The Evolution of the Data Map

ARMA Nebraska
February 21, 2018
Introduction | HBR Consulting Services

HBR’s offerings span all aspects of Law Department operations and management. We provide additional detail on each of the areas in the following slides.

Service Lines

Legal Operations Strategy + Optimization

Legal Technology

Information Governance

eDiscovery Advisory Services

Contract Lifecycle Management

CounselCommand Data Analytics Platform

Assessments

Implementations

Merger Integrations

Managed Services

Operational Improvement Initiatives

Change Management

Off-Sites/Retreats

Project Types
HBR Consulting provides expert assistance in the development of sound information governance strategies and “actionable” implementation plans that enhance legal and regulatory compliance, lower eDiscovery and storage costs, protect and secure sensitive data, and foster productivity and collaboration. Our comprehensive approach enables us to deliver end-to-end solutions that consider information governance needs, enabling technologies and business process improvements.

- **Information governance diagnostic assessments and analyses**, including strategy and implementation plans designed to achieve measurable results within realistic timeframes.
- **GDPR compliance and privacy assessments, program recommendations, data classification standards, data maps and data flows, privacy policies, and training**, all with the objective of helping protect and secure an organization’s most sensitive and critical data.
- **Evaluation, development, updating and globalization of records retention schedules** using our extensive domestic and international regulatory research capabilities.
- **Office 365 governance strategies**, including detailed plans for the deployment of O365 features for content management, collaboration and compliance.
- **Enterprise content management strategies, implementation plans, road maps and detailed project plans**.
- **Email management strategies, process improvements and enabling technology tools**.
- **Technology assessments**: Evaluation, selection and guidance on a variety of applications, systems and platforms that support the lifecycle management of records and information, regardless of media.
- **Data remediation and defensible disposition strategies**.
- **Change management services** for deploying change into the organization’s business processes and systems.
Agenda

• Data Maps
  - Background / Definitions
• Today’s Drivers
• Multi-Dimensional / 3D Data Maps
• Developing / Maintaining the Data Map
Background

• Federal Rules of Civil Procedures, 2006 revision
  - Explicitly clarifies that electronically stored data is discoverable
  - Mandates early pre-discovery attorney conference
    • Rule 26 (f) “Meet and confer”
    • Purpose: Limit the scope of eDiscovery
ESI Data Map

• An accurate, comprehensive, and understandable inventory of an organization’s IT systems which may contain ESI potentially relevant in any legal or regulatory discovery

Objectives:
• Understand the universe of potentially responsive ESI
• Effectively manage the ESI preservation and collection processes
• Minimize disruption to business and revenue-generating employees
• Reduce the cost of electronic discovery
  - Discovery cost = 50% cost of litigation
• Costs for phases are volume-driven
• Limited scope = smaller volumes = reduced cost
Information Mapping
What is an Information Map?

An Information Map is an inventory of enterprise information that tells you *what you have, where it is*, and *who is responsible* for managing it.

Common types of information maps include:

- Application
- Discovery
- Records and Information Management
- Privacy and Security

Information Maps can be invaluable tools for both prevention activities and proactive information management.
Application Inventory

• An Application Inventory catalogs the enterprise applications and systems across the enterprise, primarily for IT management purposes

• It’s primary use is to manage licensing, IT responsible parties, disaster recovery, and system health

• While useful for managing the various applications across an organization, it tends to reveal relatively little about the information each system contains
Discovery Maps

• A discovery map informs legal counsel as to the repositories of Electronically Stored Information in an organization.

• Primary role of a Discovery Map is:
  - the identification of systems commonly relevant to discovery requests
  - assistance with application legal holds
  - 26(f) meet and confer conferences

• These tend to be repository based, with a focus on backup, archival, and legal hold capabilities.

• Discovery maps were popular immediately after the FRCP rules were expanded in 2006 to specifically address electronic information, but have declined in use since that time due to the lack of commitment to keep them current.
Records Retention Schedules

- Records Retention Schedules identify information across the enterprise that needs to be retained and/or disposed of based on specific time periods.

- In most organizations, these Schedules represent the most comprehensive guide to enterprise information, and may serve as a starting point for more comprehensive mapping efforts.
  - They do not necessarily cover information which does not meet the enterprise's definition of a record.

- For ease of use, they are typically organized into functional categories. While effective for retention, these categories may not suit all information types.

- Some organizations attempt to map repositories to their schedules to some success. However, a Schedule generally lacks the granularity needed to do this with accuracy, and are not updated frequently enough to be current as to location.
Privacy and Security

- A **Privacy Data Maps** reveals where privacy related data is at a particular point in time. Through a series of interviews, surveys, and analysis, the key areas where privacy related information are identified, and key aspects of the people, processes, and systems are documented.

