party Two = AMI Risk
Consultants, Inc.

Term = 1 year

Renewals = two 1-year terms

Keywords = Actuarial Services

Elsewhere: exhibits, amendments,
signatories, jurisdiction, cover page,
etc.



RecordType = Contract

RecordClass = AP/AR Contract

Retention = Expiration + 5Y

Handling = Delete after
authorization

Sensitivity = Confidential
(protection/redaction based on user
access class)



Retain or Delete?

Where? Access?
How long?

When? How?
Approval Needed?

Another example of Automated Data Classification (AutoClassification)

DEPARTMENT OF HEALTH AND HUMAN SERVICES Food and Drug Administration		Form Approved OMB No. 0910-0814 Expiration Date: June 30, 2025 See PRA Statement on last page.
Individual Patient Expanded Access Investigational New Drug Application (IND) (Title 21, Code of Federal Regulations (CFR) Part 312)		
1. Patient's Initials J.M.S.		2. Date of Submission (mm/dd/yyyy) 01/05/2024
3. Type of Submission NOTE: Checking box 3a or 3b will "turn on" ONLY the fields that must be completed. 3.a. Initial Submission ☒ Select this box if this form is an initial submission for an individual patient expanded access IND, and complete only fields 4 through 8, and fields 10 and 11.		Investigational Drug Name NeuroCure-X
3.b. Follow-Up Submission ☐ Select this box if this form accompanies a follow-up submission to an existing individual patient expanded access IND, and complete the items to the right in this section, and fields 8 through 11.		Physician's IND Number 3747-7892
4. Clinical Information Indication Treatment of Alzheimer's disease Brief Clinical History (Patient's age, gender, weight, allergies, diagnosis, prior therapy, response to prior therapy, reason for request, including an explanation of why the patient lacks other therapeutic options) Age: 72 Diagnosis (1 year ago): Mr. Smith was diagnosed with early-stage Alzheimer's disease. The diagnosis was based on a combination of clinical symptoms, cognitive assessments, and imaging results. Progression of Symptoms (Over the past year): Memory loss has become more pronounced, impacting daily activities. Challenges in managing finances and keeping track of appointments. Difficulty in recognizing familiar faces and places. Emotional changes, such as mood swings and increased anxiety. Current Status: Mr. Smith's condition has progressed to the moderate stage of Alzheimer's disease. He requires assistance with daily activities and personal care. Communication has become more challenging, and he may repeat questions or stories.		
5. Treatment Information Investigational Drug Name NeuroCure-X Name of the entity that will supply the drug (generally the manufacturer) ABC Pharmaceuticals Inc. FDA Review Division (if known) Treatment Plan (Including the dose, route and schedule of administration, planned duration, and monitoring procedures. Also include modifications to the treatment plan in the event of toxicity.) Mr. Smith is enrolled in a clinical trial for the experimental drug NeuroCure-X. Dose: 100mg twice daily Alongside drug therapy, he receives support from caregivers, participates in cognitive exercises, and is involved in a support group for individuals with Alzheimer's and their families.		
FORM FDA 3926 (7/22) Page 1 of 3 PSC Publishing Services (301) 443-6740 EF		

Form = OMB No. 9010-0814

DocType = Patient Access IND

Drug Name = NeuroCure-X

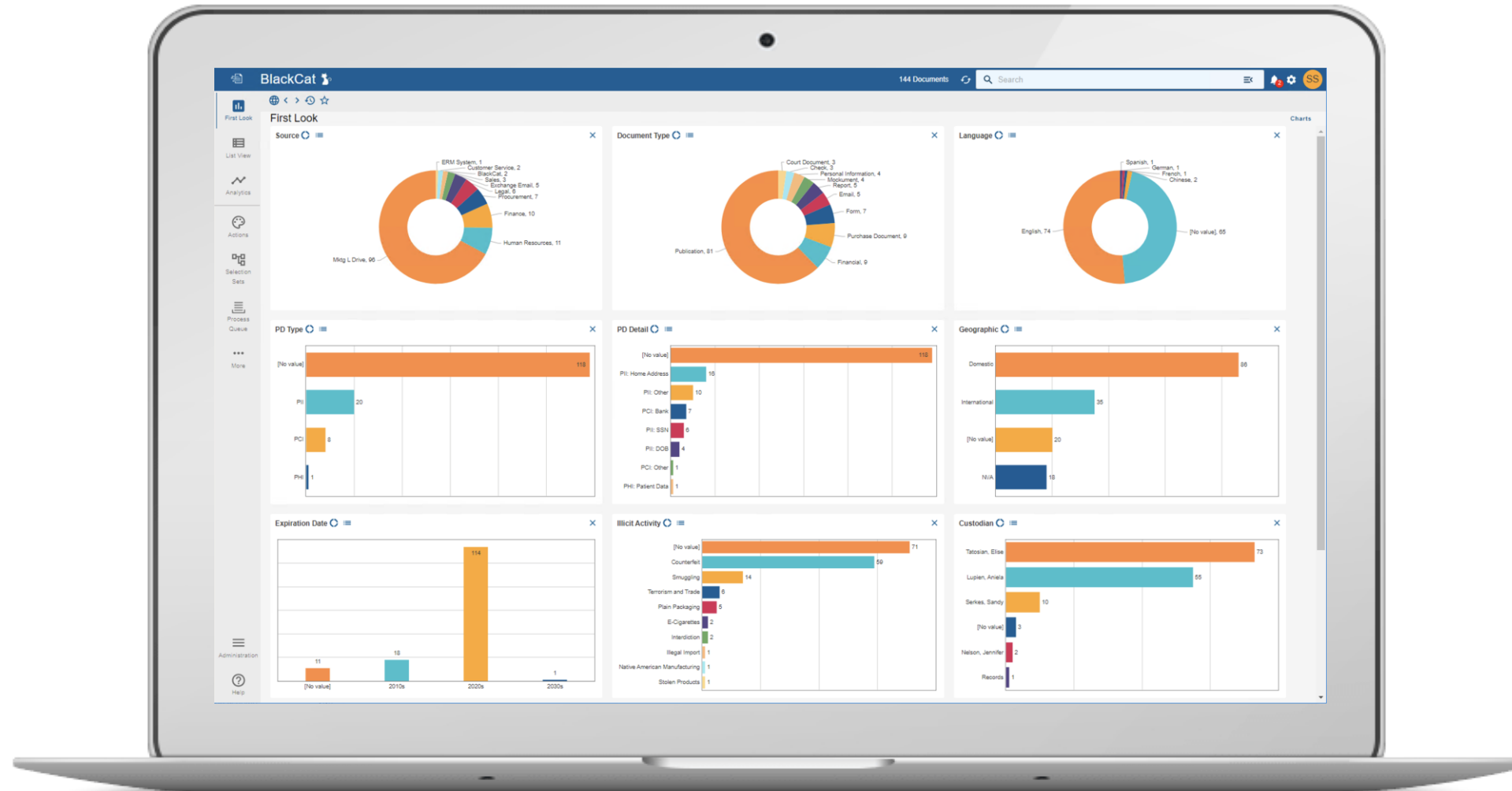
Patient = J.M.S
Cross reference patient database:
Smith, James Morgan

