



PROGRAM MATERIALS

Program #2903

January 9, 2019

How Do You Solve a Problem Like E-mail?

**Copyright ©2018 by Therese Craparo Esq. and Anthony
J. Diana, Esq., Reed Smith LLP
All Rights Reserved.
Licensed to Celesq®, Inc.**

Celesq® AttorneysEd Center
www.celesq.com

**5301 North Federal Highway, Suite 180, Boca Raton, FL 33487
Phone 561-241-1919 Fax 561-241-1969**



How Do You Solve a Problem Like Email?

January 9, 2019



ReedSmith
Driving progress
through partnership

Today's Speakers



Therese Craparo
Partner



Anthony J. Diana
Partner

Contents

- What is the **“Problem”** with Email?
- Regulatory, Compliance & Litigation Risks
- Managing the Risks of Email
- Case Study: Messaging Migration
- Case Study: Messaging Classification

What is the “Problem” with Email?

What is the “Problem” with Email?

Data Management – Including the Management of Email – is Increasingly Recognized as Critical to Corporate Governance and Risk Management but Effective Management of Email is Increasingly Challenging Considering Historical Data Volumes

- **Volume:** Email volumes are not decreasing – they are increasing exponentially. Despite the proliferation of new communication technologies, the average employee sends/receives 200 emails *per day*, which amounts to 1,000 emails *per employee per week*, and 48,000 emails *per year*
- **Corporate Hoarding:** Historically, many organizations have not routinely disposed of emails, or they have routinely disposed of emails but allowed employees to store PSTs on hard drives, shared drives, home drives or other locations. That means many organizations have tens of millions or, in some cases, tens of billions of historical emails
- **Organization & Storage:** Over time, many organizations have allowed employees to use (or have not been able to prevent employees from using) email as an electronic filing cabinet. That means individual email may contain critical business information and company records
- **Technology:** While the technology relating to email is improving, technology that would allow an organization to easily and automatically search, sort, group, categorize and organize large volumes of email across an organization in place – particularly with millions or billions of emails – remains elusive

What is the “Problem” with Email?

The Legal and Regulatory Risks Associated with Email – And Particularly with the Over-Retention of Email – are Also Increasing as Organizations, Regulators, Employees and Consumers are Beginning to Recognize the Privacy, Cybersecurity and Litigation Risks Associated with Retaining Vast Volumes of Email

- **Compliance/Regulatory:** Organizations have record retention obligations – both those that are statutorily required and those that are based on substantial business needs – that must be met both for compliance/regulatory purposes and to facilitate business operations
- **Privacy:** Increasing global focus on data privacy, in particular the GDPR and the CCPA, have significantly increased the awareness of the risks of over-retention of data and the consequences of non-compliance
- **Data Breach & Cybersecurity:** Recognition that over-retention contributes to increasing the risk of data breaches (including both personal data and sensitive business information) from both industry and regulatory perspective. Regulators are now looking to data minimization as a key cybersecurity strategy
- **Litigation Liability & eDiscovery:** Email is one of the most sought-after data sources in litigation. Unlike the risk landscape 10-15 years ago, litigation liability and eDiscovery costs are increasingly viewed as weighing in favor of better data management and against over-retention

Regulatory, Compliance & Litigation Risks

Regulatory & Compliance Risks

Organizations Have Obligations to Retain Certain Types of Data and Have Obligations to Dispose of Data No Longer Needed for Business, Regulatory or Legal Purposes for Both Privacy Concerns and Cybersecurity Risk

- **Record Retention Requirements.** Organizations have obligations – legal, regulatory, business, contractual – to retain certain types of Records and typically have policies and procedures governing the retention and disposal of those Records. These include data related to: Audit, Financial, Tax, Human Resources, Environmental, Compliance, Customers, Sourcing, Business Continuity, Corporate Governance, etc.
- **General Data Protection Regulation (GDPR).** Requires the minimization of personal data, including retaining such information for no longer than necessary for the purposes for which the personal data was collected; Includes right to be forgotten; Potentially significant fines for non-compliance
- **California Consumer Privacy Act (CCPA).** The CCPA requires businesses to take reasonable steps to dispose of, delete or destroy, or arrange for the disposal, deletion or destruction of customer records when the records are no longer to be retained pursuant to the statute
- **Data Breaches & Reputational Damage.** See, e.g., Equifax, Target, Sony, Uber, Yahoo, US Office of Personnel Management, Anthem, Facebook, Google, Home Depot, Marriott, Ashley Madison, eBay, Under Armor, Saks, Amazon ...

