Why Automated Information Management is the answer for GDPR, CCPA, and whatever comes next?
Let’s start with where most of us are today...
GOVERNING INFORMATION IS HARD

INFORMATION MANAGEMENT IS DIFFICULT

PEOPLE DON’T LIKE DOING IT

IT DISRUPTS THEIR DAILY ACTIVITIES

NEGLECTING IT CREATES RISK

DIFFERENCES IN PROCESSES CREATE UNCERTAINTY
COMMON CHALLENGES OF INFORMATION GOVERNANCE

“Where is all of our information located?”

“Who has visibility to see my sensitive content?”

“Why won’t users upload content accurately?”

“Is information accessible enough to be useful?”

“How much unstructured content do we have?”
WHAT’S THE CAUSE?

RETENTION is managed disparately, or not at all

PHYSICAL RECORDS exist outside this infrastructure

TRANSACTIONAL SYSTEMS are hard to use

COMPLIANCE teams don’t have a simple view of records

INFORMATION is scattered and difficult to manage
DO YOU HAVE TIME FOR POOR INFORMATION GOVERNANCE?

83% of knowledge workers lose time to versioning issues every day

50% Professionals spend 50% of their time searching for information

21% Document issues account for 21.3% of productivity loss
Unstructured Data Problem

The world will create and replicate 163ZB of data by 2025\(^1\)

Putting that data on Blue-Rays would be a stack 1.5x earths diameter\(^1\)
(That’s 6.5 trillion Blue-Ray discs)

90% of this new data will likely be unstructured\(^1\)

53% of new unstructured data will likely be ROT\(^2\)

This will be a 10x increase from 2016\(^1\)

\(^1\)IDC - Data Age 2025: The evolution of data to life critical
\(^2\)Varonis – Global Data Risk Report
If you still have a DVD player...

Just the ROT alone is 500+ times the distance
File and Folder Problem

- ROT is found in 72% of folders
- 53% of companies found over 1000 sensitive files open to all employees
- 15% of companies found more than 1 million folders open to all employees

Varonis – Global Data Risk Report
Automation is a journey that requires complimentary efforts with people and process in order to succeed.

Automation is the only way to ensure near 100% data quality, which ultimately supports all information requests including privacy concerns.

Automation is the path forward to simplify and improve the creation/capture of new information.
What kind of Automation and Technology are we talking about?

- AI (Artificial Intelligence)
- RPA (Robotic Process Automation)
- Semantic Search
Which AI are we talking about anyway?

**Artificial Intelligence**
- Using large, high quality data sets to perform high order data management activities

**Automated Intelligence**
- Using automated tools such as RPA to automate repeatable and somewhat mundane tasks currently performed by people.

**Augmented Intelligence**
- Helping accelerate and improve the quality of business process activities by augmenting the work performed by people using intelligent agents (see Automated Intelligence)
NLP (Natural Language Processing)

Structure of text
I am going to the Kick Off today

Meaning of text
The Kick Off Meeting

Intent of text
Siri, take me to Gimmal
Machine Learning

Trains machines how to learn

Load documents

Prediction

Tuning

Evaluation

Read documents

Train algorithm
Deep Learning

Subset of machine learning focusing on patterns
[RPA Defined] Software that can automate repetitive, rules-based tasks that are performed by people on a computer using the same systems.

- Mimics human actions
- Can operate application systems
- Operates on structured data
- Continuous processing
- Fast to implement and scalable
Who uses RPA?

Organizations with large teams working on highly repetitive tasks

- Accounting Teams
- HR Teams
- Trading Teams
- Other Business Areas

How about document centric processes?