Physician IND # = 3747-7892
Cross reference doctor database:
Dr. Pamela Jones

RecordType = Patient Access IND
RecordClass = FDA Form
Retention = Submission + 5Y
Sensitivity = Confidential
Handling = Delete after authorization

Retain
Delete on Jan 5, 2029

“Single Pane of Glass” Dashboard View



How does AutoClassification help?



IDENTIFIES & TAGS

*Separates the ROT from
content of business value*



NO GUESSWORK

*Remove inconsistencies, drudgery of
manual efforts*



COST SAVINGS

*Storage costs,
human costs*



CREATES INSTRUCTION

*at the content level,
based on metadata & rules*



PRODUCTIVITY

*Maximizes human effort,
increases productivity*



EVERGREEN

*Future-proofs your data. Content is
always current, clean*



DEFENSIBLE

*Actions, decisions, movement
& access is logged*



SET IT & FORGET IT

*STOP: "we really should..."
it will actually get done*



How to AutoClassify R0T

Find, Analyze, Remediate,
Audit

What is ROT?

Content with no business, legal or regulatory value



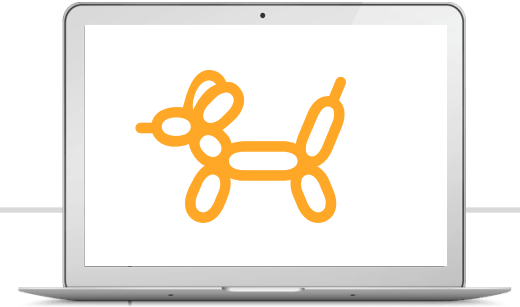
Redundant

- Duplicates
- Backups & cloud storage
- Litigation Databases
- Shared/collaborative sites



Obsolete

- Past Retention
- Last accessed date
- Versions & updates



Trivial

- Personal content
- Cookies & markers
- Spam
- Low value docs

Should it be ROTI?



Inappropriate

- Gotten illegitimately
- Uncertain provenance
- Offensive or illegal



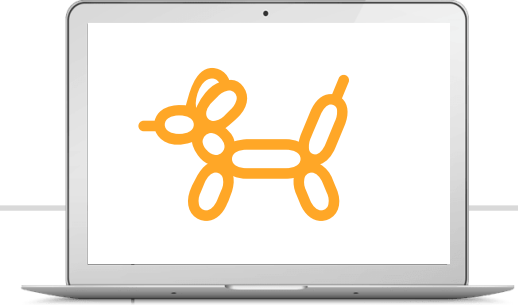
Redundant

- Duplicates
- Backups & cloud storage
- Litigation Databases
- Shared/collaborative sites



Obsolete

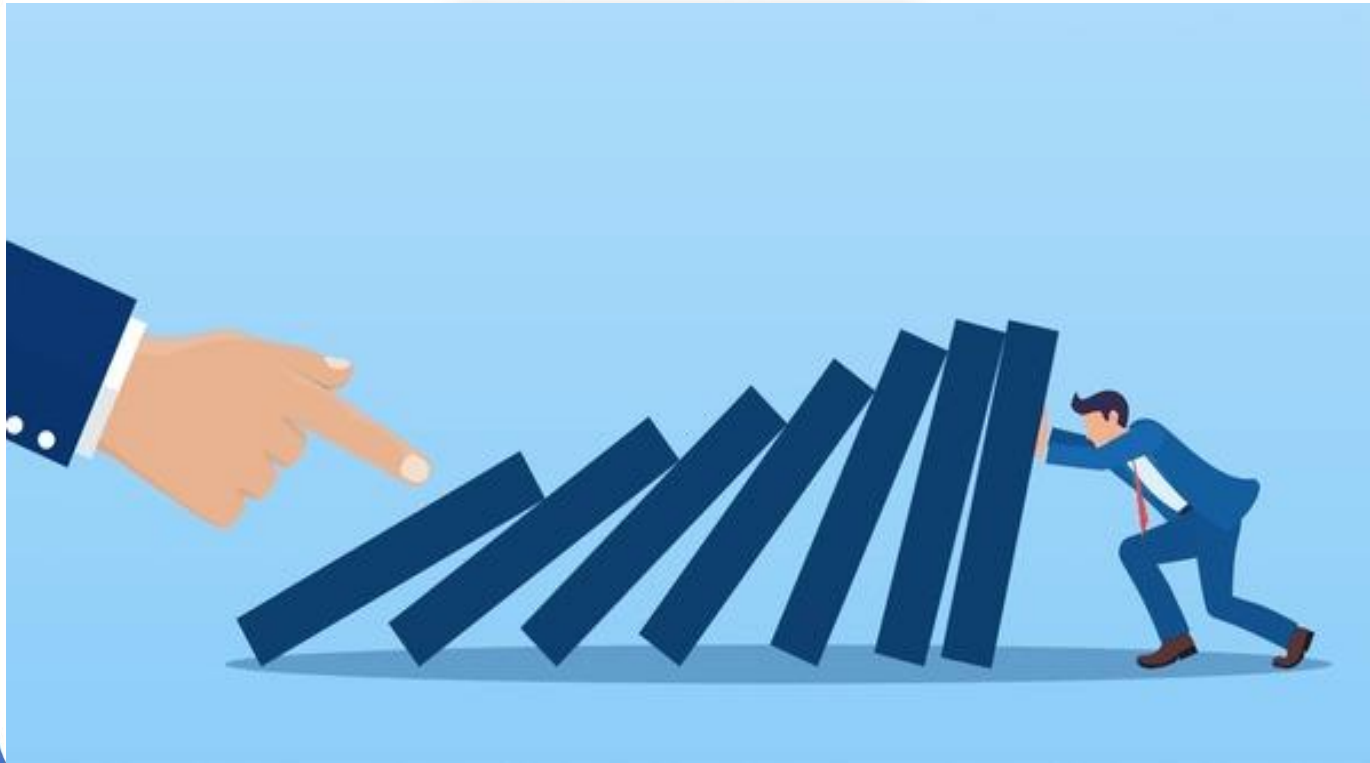
- Past Retention
- Last accessed date
- Versions & updates