Regulatory & Compliance Risks

Regulators are Actively Engaged in Cybersecurity, Record Retention and Data Management Enforcement, Including Data Risk & Managing Data Flows and Deletion

- **HHS Office for Civil Rights** goal of auditing covered entities for compliance with HIPAA Privacy and Security Rules by looking at their policies and procedures for handling and storing data
- **Federal Trade Commission** recommends data minimization (limiting the collection of consumer data, and retaining that information only for a set period of time) for addressing two key privacy risks: first, the risk that a company with a large store of consumer data will become a more enticing target for data thieves or hackers, and second, that consumer data will be used in ways contrary to consumers' expectations
- **NYS Department of Financial Services** issued rules on cybersecurity for covered financial institutions to establish cybersecurity programs; Focus on *information* (not just information systems) and expressly calls out *deletion* of data
- **Financial Regulators** are actively enforcing data retention requirements for email, e.g., LPL Financial fined \$7.5 million by FINRA for messaging failures, including losing 80 million messages during migration to new email archive; Barclay's fined \$3.7 million for failing to retain Bloomberg messages and attachments and failing to detect wider issues with email retention; Citigroup fined \$750,000 for failure to perform Quality Assurance tests after journal migration which failed to assign correct retention periods

Litigation Risks

Unlike the Risk Landscape 10-15 Years Ago, Litigation Liability and eDiscovery Costs Are Increasingly Viewed as Weighing in Favor of Better Data Management and Against Over-Retention

- **Litigation Liability:** Seemingly innocent comments, jokes, or candid opinions expressed by non-legal personnel can be taken out of context or look unlawful in hindsight, with significant consequences in later litigation
 - *Oracle v. Google (N.D. Cal.)*. Suit filed in 2010 relying on informal email discussions between engineers in the 2005-2007 time frame, survived several appeals and again in trial, with the initial damage estimate almost doubled to \$9.3 billion
 - *DOJ v. Standard & Poor's (C.D. Cal.)*. DOJ's \$5 Billion lawsuit filed in 2013 against Standard & Poor's used documents from 2004 and 2007 to demonstrate executives criticizing investment-grade ratings and documenting deterioration in housing markets before 2008 financial crisis; lawsuit settled for \$1.4 Billion with lengthy stipulation reflecting damaging facts revealed in "voluminous" discovery
- **E-Discovery Costs:** When accounting for e-discovery costs across all legal matters, the cost exposure for over-retention of email can exceed tens of millions of dollars
 - For a single case with 10 custodians who have 1 year of retained email, the overall e-discovery cost could range from \$75,000 to \$450,000. For 6 years of data, the cost balloons to a range of \$450,000 to \$2,700,000, according to survey by RAND Corp.

Litigation Risks

The 2015 Amended Federal Rules, and Subsequent Case Law, do Support Organization-Wide, Routine Disposal of Email Based on Policy and Proportionate (Not Perfect) Preservation

- **Federal Rule of Civil Procedure Rule 37.** Where ESI that should have been preserved is lost because party failed to take reasonable steps to preserve it and it cannot be restored/replaced: (1) if prejudice, measures no greater than necessary to cure; or (2) if intent to deprive, then presume unfavorable, adverse inference, default judgment
- **Rule 37, 2015 Advisory Committee Notes.** “This rule recognizes that ‘reasonable steps’ to preserve suffice; it does not call for perfection”
- ***Matthew Enter., Inc. v. Chrysler Grp. LLC*, 2016 WL 2957133 (N.D. Cal. May 23, 2016).** “Rule 37(e) now provides a genuine safe harbor for those parties that take ‘reasonable steps’ to preserve their electronically stored information.”
- ***EPAC Techs., Inc. v. Thomas Nelson, Inc.*, 2016 BL 502408, *4 (M.D. Tenn. Jan. 27, 2016).** “[T]he scope of preservation should be fairly tied to a good faith, subjective assessment of the likelihood of litigation and to a reasonable, objective assessment of the proportionality factors laid out in FRCP 26(b)(1), of the claims and defenses most likely to be asserted and of the sources and custodians most likely to hold responsive information.”
- ***Barnett v. Deere & Co.*, No. 2:15-CV-2-KS-MTP, 2016 WL 4544052 (S.D. Miss. Aug. 31, 2016).** “Numerous federal courts have recognized that the routine destruction of information pursuant to a reasonable data retention policy is not grounds for sanctions.”