- A **Privacy Data Flow** tracks the step-by-step flow of a particular information type from the point it enters an organization through to the point it leaves or is destroyed. As data flows are far more detailed than information maps, they are usually limited in scope.

- **Data Classification** is concerned with how data organized and secured.
What’s the Difference?

Application maps are typically organized by the software application.

Discovery maps are organized by the repository, which may be an application or a storage location.

Records Retention Schedules are organized by record category and type.

Privacy maps are organized by privacy elements, which may include PII, PHI or PCI, as well as other sensitive elements.
Can These Efforts Be Combined?

Privacy
Applications
RIM
Discovery
Typical Silos

**RIM**
- Record Retention Schedules & Policies
- Content Management
- Information Lifecycle Management
- Defensible Disposition

**eDiscovery**
- Discovery Response
- Regulatory Requests
- Identification, Collection & Preservation
- Processing, Review & Production

**Privacy**
- Privacy Policies
- Risk Impact Assessments
- Management of Sensitive Information
- Sensitive Data Mapping
- Breach Response

**Compliance**
- Sarbanes-Oxley
- FCPA
- Enterprise Risk Management
- Audits and Controls
- Regulatory and Industry Specific

**InfoSec**
- Enterprise Security
- Cybersecurity Information Classification
- Access Controls
- Breach Detection & Response
Coordination is Good, but Collaboration is Better

Metcalf’s Law: *The more people who use something, the more valuable it becomes.*
Building the 3-D Data Map
What are the challenges?

- Tribal knowledge of ESI
- Volume of ESI
- Undocumented systems
- Lack of litigation savvy technical resources
- ESI is scattered across the organization
- Inaccessible or in a format that cannot be readily produced
Form a Cross-Functional Team

- Records and Information Management
- Legal and Compliance
- Privacy
- Information Technology
  - Application / Infrastructure Management
  - Information Security
  - Network Administration
Define Success

• **Identification of Privacy and Sensitive Information.** Particularly where it either exists outside systems where the appropriate security controls are applied, or that the broader organization was unaware was being collected or created.

• **Defensible Disposition.** An information map can help identify data that is no longer needed for a legal or business reason, and prioritize disposition efforts by the risk it presents to the organization.

• **Redundancy and Efficiency.** Through the investigative process, duplicate caches of information and redundant information management activities across business functions are often identified.

• **Identify Needs and Reduce Workarounds.** Be clear with interviewees that this is not an audit, but an improvement effort. They often reveal the use of unsanctioned tools and processes that can be turned into supported business efforts.
Define the Requirements

- Who will the main users of the map? How will they use it?
- Will this be a single purpose map, or a multipurpose map? What fields of information will it need to satisfy the business and legal needs?
- Determine whether this will replace other tools that manage similar information, or act as a supplement to them.
- Determine what kinds of information and repositories will be included. Is the goal just structured systems, or will it include unstructured data as well? What about paper?
- Contemplate all the uses before starting – it helps with both the business case for moving forward, as well as ensuring the design can scale to meet future needs.
Sample Structure

**Geography/Business Units**
- Cantera/Lisle Headquarters
- Ft. Wayne R&D
- Knoxville Accounting
- Brookfield Data Center
- NFC
- Plant

**Container/Type**
- File Servers
- External Media
- Plant Application Servers
- Exchange Servers
- Database Servers
- Data Warehouse

**Data Source**
- U Drive
- Accounting Applications
- Individual Email Box
- Peoplesoft
- Paper
Design and Build the Framework

• Can this be done with a spreadsheet, or do you need a database or third party application?
• Will there be integration with other systems? For example, how will the people responsible for various systems and information be entered and managed?
• How will information be entered into the map? How will it be kept fresh?
• How will users access the map? What are the reporting needs?
• What specific fields must be captured? How can these be logically grouped?

TIP: Pick the level of detail and number of fields with care; too much can make it hard to populate and maintain; too little can diminish usefulness.
# Potential Fields for the Data Map

<table>
<thead>
<tr>
<th>System / Application / Repository Name</th>
<th>System Business Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>System IT Support</td>
</tr>
<tr>
<td>Status</td>
<td>System Users</td>
</tr>
<tr>
<td>Deployment Date</td>
<td>Data Format</td>
</tr>
<tr>
<td>Decommission Date</td>
<td>Content / Information Types</td>
</tr>
<tr>
<td>Cloud / On Prem</td>
<td>Sensitive Data</td>
</tr>
<tr>
<td>System of Record</td>
<td>Retention</td>
</tr>
<tr>
<td>Backup Policy</td>
<td>Legal Hold Capability</td>
</tr>
</tbody>
</table>
Why a Spreadsheet May Not Be Enough

• Most organizations approach this through a series of spreadsheets to gather and organize their information mapping efforts, however the approach only shows one dimension
  - A flat table model can’t reach the level of granularity needed to accurately serve multiple purposes.