Trivial

- Personal content
- Cookies & markers
- Spam
- Low value docs

How much ROT have we got? What's the big deal?



- Typically, 40-50% of enterprise data is ROT
- ROT impacts:
 - Data storage costs
 - Data breach exposure
 - Litigation & eDiscovery fees
 - Compliance & Data Privacy
 - Organizational efficiency
 - AI Readiness

What's the risk?

Estimated that 50% of all stored enterprise content is ROT



Cost

- Storage & maintenance costs
- A large company with 10PB of data could be spending as much as \$34.5m on data that could be deleted

Compliance, Governance & Oversight Council (CGOC report)



Lost Productivity

- Employees can't find things
- Working with latest version?
- Workers spend 30% (2.5 hours) of their workday looking for content

(IDC Study)



Risk & Exposure

- Data Governance Risk: don't know what it is / where it is
- Privacy Risk: does it contain PII or sensitive data?
- Compliance risk: does this exist when it shouldn't?

What's the risk & impact?

Estimated that 50% of all stored enterprise content is ROT



Cost

- Storage & maintenance costs
 - A large company with 10PB of data could be spending as much as \$34.5m on data that could be deleted
- Compliance, Governance & Oversight Council (CGOC report)



Lost Productivity

- Employees can't find things
 - Working with latest version?
 - Workers spend 30% (2.5 hours) of their workday looking for content
- (IDC Study)



Risk & Exposure

- Data Governance Risk: don't know what it is / where it is
- Privacy Risk: does it contain PII or sensitive data? Big fines.
- Compliance risk: does this exist when it shouldn't?

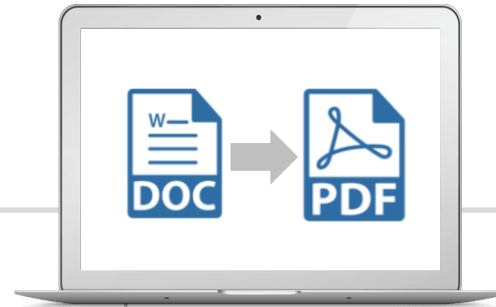
Understanding Duplicates

Goal: remove as much duplication as possible



Exact Duplicates

- Forensically identical
- Hash Values match



Functional Duplicates

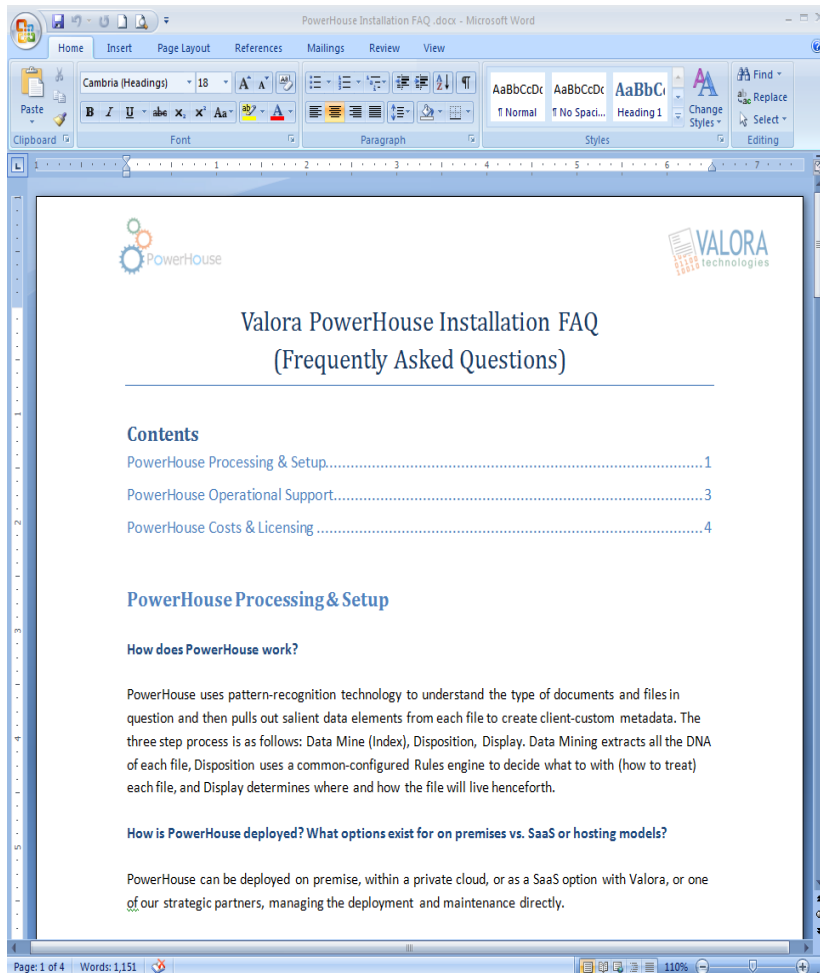
- 99% duplicates
- Functionally identical, but forensically different (no Hash match)



Near duplicates

- (Similar) files are neither functionally or forensically identical
- May be similar or related enough to warrant their treatment as a family unit. Ex: Revisions