Litigation Risks

Nonetheless, Organizations Must Comply with Their Legal Hold Obligations and the Risk of Sanctions for Improper Disposal Remains a Genuine Concern

- ***O'Berry v. Turner*, No. 7:15-CV-00064-HL, 2016 WL 1700403 (M.D. Ga. Apr. 27, 2016).** Imposing adverse inference instruction despite loss of evidence resulting from a series of errors/mistakes by defendant and no apparent "intent to deprive."
- ***Klipsch Group, Inc. v. ePRO E-Commerce Ltd.*, 880 F.3d 620 (2d. Cir. Jan. 25, 2018).** Plaintiff moved for sanctions after defendant failed to implement an adequate legal hold, including failing to preserve emails. Although damages in matter were likely only \$20,000, the Circuit Court found \$2.7 million sanction was justified and not punitive as "The proportionality that matters here is the amount of the sanctions was plainly proportionate – indeed, it was exactly equivalent – to costs ePRO inflicted upon Klipsch in its reasonable efforts to remedy ePRO's misconduct".
- ***GN Netcom, Inc. v. Plantronics, Inc.*, No. 12-1318-LPS, 2016 WL 3792833 (D. Del. Jul. 12, 2016).** Plaintiff in antitrust action moved for sanctions based on intentional deletion of emails by defendant's senior manager. After defendant issues legal hold, the senior manager sent email to his sales team calling for the deletion of email chains that address competition from plaintiff, and also deleted a number of his own emails. Court found defendant failed to take "all reasonable steps" to remove the lost email (including choosing not to pay a forensic expert to analyze what was deleted), and imputed the manager's conduct to defendant. Court rejected further discovery, and imposed a \$3 mm monetary sanction

Litigation Liability

Even Mishandling of BAU Data, or an Inability to Explain Data Anomalies After Data Migrations, Can Have Significant Litigation Risk

- ***Cat3, LLC v. Black Lineage, Inc.*, 164 F. Supp. 3d 488 (S.D.N.Y. 2016).** Sanctions imposed where court found that the metadata of emails at issue was manipulated and intentionally spoliated, rejecting the argument that such an anomaly might have resulted from the migration of the emails from Gmail to Microsoft Office 365
- ***Harkabi v. SanDisk Corp.*, 275 F.R.D. 414, 419 (S.D.N.Y. 2010).** Sanctions awarded for failure to supervise discovery where “there [was] no evidence that in-house counsel was involved with the transfer of email archives to the Evault system, when [relevant] emails were deleted from readily-searchable servers”
- ***Krumwiede v. Brighton Assocs.*, No. 05-cv-3003 (N.D. Ill. May 08, 2006).** Court found that even if files at issue were only modified and not deleted, “the changes to the file metadata call the authenticity of the files and their content into question and make it impossible for [a party] to rely on them”

Managing the Risks of Email

Managing the Risks of Email

Retention Requirements

- Industry-Specific Requirements (Financial, Pharma & Life Sciences, Industrial, etc.)
- High-Risk Data (Tax, Employment, Environmental, Audit, Financial, Customer, etc.)



Privacy/Data Protection / Cybersecurity Challenges

- HR, NPPI, PII, Privilege, IP, Business Sensitive Information
- CCPA, GLBA/Cross-Border, Localization, etc.



Litigation Exposure/ eDiscovery

- Preservation/Spoilation
- Collection/Production



Market Pressure & Business Need

- User reliance on messaging data
- Product Development, Data Analytics, Marketing, etc.
- Leveraging customer information



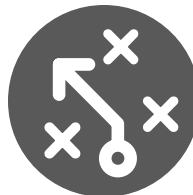
New Technology & Vendor Management

- Cloud, AI, O365, etc.
- Migrations, Implementations, etc.



