TOTAL DATA INTELLIGENCE EGUIDE

DRAWING THE LINE ON EDISCOVERY LINEAR REVIEW:

A Guide to Implementing Smart Review

nuix

A NEW AGE, A NEW APPROACH TO EDISCOVERY

There's no such thing as simple discovery in the New Information Age. Document review is inherently complex—with rolling productions, multiple phases of review, and distributed teams with different levels of expertise. There are also varying document types to review including emails, audio, video, spreadsheets, text messages, social media, and online chat services.

Traditionally, eDiscovery teams have used a linear review process for their cases' legal and strategic needs. With this approach, legal teams tend to derive key search terms to produce the documents to review while only having superficial knowledge of the matter. Review teams then proceed "in the dark," with each document reviewed in isolation from its larger context until they get to the last document.

The sequential process results in an incomplete understanding of the facts—as what's learned in later documents can't be automatically applied to earlier documents. Linear review can also fall victim to biased or inconsistent reviews across reviewers with differing expertise or perspectives, and reviewer fatigue.

In addition, reviewing documents one at a time is time consuming and costly—especially for highly paid attorneys. Despite this, 38% of corporations have yet to employ new eDiscovery technologies that can alleviate much of this effort.[1]

The New Information Age requires a better way of approaching eDiscovery review—one that provides Total Data Intelligence. That ensures all documents are coded consistently and completely to satisfy your cases' needs. That clusters related information by format, keywords, or custodian. That allows you to visualize connections between documents and people. That drives down the costs and wasted effort of linear review and allows your legal team to spend more time developing case strategies and improving the outcomes of all eDiscovery projects.

To transform your linear review process into a rock-solid, repeatable workflow, you need to adapt the Smart Review approach: deploying visualization tools and analytics within a single platform to empower your eDiscovery teams to find the facts while eliminating the waste in the review process.

This eGuide will explore how a solution that enables Smart Review provides better, faster, more accurate ways to find the evidence that matters and take action on it, including:

- How better preparation brings better outcomes
- The different tools of a Smart Review platform for the New Information Age
- Applying analytics, even in the small, everyday cases
- How it all comes together with a platform that delivers Total Data Intelligence.

Half of industry professionals say that the most significant impact of the eDiscovery process is the variety, volume, and velocity of disparate data.[2]

[1,2] BDO Consulting, *Inside eDiscovery: The State of eDiscovery According to Corporate Counsel*, 2015

Copyright ©2019 Nuix. All rights reserved.

CHAPTER 1:

Better Preparations, Better Outcomes

In the New Information Age, eDiscovery needs a fresh approach. Ever-growing data volumes and diversity make traditional methods obsolete and can defy attempts to automate parts of the process. You need a more holistic approach that begins before review.

Smart Review starts with better ways to ingest and analyze data before you read a single document. These will maximize your ability to identify relevant information during review. This takes some preparation.

Pre-Review Preparation	
ASSESS THE DATA SOURCES.	Assess the likely evidence sources to ensure your processing engine can accurately index, enrich, and analyze all relevant data and make it accessible in a single review platform.
ASSESS THE SCOPE AND EFFORT REQUIRED TO REVIEW THE DOCUMENT COLLECTION.	You need to know where to start by defining the early strategies, how to prioritize the work, etc. Answer the following questions to assess the scope and effort required: • How many reviewers will be working on the case/project? • What are the case/ project deadlines? • How many documents need to be reviewed per day? • How exhaustive does your review need to be? Do you need to see 50% of relevant documents? 75%?
REVIEW THE QUALITY AND ACCURACY OF KEYWORDS IN ADVANCE OF DOCUMENT REVIEW, AND CLASSIFY KEY CONCEPTS AND TERMS IN THE DOCUMENT COLLECTION.	If you agree to search terms that are too broad, you'll spend too much time reviewing documents that don't matter, which is an expensive waste. Reviewing a random sample of documents before diving into full review helps you identify the right key terms within relevant documents. Leveraging document mapping, you can more easily dismiss clearly irrelevant document clusters and more closely investigate documents that contain concepts of interest. This will help you identify the most relevant documents—providing a solid place to start for reviewers. We'll explore this further later in this eGuide.