Understanding Functionally Identical Duplicates

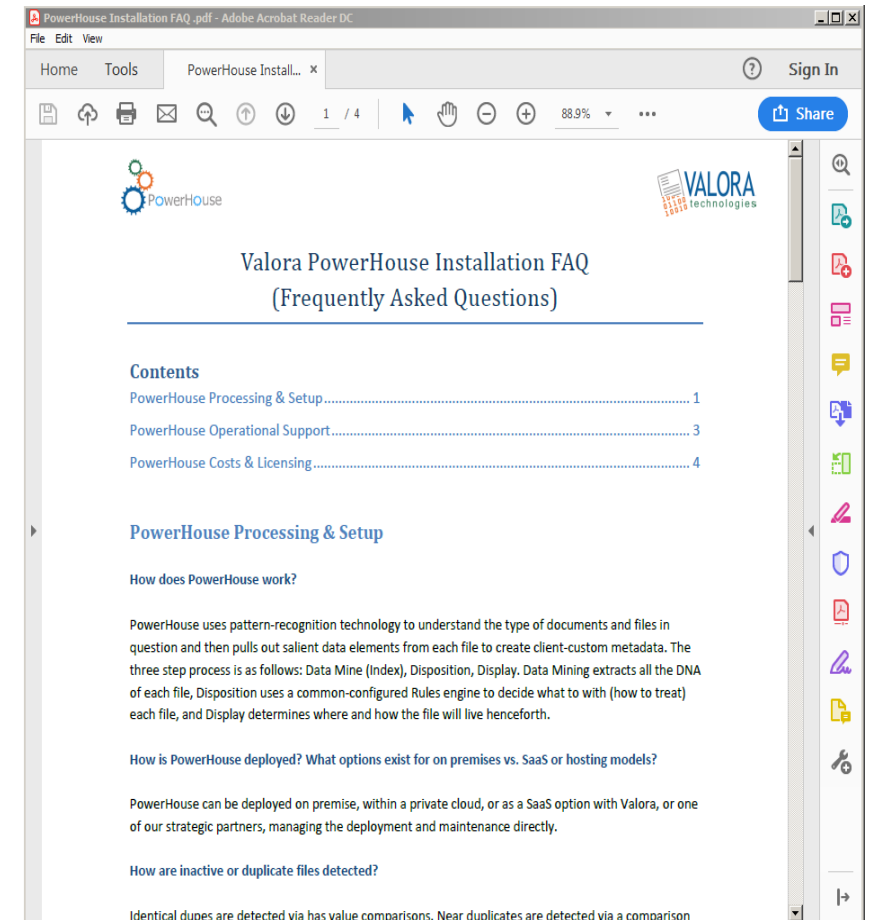


Questions v12.doc
96 KB



FAQ Final.pdf
616 KB

Functionally identical
based on content



Auto Classifying ROT (junk)



“Watch list” terms

Date

+ No further content

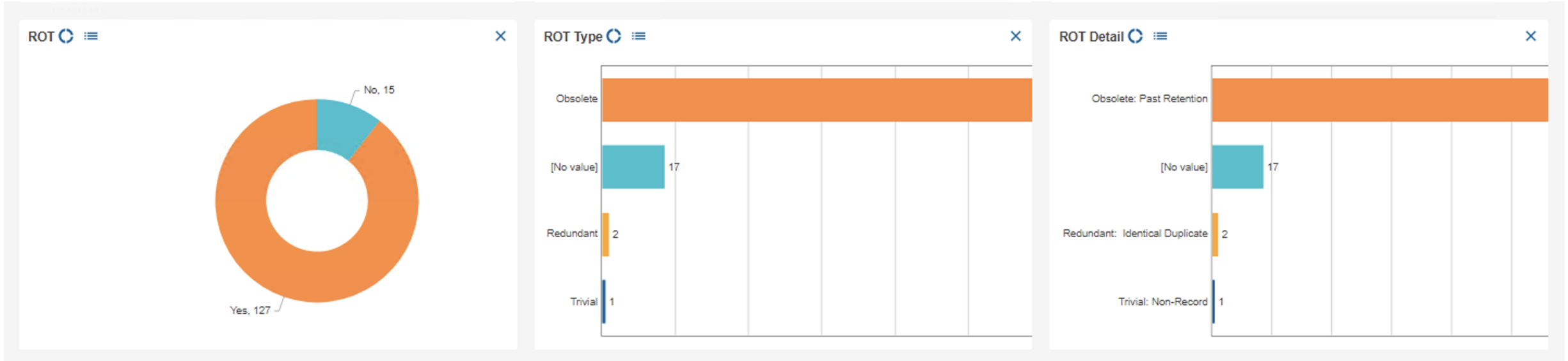
+ Heavy graphics

+ 4,000 ID copies



RecordType = ROT
RecordClass = [None]
Retention = [None]
Handling = Delete
Sensitivity = [None]

AutoClassified ROT & Duplicates



It this file ROT?

- Yes
- No

If so, what type of ROT?

- Redundant
- Trivial
- Obsolete

What is the **basis** for the ROT tag?

- Redundant
because it is an identical duplicate
- Obsolete
because it is past retention
- Trivial
because it is a non-record

Identifying ROT: Redundant

First Look

List View

Analytics

BlackCat

9 Documents (of 199)

Functional duplicates: Demo...

List View

	Special	Source	Title	ROT	ROT Type	Duplicate Type	ROT Detail	Retention Period	Expiration Handling
☐	  	Sales	NeuroCure-X Clinical Trial Form ...	Yes	Redundant	Identical Duplicate	Redundant: Identical Duplicate	Last Modified + 5 Years	Automatic Deletion
☐	  	Sales	NeuroCure-X Clinical Trial Form ...	Yes	Redundant	Identical Duplicate	Redundant: Identical Duplicate	Last Modified + 5 Years	Automatic Deletion
☐	  	Sales	NeuroCure-X Clinical Trial Form ...	Yes	Redundant	Identical Duplicate	Redundant: Identical Duplicate	Last Modified + 5 Years	Automatic Deletion
☐	  	Sales	NeuroCure-X Clinical Trial Form	Yes	Redundant	Identical Duplicate	Redundant: Identical Duplicate	Last Modified + 5 Years	Automatic Deletion
☐	  	Sales	NeuroCure-X Clinical Trial Form ...	Yes	Redundant	Identical Duplicate	Redundant: Identical Duplicate	Last Modified + 5 Years	Automatic Deletion
☐	  	Sales	NeuroCure-X Clinical Trial Form ...	Yes	Redundant	Identical Duplicate	Redundant: Identical Duplicate	Last Modified + 5 Years	Automatic Deletion
☐	  	Sales	NeuroCure-X Clinical Trial Form	Yes	Redundant	Identical Duplicate	Redundant: Identical Duplicate	Last Modified + 5 Years	Automatic Deletion
☐	  	Sales	NeuroCure-X Clinical Trial Form ...	Yes	Redundant	Identical Duplicate	Redundant: Identical Duplicate	Last Modified + 5 Years	Automatic Deletion
☐	  	Sales	NeuroCure-X Clinical Trial Form	No		Identical Duplicate		Last Modified + 5 Years	Automatic Deletion