Business Transformation

- Mergers, Acquisitions, Divestitures
- Closures, Infrastructure/Corporate Changes



Email management requires balancing the value of information to the business with the legal/regulatory/operational risks



Managing the Risks of Email

With the Growth in Technology Platforms that Allow for Comprehensive indexing of documents/data, such as Microsoft O365, developing processes around the technology will be transformative for the management of information

- Goal generally will be to only retain data as required for legal, regulatory, or business purposes
 - Ability to identify and categorize email as containing Records or other data that needs to be retained
 - Ability to identify/screen data containing personal data, business sensitive data, highly confidential data
- Requires leveraging people, process and technology
- To effectively manage messaging data, two critical components:
 - Centralize messaging data in technology platform with key functionality order to manage it effectively—likely necessitates migrations to new platform
 - Development of sustainable process to identify/classify messaging data for legal, regulatory or business needs (i.e., Records) with required retention applied

Case Study: Messaging Migration

How a multi-national financial institution defensibly disposed of billions of emails, and developed a process for BAU disposal of messaging data

Case Study: Messaging Migration

Multiple legacy messaging archives from mergers, acquisitions, and system changes housed in datacenters across the globe (NAM, EMEA, APAC)

- Operational costs and burden of operating legacy archives rising, all while data footprint continued to grow in current archives
 - More than 3 billion messages in legacy archives
 - More than 8 billion messages in current archive
 - Projected to reach 25 billion by 2020
- Legal and regulatory obligations required the retention of significant data
 - Thousands of employees on legal hold
 - Regulatory retention obligations for Broker-Dealers, Investment Advisors, etc.
 - Retention obligations in several foreign jurisdictions

Case Study: Messaging Migration

Senior Management's goal was to retain only messaging data needed for legal, regulatory, or business needs and to enable go-forward messaging management and routine disposal

- Key drivers of change
 - Initial plans to expire data interrupted due to significant litigations; other litigation concerns
 - Merger integration included evaluation of data governance practices
 - IT costs, efficiency, and consolidation
- Focus on identifying messaging data needed only for legal, regulatory or business needs and enable expiration
 - Validated implementation of policy that key Records were not routinely stored in email
 - Set retention periods on all email for all employees and shared mailboxes based on business, legal and regulatory requirements
 - Migrate data from seven legacy compliance archives into three regional compliance archives (NAM, EMEA and APAC)
 - Implement legal holds
 - Expire data on a regular basis
- Initial planning centered on message-by-message migration from all legacy archives overseen only by IT
 - Initial scope was all message data; more than 3 billion legacy messages
 - Timeline and costs determined to be prohibitive

Case Study: Messaging Migration

Migration plan was re-evaluated and re-designed to facilitate efficient, cost-effective and defensible migration