Copyright ©2019 Nuix. All rights reserved.

UNDERSTAND THE CONCEPTS AT THE CASE LEVEL.	Understand concepts by taking the following steps: • Index data • Content analysis • Concept analysis • Clustering—visualize data in a word cloud or document map.
ANALYZE COMMUNICATION PATTERNS ACROSS INDIVIDUALS AND NETWORKS.	Pinpointing key custodians or persons of interest—including those you may not have originally identified—is key to any case. Relationship analytics can provide a helpful map to do just that, uncovering links you might otherwise not have expected. This is also a useful tool to help identify inappropriate communications which, if highlighted early, could make the difference between proceeding with a matter or attempting to bring it to a quicker resolution.[3]
IDENTIFY AND ISOLATE KEY PERIODS OF ACTIVITY DURING THE MATTER.	Determining a reasonable timeframe and relevant custodians can help you narrow the document set based on date ranges, search terms, custodians, and file types. Mapping documents to a timeline can also help isolate key periods of activity.
RECOGNIZE FOREIGN OR "CODE" LANGUAGES FOUND IN THE CASE DOCUMENTS.	Language is slippery. Projects and teams use jargon and code names. Unless you're part of that team, there's no way to anticipate those terms. There may be an entire lexicon beyond your grasp. Concept analytics can show you aggregations of concepts to focus your review and reveal code names of potentially nefarious conduct.

When you spend the time upfront to dig into the make-up of the case information and analyze where to concentrate review efforts, you can make a significant impact on the eDiscovery process, increasing efficiency in the long term.

YOU WILL ULTIMATELY:

Drive down the cost of review.

Increase the speed and dexterity of review teams.

Provide revenue opportunities for service providers and law firms.

[3] Nuix, *Early Case Assessment: Evolving from Tactical to Practical*, 2015

Copyright ©2019 Nuix. All rights reserved.

CHAPTER 2:
The Tools for Smart Review

How can the same data provide one litigator the secret sauce while the other side suffers from the sting of not seeing how the information fits together? Smart Review uncovers better evidence to discover the truth that is integrated faster into case strategy, giving legal teams a key competitive advantage.

Nuix visualization and analytics tools support smarter, more efficient review. These visual cues can help focus the eye on what's most important, allowing you to quickly identify and prioritize key documents and terms for review. By leveraging visual analytics, content analytics, concept clustering, metadata metrics, and data patterns, you can quickly get a sense of what you're seeing without having to read every single document. These tools support you to identify relationships and make connections faster and easier.

WHY USE VISUAL REPRESENTATIONS?

It comes down to human cognition. Our minds are extremely well equipped to recognize patterns, but comparatively poor at computational assessments. Visual representation provides a shortcut to pattern recognition and identification of trends[4]—so you can see the connections clearer than ever before.

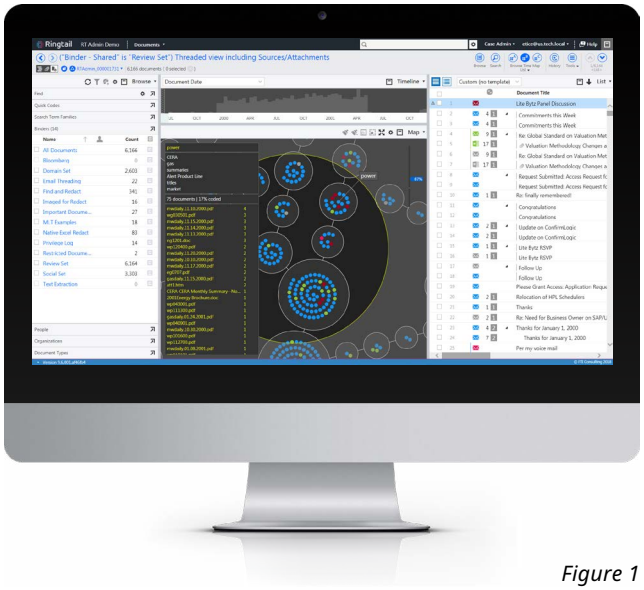

Figure 1

CONCEPT CLUSTERING:

Concept clustering is designed to accelerate the review and analysis of large document sets, transforming disparate, unstructured content—such as emails, documents, and social media content—into a highly organized index. This index then allows Document Mapper (Figure 1) to tell the story through interactive visualization, rather than your reviewers trying to decipher it.

Concept clustering begins with content analytics during the data ingestion phase of eDiscovery. During this process, the software analyzes and classifies all incoming content to identify key features in every document. For example, it can:

- Extract nouns and noun phrases from all emails, documents, and social media content
- Identify concepts and weight them based on frequency within the documents
- Identify overlapping concepts
- Score and assign clusters based on similarity.

Once your software has completed concept clustering, you can use Document Mapper to visualize search results by clustering conceptually similar documents together—creating a micro view of how they're conceptually related. These clusters are presented as a spiral of dots, with each dot representing a unique document. The document with the most common concept is at the center. At a macro level, multiple clusters sharing dominant concepts are arranged in spines. The entire visualization depicts the results of tens to hundreds of parallel concept searches—so you can easily see what's related, where discrepancies arise, and who's talking to who.

[4] David Houlihan, James Haight, *Evaluating the Opportunities Presented by Analytics and Visualization in Legal Review*, Blue Hill Research, November 2015

Copyright ©2019 Nuix. All rights reserved.

When the data volumes are really big, it's time to switch visualizations. Mines (Figure 2) are also powered by concept clustering but operate on a larger scale, up to millions of documents. The mines tool ingests documents, emails, and attachments to identify the concepts that best describe each cluster of items. You can then identify trends and investigate the strength of conceptual relationships between document populations. This helps to prioritize documents for further review.

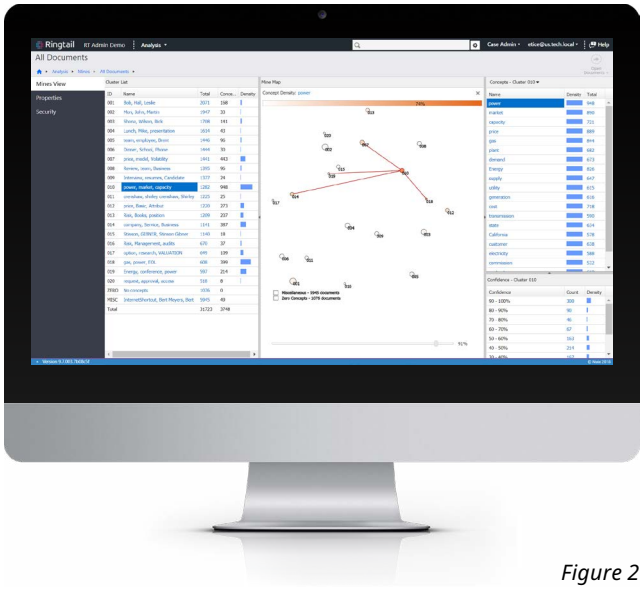

Figure 2

SOCIAL NETWORKING VIEW:

While documents are an important source of information for litigation, investigations begin by understanding who is communicating with whom about particular topics. Social network views depict the relationships between people and organizations along with the flow of information—such as emails, phone calls, chats, and financial records.
Email metadata (from, to, CC, BCC, date, time, etc.) provides the organizing information to show three types of communication relationships: person to person, organization to organization, and person to organization.

Combining this view with a concept cloud shows the concepts discussed in these different types of communications. You can view the communicators involved to see connection between them—and even concepts that individuals discussed together.

If you need to determine which concepts were discussed within a certain timeframe, or even determine what timeframe you should be exploring, the timeline feature (Figure 3) creates a histogram of the documents and communications. With this feature, you can select time ranges and see where the documents are on the map.