Team captain
identical dupes

Team captain
functional dupes

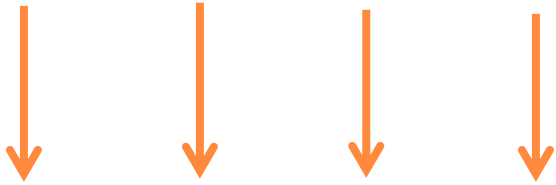


AutoClassifying Records the Right Way

Mapping to the Records Retention Schedule

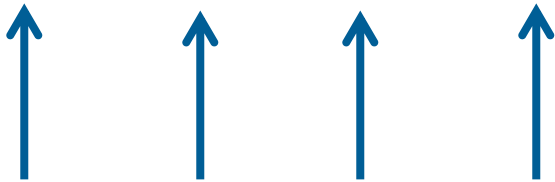
Top-Down vs. Bottom-Up Approach

Determine Broad Record Types



- Implement downward onto departments
- Individual employees “make it fit” into provided classifications
- RIM personnel are encyclopedic experts on Record Types
- Onus is on: department/employee
- Non-records managed by time

Understand Specific Content



- Implement upward from content analysis
- SMEs design rules & classifications around known content types
- Onus is on: centralized IG team
- Non-records managed by value

Auto Classifying an Employee Termination Letter

Employee Termination Letter for Attendance

January 8, 2016

Mr. Eric Jacobson
Sales Representative
M & Co
20 E Utica Street
Oswego, NY 12236

Dear Mr. Jacobson:

This letter is being written to inform you of your termination as Sales Representative at M & Co owing to your poor attendance. We have sent you three written warning (copies enclosed) detailing how your behavior may eventually affect your work but you have paid no heed to any of them. Mr. Leslie Neal (Sales Supervisor) has also spoken to you personally about your tardiness but we have not seen any change in your routine.

Please understand that M & Co is a team environment and any team member who is not delivering his work brings down the efficiency of the entire company. Since you make it to work barely three times a week, teamwork is not possible. We observed that the sales campaign that you were assigned to has not produced any results – the entire team was effected due to its failure.

It is my understanding that Mr. Neal offered you an unpaid leave so that you could sort out any issues that you may be facing which have resulted in poor attendance on your part. Your refusal to accept this offer confirms that you are non-serious to your work.

Please note that this termination is effective immediately. Mr. Neal will be with you after lunch so that you can hand over your tasks and any company information that you may have in your possession.

Regards,

Primary DocType = Correspondence
Secondary DocType = Termination Letter

Employee Name = Jacobson, Eric
Employee ID = 44701 (via HRM/DB lookup)

Effective Date = 2016-01-08

Keywords = “poor attendance,” “three written warnings,” “tardiness”



RecordType = Termination
RecordClass = Personnel Record
Retention = Termination Date + 7Y
Handling = Delete after authorization
Sensitivity = Business Confidential
Personal Data = Yes
PD Detail = Home Address

ROT, Records & Retention Classification Tags



ROT YES/NO

- Simple umbrella classification Yes/No
- Summary of ROT Detail



DUPLICATE TYPE

- Identical (Hash) Duplicate
- 99% (Functional) Duplicate
- 75-98% (Near) Duplicate
- Unique



RECORD CLASS

- Based on Document Type, Source, Custodian/Author, Jurisdiction, Keywords, Personal Data/Sensitivity



ROT DETAIL

- Redundant: Identical Dupe
- Obsolete: File Extension, Past Retention, Temporary, Draft
- Trivial: Auto-Reply, Low Business Value, Non-Record, Code or Executable, Thumbnail, Logo, File Path, File Type, Logo



EXPIRATION DATE

- Calculated from Date Hierarchy (differs by Records Type)
- Creation Date, Hire/Start Date, Termination Date,
- Often triggered by events



RETENTION PERIOD

- Direct inherit from Record Class (Rules)

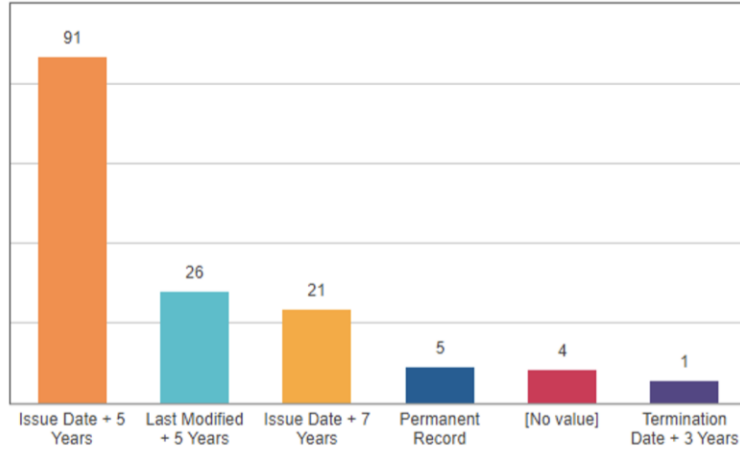


EXPIRATION HANDLING

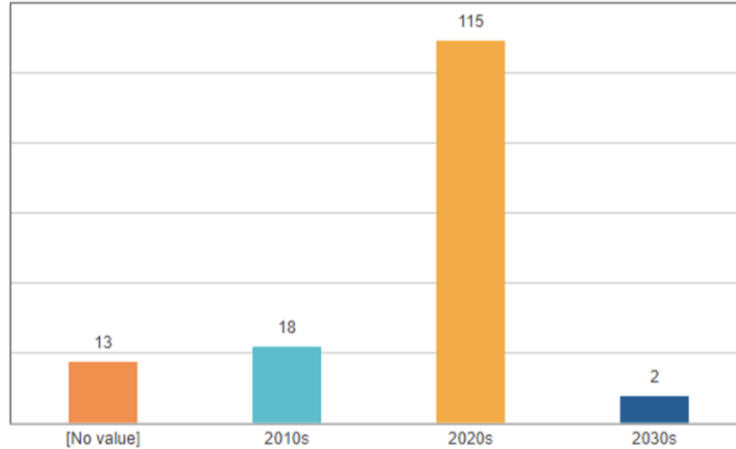
- Automatic Deletion
- Approval Required for Deletion
- Quarantine
- Permanent Record