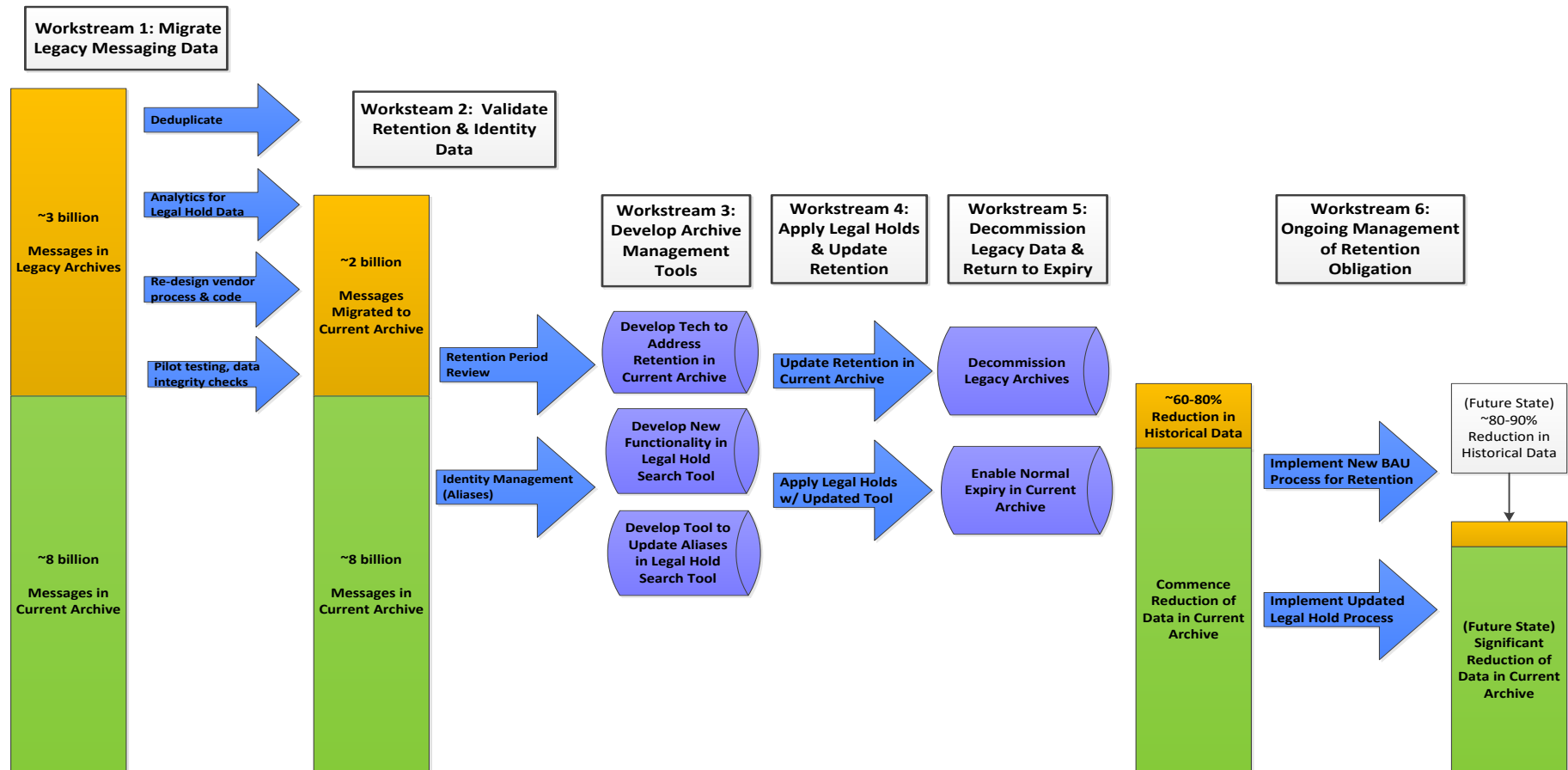
- Re-evaluation of plan led to key insights
 - Needed a cross-functional team, including Legal, Compliance, IT, and Records
 - Technology innovation and analytics on messaging data were critical
 - Go-forward messaging management improvements needed
- New migration plan focused on migrating only the data needed for legal, regulatory, or business purposes
 - Migrate data associated with custodians on legal hold
 - Migrate data subject to regulatory retention obligations
 - Migrate only unique, deduplicated messages in above categories
- Plan took into account overall management approach that would be applied in the current archive
 - Designed migration process to allow for routine expiry based on message date and the assigned retention period (i.e., avoid over-retention)
 - Designed migration to permit unique legal hold data to expire after legal holds are lifted (if not subject to other retention requirements)
 - Designed migration process to allow for clear identification of migrated data

Case Study: Messaging Migration

Retention periods were re-evaluated firmwide to determine appropriate retention periods for both individual employees and shared mailboxes, retention periods and reasons were approved and recorded, and processes were implemented to monitor and maintain accurate retention periods

- Assigning retention based on employees
 - Each user was assigned a specific retention period by business unit or corporate department
 - Default retention was set at six months
 - Extended retention based on legal/regulatory requirements, including risk that messaging data would likely contain legally required records
 - Exception process implemented by senior management based on business line, role or function for business exceptions
 - Extensive process for changing/correcting retention
- Assigning retention based on shared mailboxes
 - Default retention was set at six months
 - Shared mailboxes associated with regulated businesses assigned extended retention
 - Liberal allowance for extended retention based on business need (repository of records)