• Some try to add columns to a Retention Schedule to cover privacy or repository locations
  - Records are often in multiple locations.

• Different audiences will need different views on the information

• A spreadsheet cannot show the relationships between different kinds of information
<table>
<thead>
<tr>
<th>Vendor</th>
<th>Data Mapping Solution</th>
<th>Software / Service</th>
<th>Related Offerings</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>TrustArc</td>
<td>Data Flow Manager</td>
<td>Software Professional Services Training &amp; Certification Partner Networks</td>
<td><a href="https://www.trustarc.com/products/">https://www.trustarc.com/products/</a></td>
<td>Methodology and framework. Offer workshops on implementations, etc.</td>
</tr>
<tr>
<td>Spirion</td>
<td>Identity Finder / Spirion Console</td>
<td>Software Partner Networks</td>
<td></td>
<td>Free Trial</td>
</tr>
<tr>
<td>AvePoint</td>
<td>Compliance Guardian</td>
<td>Software Training and Certification Partner Networks</td>
<td><a href="https://www.avepoint.com/products/">https://www.avepoint.com/products/</a></td>
<td>Specialized functionality for O365/SharePoint</td>
</tr>
<tr>
<td>OneTrust</td>
<td>OneTrust Data Inventory &amp; Mapping</td>
<td>Software Training and Certification Partner Networks</td>
<td><a href="https://onetrust.com/products/">https://onetrust.com/products/</a></td>
<td>Free Trial available</td>
</tr>
</tbody>
</table>


Collecting Information and Populating the Map

• Don’t reinvent the wheel - start with existing materials, including application maps, retention schedules, and compliance lists

• Consider a mix of surveys, interviews, and workshops
  - Surveys work well for very defined information, such as technical details for repositories, but not as well for contents and use patterns
  - Interviews provide deeper understanding and reveal pain points
  - Group interviews and workshops are useful records-rich areas like Finance and HR.

• Consider starting as a pilot and ensure the framework can support the nuances of the enterprise. It can save the need to go back to users a second time.

*TIP:* While automation may help, there’s no technological silver bullet.
Potential Roles

• **Data Stewards (DS)** – Person responsible for his/her assigned data sources and ensuring timeliness and quality of the information populated and maintained in the data map. Also a potential SPOC.

• **Single Points of Contact (SPOCS)** – Person most knowledgeable of a specific data source. Familiar with the necessary logistics, resources and context of the information contained in the data source. Ability to determine if that data source meets Legal’s request for legal hold, actions to preserve and collect and to engage the appropriate SMEs to produce. Also a potential SME.

• **Subject Matter Experts (SMEs)** – Person intimately familiar with the business and/or technical aspects of the data source and the underlying files, databases and other information contained therein. Also understands how they information is resides natively, is used, can be obtained and produced.
Iterate and Extend

- Do you need to map the entire enterprise, or just certain areas to meet your needs? Given the potential scale of this kind of effort, it’s best to break it into manageable pieces.

- Try to show continual short-term value while moving toward a greater long-term goal. This helps maintain momentum.

- Look for ways to improve the information collection methods as you gain experience with the various methods.

- Prioritize based on areas of highest risk or greatest business benefit.

TIP: Track your progress so you know which areas are fairly complete, and which are works in progress that should not be depended on.
Maintenance

• Determine ownership and responsibilities up front
• Develop a plan to keep the information map evergreen
  - Information can become stale in a matter of months.
• Documents processes and procedures for the use and maintenance of the information map
• Work with IT and other areas to develop “triggers” so you are notified when new systems come online, or when new privacy-related information collection processes are adopted

TIP: Log when data was last updated, and use that information as a guide to refreshes.
Lessons Learned

• Don’t try to address all needs at once; take an iterative approach and continually show value to the effort. **Start with high risk areas, and move on from there.**

• Identifying information “owners” is often difficult, but worthwhile in the long run.

• **Don’t underestimate the amount of time and level of effort needed;** doing this right takes time.

• Contemplate all the various uses and users at the outset; **getting the framework right at the beginning** will avoid rework and needing to go back to the business multiple times.

• Ensure you have the people and processes to **keep the map fresh;** dated information can undermine confidence and make it more of a burden than a benefit.

• Don’t forget cloud and SaaS applications – understanding the flow of information to external vendors is extremely valuable.