Automated Retention Handling

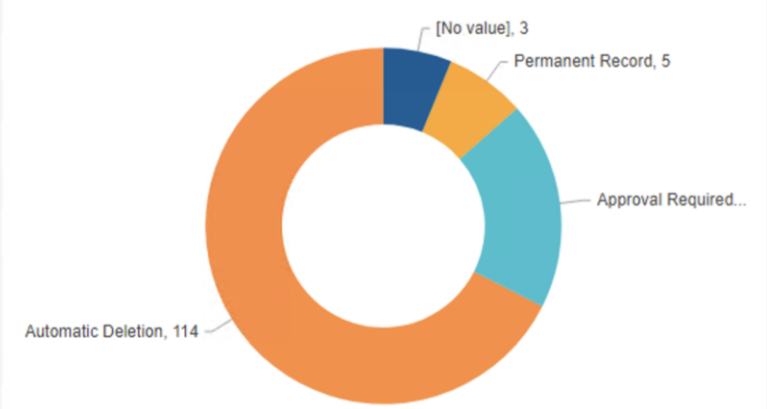
Retention Period



Expiration Date



Expiration Handling



List View

Original File Name	Document Type	Record Class	Expiration Date	Expiration Handling	ROT	ROT Type	ROT Detail
HR_Job Application 1.jpg	Form	Miscellaneous	10/7/2021	Automatic Deletion	Yes	Obsolete	Obsolete: Past Retention
NeuroCure Clinical Trial Form....	Form	Miscellaneous		Automatic Deletion	No		
Meeting Materials1.jpg	Meeting Materi...	Legal	11/1/2021	Permanent Record	Yes	Obsolete	Obsolete: Past Retention
HR_Resume 1.png	Personal Infor...	Miscellaneous	10/10/2021	Automatic Deletion	Yes	Obsolete	Obsolete: Past Retention
HR_Resume 2.png	Personal Infor...	Miscellaneous	10/7/2021	Automatic Deletion	Yes	Obsolete	Obsolete: Past Retention
HR_Resume 3.png	Personal Infor...	Miscellaneous	10/10/2021	Automatic Deletion	Yes	Obsolete	Obsolete: Past Retention
Doctors warned about counterf...	Publication	Public Relations		Automatic Deletion	No		
ICE targets sale of NBA count...	Publication	Public Relations		Automatic Deletion	No		
The Scary Truth About the Bea...	Publication	Public Relations		Automatic Deletion	No		
Closing Statement.png	Purchase Doc...	Accounting & Finance		Approval Required for Deletion	No		
Closing Disclosure2.png	Purchase Doc...	Accounting & Finance		Approval Required for Deletion	No		
HR_Incident Report Form 1.jpg	Report	Miscellaneous	10/10/2021	Automatic Deletion	Yes	Obsolete	Obsolete: Past Retention
Banking Statements 2.jpg	Purchase Doc...	Accounting & Finance	5/31/2010	Approval Required for Deletion	Yes	Obsolete	Obsolete: Past Retention
HR_Job Application 2.gif	Form	Miscellaneous	10/7/2021	Automatic Deletion	Yes	Obsolete	Obsolete: Past Retention

3 Approaches to Disposition or “Actioning”



Solution's
Software Agents

Action taken by:
Automated workflow



Client's
IT Team

Action taken by:
IT Team



Client Employees
Self-Service

Action taken by:
Employees

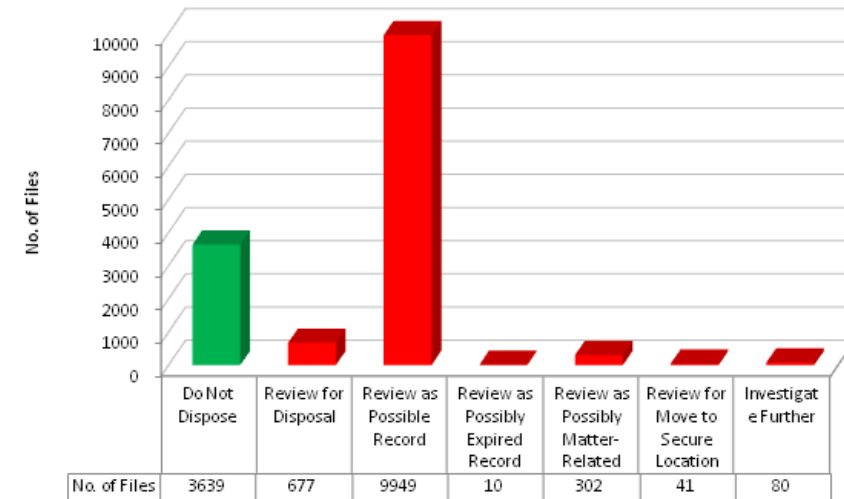
Guiding Employees to Self-serve

STARK INDUSTRIES NON-RECORD DISPOSAL REPORT

Custodian: Potts, Pepper
Report Run Date: 17 Mar 2020

Total File Counts:
14350

Breakdown of Recommended Actions



Files Found		What to Do For Each File		Where the File is Located		Click to Perform
File Name	DocTitle	Status: Recommended Action	Details	File Path (click to view file)	Recommended Action	
Pepper @ Stark.url	Pepper @ Stark	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
Copy of 2019 Firm Invite List_Aug 6.xlsx	Total Potential Firms 205 Notes	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
Discovery map for banner.pptx	Discovery map for banner	Investigate Further		C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
Draft Learnings Presentation	Litigation Learnings Masterclass 1	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
MJH Portal II.docx	LITIGATION INFORMATION GROUP (LIG) PORTAL	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
Lit Realities 102 Abbreviated.pptx	Litigation Realities102:	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
Litigation Realities and Learnings Survey XLT	Litigation Risks;Preservation of documents in	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
MTHM All Learnings Reports final.xlsx	Pivot	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
Original First Preservation Notice	Original/First Preservation Notification	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
PN Project Plan.xlsx	Matrix	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
Preservation Notice Review Team Charter.pptx	Preservation Notices/Legal Hold Review Team Charter	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
Preservation Notice Review Team Project.pptx	Preservation Notice/Legal Hold Review Project Timeline Due: 9/27/19	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+
project glic.xlsx	Amount Claimed Against Stark \$332,045,063 \$166,602,984 Amount Claimed By Stark \$101,234,467	Protected: Do Not Dispose	Last Modified < 3 Years	C:/Stark-NRD-ODB/Pepper Potts OneDrive - Stark.zip/OneDrive - Stark		+