Case Study: Messaging Migration



Case Study: Messaging Migration

Challenges & Lessons Learned

- Vendor Management: Never trust, always verify
- Technology Limitations: Learn the difference between the ideal world and the real world – Question everything
- Identity Management: Your data is a snapshot in time, but your organization is always changing
- Senior Management Expectations: It's better to under-promise and over-deliver
- Stakeholder Involvement: This is not “just” an IT project, managing data risk is cross-functional
- Defense of Process: Develop the process as if you are under legal and regulatory scrutiny - document every decision and challenge every assumption
- There Will Be Problems: They don't have to derail your team – it's not whether you get knocked down, it's whether you'll get back up

Case Study: Messaging Classification

How a health care company
identified/classified Records within
1.5 billion messages, and developed
a BAU process for record
classification

Case Study: Messaging Classification

Senior Management's goal was to only keep messaging data as long as needed for legal, regulatory, and business purposes

- Volume of email and historical treatment of email posed challenges to immediate implementation
 - All messaging data had been retained for over 10 years in an email archive (over ~1.5 billion messages, with ~3.5 million added daily)
 - PSTs had been retained throughout the enterprise personal drives, network shares, mailboxes, etc.
 - Key Records, including Records required for regulatory purposes, were routinely retained in email by employees
- Solution focused on consolidation and leveraging technology
 - Migrated data into an archive with additional necessary functionality
 - Eliminated PSTs outside of archive
 - Apply standard email retention of six months on all messaging data unless identified as a Record or subject to legal hold
 - Apply email Records auto-classification to identify and classify email Records

Case Study: Messaging Classification

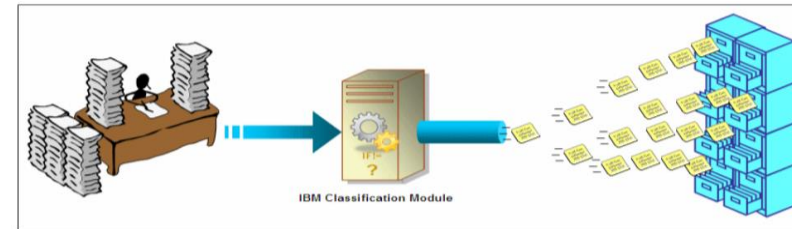
Employed Records Auto-Classification Tool to leverage machine learning algorithms to automatically identify and categorize email Records

- Ongoing Email Record Auto-Classification.
 - Automatically identify and classify Records categories in emails on an ongoing basis to retain Records in email and apply the appropriate Records category retention periods
 - May be used as an alternative or complement to user tagging of Records
- Employee Tagging Tool:
 - Enables employees to manually identify and classify emails that contain Records
 - May be used as an alternative or complement to ongoing auto-classification
- Historical Email Record Auto-Classification. Identify and classify Records in historical email systems to apply the appropriate Records category and retention