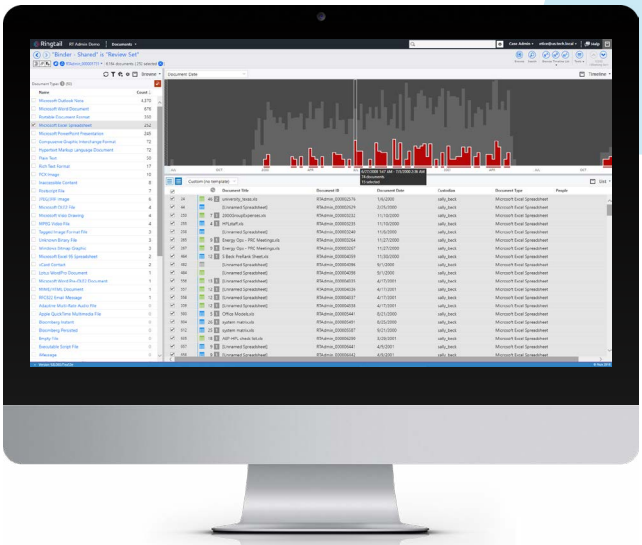

Figure 3

DOCUMENT MAPPER EXAMPLE USE CASES INCLUDE:

Case assessment—Corporations and their law firms have used Document Mapper for more than 15 years to develop case strategy on a wide range of matters of all sizes and significance.

Litigation review—Document Mapper has been proven on the world's largest litigation matters, including anti-trust, patent infringement, and subprime cases. It enables review teams to quickly and accurately categorize, annotate, and redact documents in accordance with review guidelines and legal requirements.

Investigative review—Proven on high-profile internal investigations, Document Mapper helps you quickly and accurately identify key information in support of legal strategy.

You can also see which concepts are present within a certain timeframe. It's another way to get a layer of data on top of the concept cloud, so you can determine when the activities you're investigating took place.

Copyright ©2019 Nuix. All rights reserved.

PREDICTIVE CODING:

Predictive coding (Figure 4) uses machine learning to analyze a sample set of documents that a human being has already coded—for example as relevant or irrelevant, privileged or not privileged. It then applies the resulting knowledge to very large sets of documents that have not yet been coded.

A recent and exciting advance in predictive coding is Continuous Active Learning (CAL), which makes the process quick and easy for reviewers of any skill level. CAL works by learning as each reviewer codes documents, using those decisions to optimize and train the predictive coding model.

CAL identifies relevant documents earlier in the review process than traditional predictive coding. You can add new documents to the review set as you receive them, rather than having to wait for all the documents before you start. Even with multiple reviewers working simultaneously, CAL incorporates all their decisions to continuously fine-tune the model and identify more and more relevant documents.

Our visual approach to predictive coding allows legal teams to balance the precision and recall of each coding exercise, so they can align their review goals with the budget and importance of the project. It demonstrates that precision and recall are always in opposition: the greater your recall, the lower your precision. This reinforces that trying to review all relevant documents is impossible and prohibitively expensive.

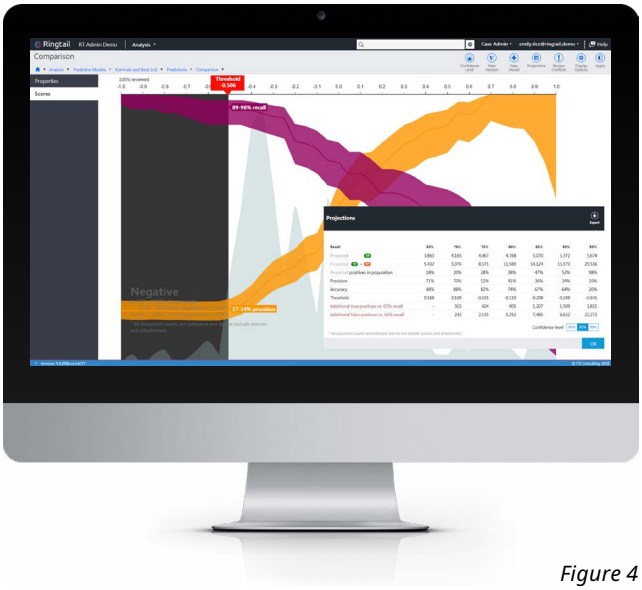

Figure 4

Copyright ©2019 Nuix. All rights reserved.