Putting it all Together

Retention & lifecycle management
with AutoClassification

How do you do it?



1



Data Discovery
Locate and inventory

2



Intelligent Analysis
Full-text content analysis

3



AutoClassification
Content-based Classification

4



Defensible Disposition
Automated rules & workflows

5



Delta Monitoring
Ongoing scans & audits

AutoClassification: Technology + Methodology



CRAWL

Locate, ID content

- On-prem, cloud-based repositories
- Unstructured & structured repositories
- Info on who, what, where, why, how



TAG

Apply rich metadata

- Tag rich metadata, content/context
- Determine DocType, dupes/near dupes, ROT
- Tag for keywords, custom metadata



ACTION

Automate rules

- Simple/complex rules
- Apply retention schedules & legal hold
- AutoRedact PII, apply security access control



EMPLOY

Search & retrieve

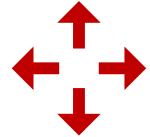
- Records Retention
- Data/DSAR requests
- Legal hold
- Compliance
- Litigation
- M&A



MONITOR

For new data

- Automate processes for new & edited content (perpetual)
- Update/dynamic retention schedules
- Runs in background, no performance draw

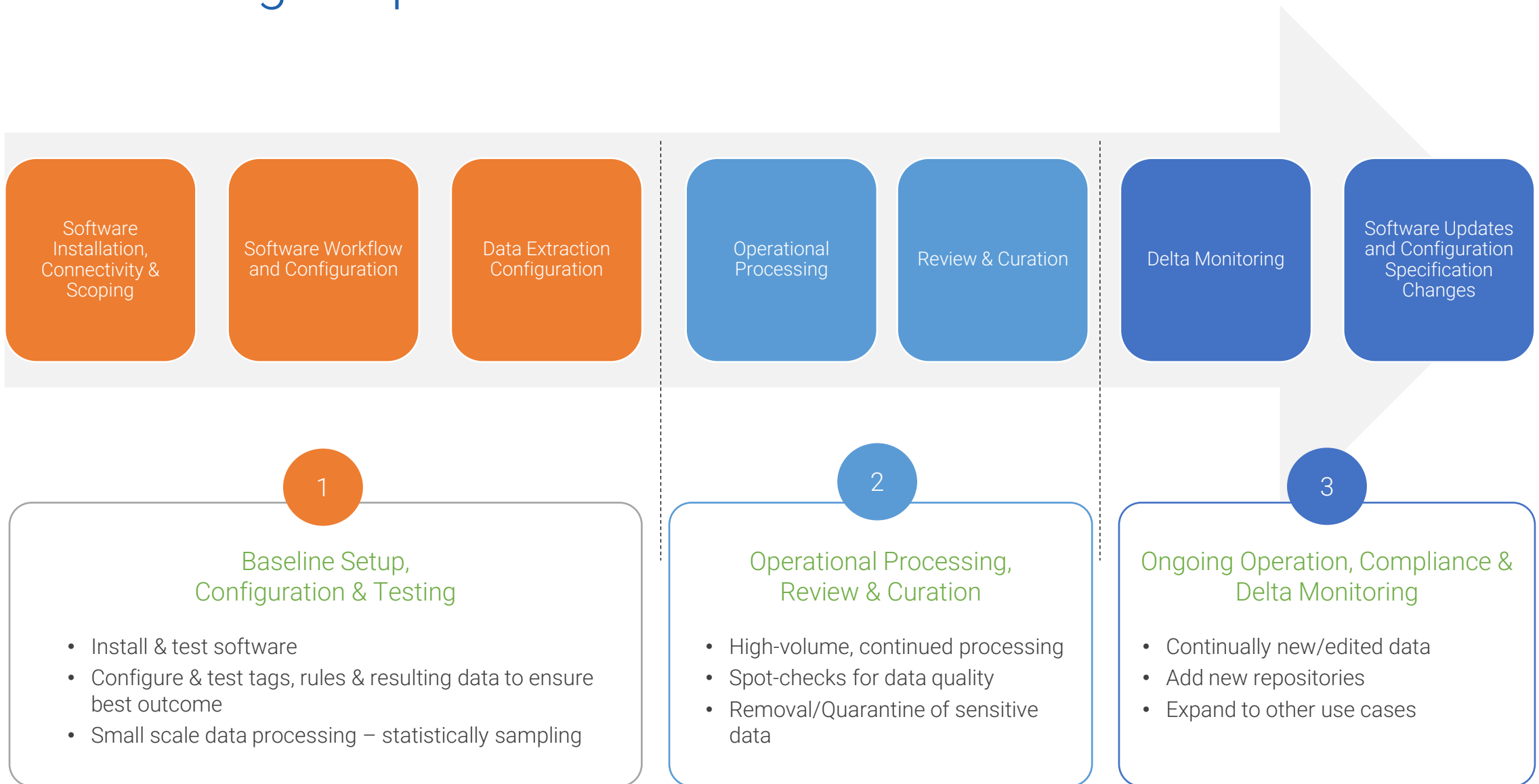


EXPAND

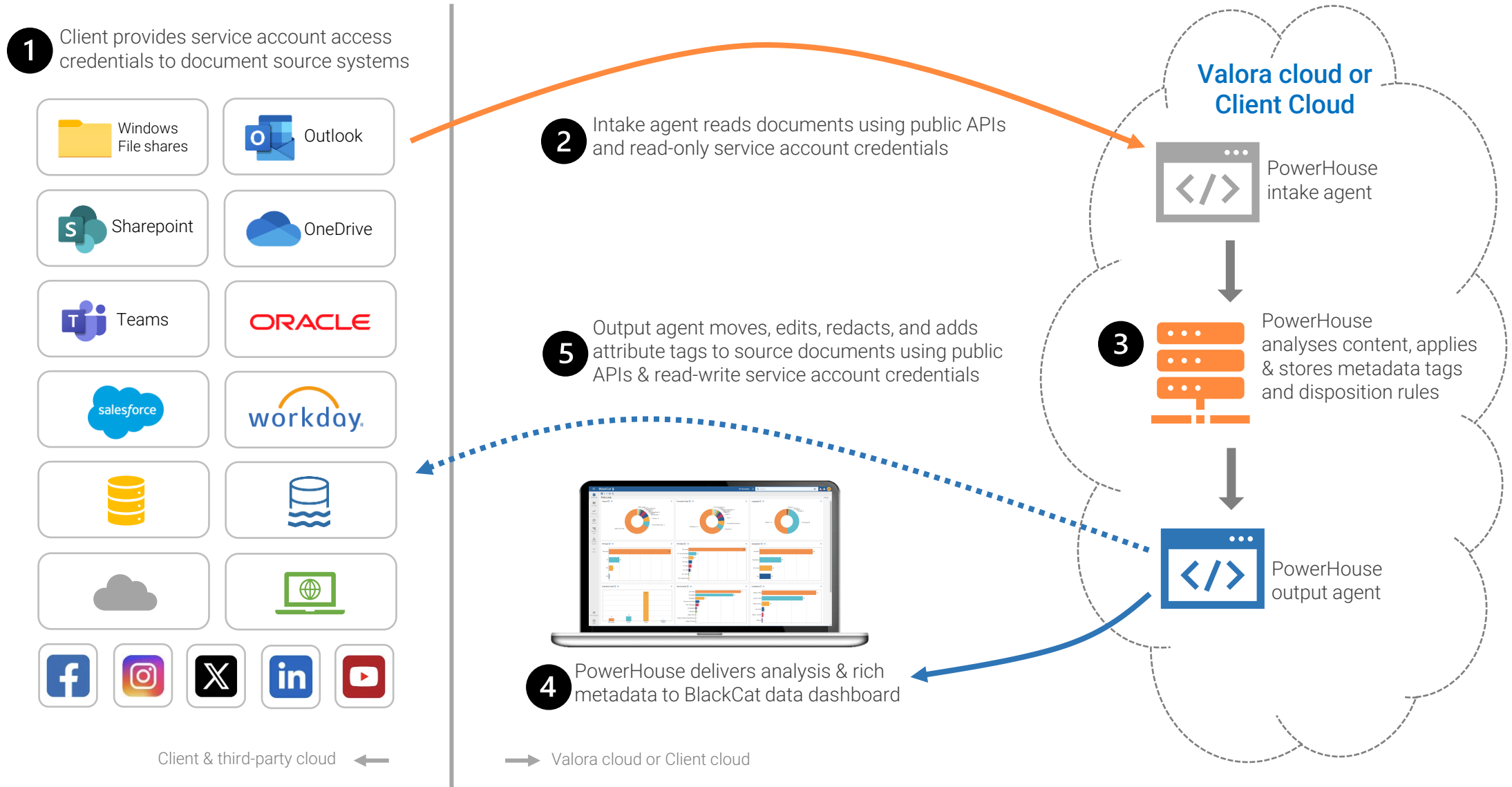
To other areas

- + Data repositories
- + Use cases
- + Departments
- + On-demand search & retrieval
- + Self-serve

Onboarding & Operational Phases



Connecting to Data Stores



Roadblocks & Gotchas



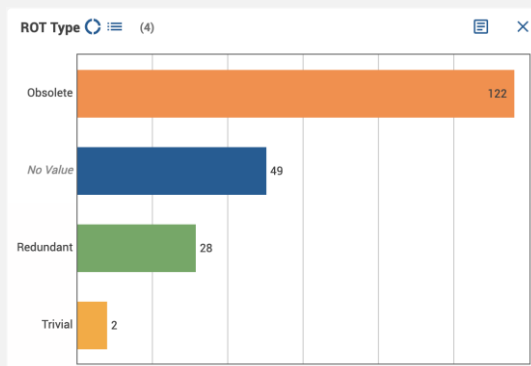
IT: “storage is cheap”,
ROT is not a problem



People will resist
cleaning up their data



No single group “pays for” ROT
(IT? Legal? Compliance?)



Show people the risk



Get them involved



Share the love.
Share the budget.

How to Get Started (how to get internal buy-in)

Pilot Use Case