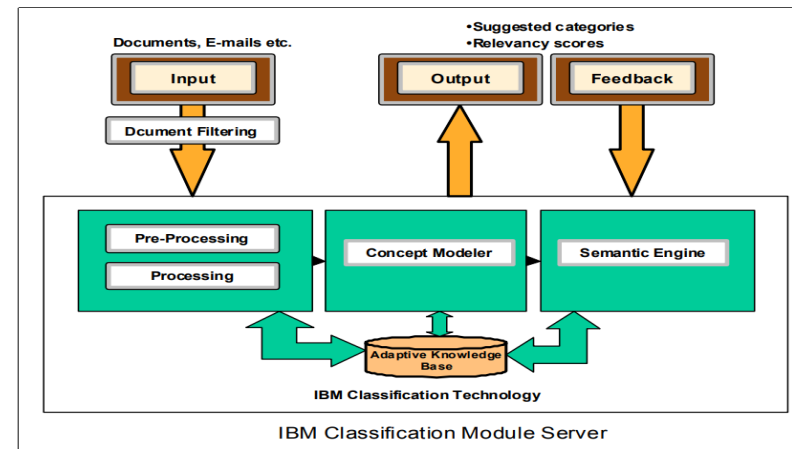
Case Study: Messaging Classification

Classification Tool Overview

- Classification Tool
 - Software tool that uses natural language processing and sophisticated semantic analysis techniques to analyze and categorize text
 - When emails are sent to the classification tool for analysis, classification technology identifies the categories that are most likely to match the text
 - Note: Presentation features the IBM Infosphere Classification Module. Other classifications solutions are also available
- Knowledge Base
 - The Classification Tool works with an adaptive knowledge base ("Knowledge Base"), which is a set of collected data (e.g., emails) that the Tool uses to analyze and categorize content items
 - To perform well in identifying and classifying email Records, the Knowledge Base must be trained with a sufficient number of sample documents/texts that are properly classified into categories



Source: IBM Infosphere Classification Module documentation



Source: IBM Infosphere Classification Module documentation

Case Study: Messaging Classification

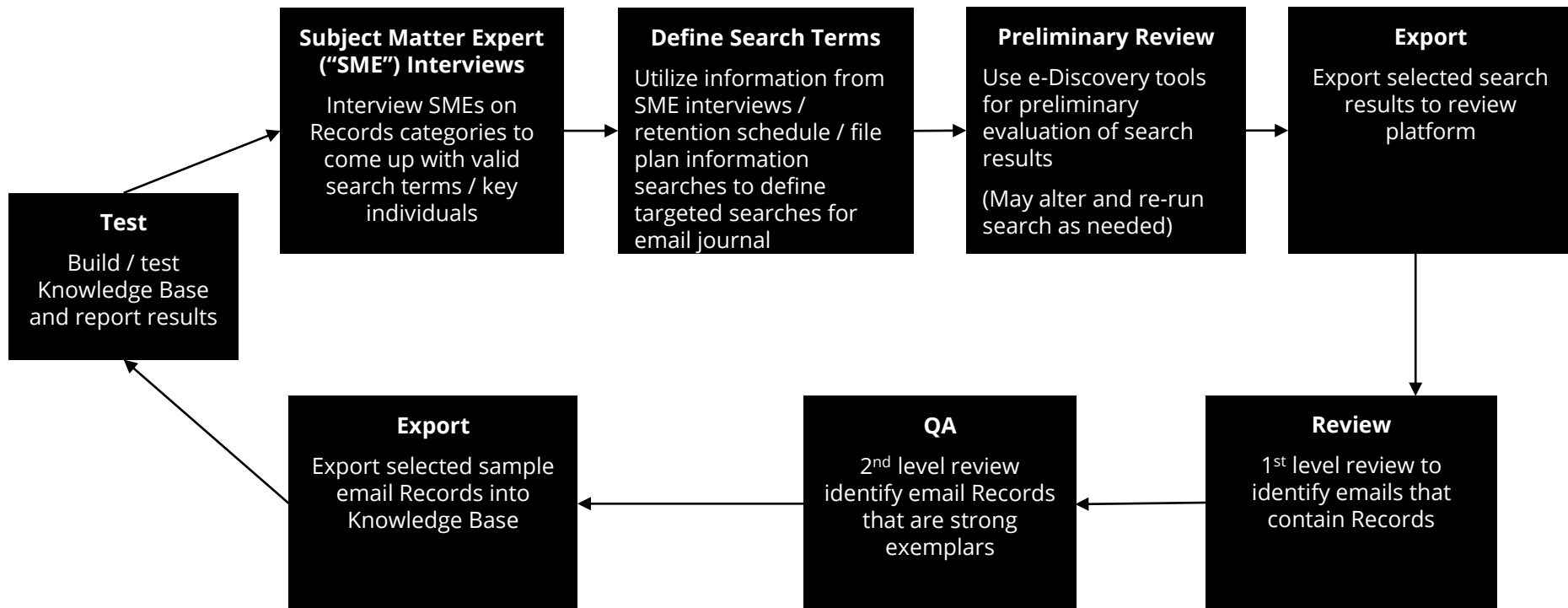
Develop process to assign each Records category a specific risk rating to enhance the effectiveness of the Records Categorization Tool in meeting client's legal, regulatory and business retention needs