- Auto Classification
- Automating Lifecycles
- Ensuring Adherence to Policy

Accounting Teams
HR Teams
Trading Teams
Other Business Areas
Using automation to apply layers of intelligence

Content enrichment by repetition via automated activities
Auto Classification Example

Adding document into SharePoint document library

An RPA solution was devised that:

- End user “adds” document to SharePoint
- Desktop Bot wakes up on “add” action
- Bot launches business application with metadata
- Bot scrapes metadata from business app
- Adds metadata to proper columns
- Finalizes document add

Results

- Reduced processing time
- Consistent use of metadata
- Eliminated manual steps
Semantics and Graphs
How does it apply to automation?
Similar Technologies

Federated Search
- Does not use semantics (so not really similar)
- Compiles text and metadata into a single search engine

Knowledge Graph
- Often used synonymously with Semantic Search
- The actual graph components of semantic search
- “Semantic Knowledge Graph Search” is often used term

Semantic Web
- Established standard protocols for data formats and exchange
Of Ontologies, Taxonomies and Folksonomies

- “2020 is the year of Taxonomies”
- Can speed the time to value of technology programs
- Improve information flows across the enterprise and act as a “knowledge scaffolding” that becomes a corporate asset of increasing value over time.
- Are the missing piece in many digital programs and can mean the difference between success and failure for many enterprise initiatives.
Automating Information Management

What are the retention rules?

- Records Attributes
- Compliance Attributes
- Security Attributes
- Usage Attributes
- GDPR, PII, HIPAA

Who has access to what?

- Show me only what I need!

Automatically evaluate content repositories looking for records and applying proper records policy

Auto classify content at point of creation/capture using native tools (Word, Excel, PDF, etc.).

Complete classification across all Information Management domains – Usage, Security, Compliance, and Record

Manage all information, including physical, regardless of location
Eliminating Unknowns

Silos of Information

Value

Official Info Repository(s)

Unmanaged Repositories

Dark Data
And identify the exceptions

One Of These Things Is **Not Like The Others!**

Can you see it?
To sum things up...

- The only way to ensure proper
  - Compliance with privacy requirements – GDPR, CCCP, etc
  - Security – Both internal and external
  - Regulatory compliance – FOIA, Audits, etc.
  - Efficient search and use...

You must automate as much as possible...
The Automated IM Imperative

• Identify – Organize – Manage (Document, Voice Message, Email, Social, etc.)

• Source Independence – Capture from anywhere with a simple save

• Content is business driven for access and collaboration; policy driven for compliance, security, records, and privacy

• Governance Policy Automatically Applied via AI/ML/NLP/RPA.

• Evergreen and evolving. Persistent and consistent. Not big bang.
Automation Benefits

• It’s about getting work done differently
• Puts productive work where it belongs -- Real knowledge work vs. the mundane
• Artificial Intelligence, Augmented Reality drive continuous improvement
• Enables better access control ⇒ ensures pro-active privacy control ⇒ delivers easier discovery for other types of compliance and information requests.
• Mitigates cybersecurity issues due to improved classification consistency, etc.
• Finally gives Information Governance a business case for investment
• Supports the elimination of ROT
The Pitfalls

• Can’t be used as a shortcut for solving the problem
• Garbage in...Garbage out
• It won’t succeed without proper alignment with the people and processes already in place
Getting Started

Find
First you must find and connect all of your data sources.

Define
You need to define the basic taxonomy – Ontology – Folksonomy that represents the important relationships of your data.

Inspect
Using sophisticated search tools such as Smart Logic, you inspect all content looking for relationships in the data.

Classify
Utilizing the knowledge gathered you can begin the assisted classification of this data to support relationships and rules such as privacy, security, disposition, and use.

Manage
By bringing light to dark data, integrating disparate data sources for consistency, control, and access you create a managed environment where responding to privacy or any other request is greatly simplified.
TAKE CONTROL OF YOUR INFORMATION

Software customers in over 14 Countries

Trusted solutions for over 17 Years

A network of Global Partners

Over 500,000 Licensed Users and 10’s of Millions of Items

Gimmal exists because you must take control of your information, but by doings things differently.