Start with a compelling use case

- Quick and easy that will prove ROI quickly
- 1 content owner/ROT?
- Prove ROI



Look Forward

Always have an eye on future use cases

- What do you need for this application or use case?
- What will you need for future use cases?



Measure Results

Benchmark results and accuracy of pilot

- Accuracy of results?
- Increase guidance or seek SME?
- What reporting was useful? What else is needed?



Build on Success

Build business case on pilot alone then expand

- Prove ROI functioning system
- Where to next?
- Additional use cases running on same platform



Spread the Word

Sharing is caring, expand internally

- Host internal showcases for mgmt & other departments
- Design pilots that highlight important content, results or exposure points



Responses to the “hell no.”



- **Leave it:**
Some organizations leave the dupes where they are as long as there is not no business risk to doing so.*
- **Too Bad, So Sad:**
“It’s policy: you have non-business content stored on corporate data stores & your dupes are taking up half a terabyte of storage”. “It is a breach risk.”
Tip: delete the dupe and tombstone to where the master lives.
- **The Gradual Goodbye:**
Keep the dupes where they are for 3 months, then move them to a “To Be Deleted” folder (with a tombstone). Keep them in the “To Be Deleted” folder for 6-12 months. If no one misses them in that time, then delete them.
- **Obsolescence Risk*:**
If the content exists, it is discoverable in a potential litigation. Regulated industries don’t mind a dupe or two hanging around, it’s the obsolescent content that is the real risk of it existing when it should not but way of a) the law, b) industry regulations, and/or c) their own retention policies.



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