- Records category risk ratings enabled the organization to:
 - Account for legal and regulatory retention obligations, as well as operational needs
 - Assess the likelihood that a Records category with legal, regulatory requirements and a significant business need would only exist in email
 - Prioritize high- or medium-risk Records categories in training the Records Categorization Tool

Risk Rating	Criteria		
	Legal / Regulatory Retention Requirement	Impact	Email Exclusivity
High	Legal or regulatory requirement exists	"High" Or "Medium"	"Yes" or "Possibly"
Medium	Legal or regulatory requirement exists	"High" Or "Medium"	"No"
Medium (Alternate)	No legal or regulatory requirement	"High"	"Yes" or "Possibly"
Low	Remainder		

Case Study: Messaging Classification

Required process to order to define and develop the knowledge database that the software will use for auto-classification ("Knowledgebase Development")



Case Study: Messaging Classification

Define Key Performance Indicator (KPIs) assist in evaluating Knowledgebase performance both overall and in the context of the Employee Tagging Tool and Journal Archive Classification

- Terminology
 - **Classification Success Rate:** The percentage of the sample email Records that the classification tool classifies correctly for a specific Records category
 - **Recall:** Number of emails auto-classified correctly for a category / total number of email exemplars for that category
 - **Precision:** Number of emails auto-classified correctly for a category / total number of emails auto-classified to that category
- Example KPIs
 - **Total Cumulative Success:** 80% average classification success rate overall across all Records categories, particular high and medium-risk categories
 - **Tagging Tool Performance:** 80% classification success rate within the first 3 recommended categories displayed to the end-user for the “High” and “Medium” risk record categories
 - **Archive Classification:** 80% Recall Rate for the “High” and “Medium” risk categories, while maintaining an 60% Precision Rate

Case Study: Messaging Classification

Evaluate, test and validate effectiveness of auto-classification

- Knowledgebase auto-classification should be evaluated using a dataset with email Records different from those used for training
- To compile a testing dataset that included a statistically significant number of email Records exemplars for high- and med-risk categories, perform the following:
- Pull email Records exemplars from the following sources:

Source	Description
Reviewer Tagged Email Records	• 120,000 emails previously categorized during bulk email review but not included in Knowledgebase
Employee Mailboxes	• All emails sent or received in a month-long period by an employee identified as handling email Records for a category
Employee Tagged Email Records	• Emails employees have already tagged as Records using Tagging Tool (and which were verified as Records during QA)

- Set performance goals for the testing phase (e.g., high recall (>80%) and moderate precision (>60%))
- Based on the results of the testing phase, fine-tune the threshold values at the category level to optimize recall/precision and reduce risk

Case Study: Messaging Classification

Challenges & Lessons Learned

- Email exemplars for each category must have distinct content for the classification tool to distinguish them from the rest of the email population
 - Identifying “good” exemplars to populate and train the database may involve a time- and labor-intensive process of:
 - Interviewing Subject Matter Experts
 - Collecting and reviewing email Records
- A “good” exemplar has the following characteristics:
 - Only associated with a single Records category
 - Text-based (cannot classify unreadable attachments, such as images)
 - Includes distinct keywords (potentially associated with Records category)
- Organizations may need to redefine or combine categories with similar exemplars to reduce overlap and improve performance, e.g., board of director materials vs. financial reporting; service agreements vs. purchase orders

Case Study: Messaging Classification

Challenges & Lessons Learned

- One possible approach to improving performance is to lower the thresholds for categories that are higher risk or have a longer retention period.
 - Increases the likelihood that email Records for that category will be retained, but also that other types of email Records will be misclassified
 - Less risk associated with misclassification if email Records are tagged to categories with same retention periods or longer
- Email Records Auto-Classification should be supplemented with other Records retention strategies
 - Email Records Auto-Classification will not correctly categorize emails 100% of the time
 - Other strategies include:
 - Policies requiring employees tag email Records and retain other types of Record documents in designated Systems of Record and provide
 - Employee Records retention training

Questions?



Therese Craparo
tcraparo@reedsmith.com
212-549-0421



Anthony J. Diana
adiana@reedsmith.com
212-549-0332

Thank you!