TEMPLATES WITH CONDITIONAL CODING:

By creating and using coding templates, you can significantly speed up workflow. Conditional coding can support reviewers with smart templates that dynamically change as the documents are coded, which directs reviewers to required fields, and deemphasizes those that are not required. Reviewers can also create one-click coding macros that execute a set of coding criteria across the document.

Where reviewers can become fatigued after several hours, conditional coding can support them by streamlining the process, improving consistency, and increasing scalability. By using analytics to group similar documents together, reviewers can apply coding to all documents with just one click.

Another helpful feature is the use of color and icon cues. These cues can help the reviewer very quickly get a sense of what they're looking at without having to look at the specific document. As a result, it allows reviewers to cull down information to what is most key, while allowing them to determine the most important documents to review.

VISUAL CODING SUPPORTS REVIEWERS IN TWO KEY WAYS:

EASY UNDERSTANDING

Provides visual cues about the document type, making coding easily understood.

COLOR AND ICON CODED CONFIGURATIONS

Icons update and colors change based on tagging, such as "not coded," "relevant," "not relevant," and "privilege relevant." You can further configure your codes to have particular colors to inform you in list view or visual view of the documents.

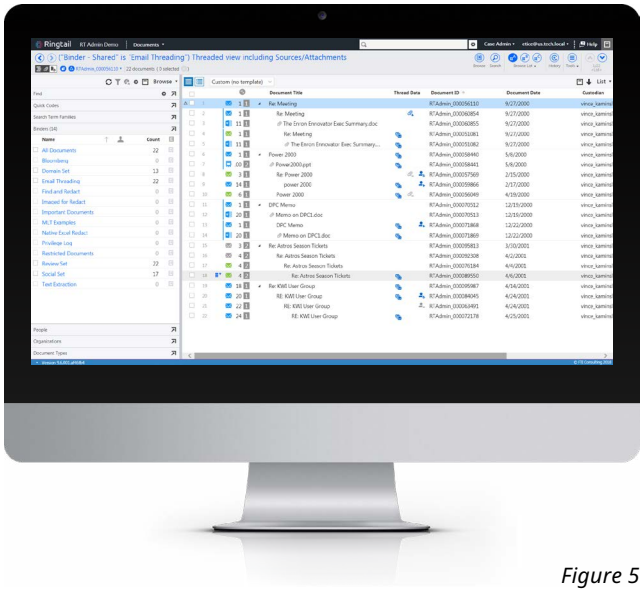

Figure 5

EMAIL THREADING: VISUALIZING EMAIL CHAINS AND ACTING ON THEM:

Email threading (Figure 5) helps reviewers parse email conversations without having to read duplicate passages of text, making the entire process more efficient.

The threading process works by analyzing the body, attachments, recipients, and subject lines of emails to more accurately group emails into threads and to identify the most inclusive, or pivot, emails. The resulting information is presented with visual cues that facilitate review whether the approach is to review only pivots, or to review all data more efficiently in email chain order. Using a threading view, each successive message in a chain is indented further. The thread data column allows reviewers to scan long threads and identify which emails are pivots, where new threads begin, and when attachments or people are added or dropped from the conversation. Threading streamlines a time-consuming activity, allowing legal teams to understand email communications much faster.

Copyright ©2019 Nuix. All rights reserved.

AUTO REDACTION:

In the New Information Age, there are many reasons why you might need to redact portions of a document. However, simply blacking out the text won't fulfill today's needs. Courts or lawyers may want to see reasons for redaction, and some sections of text may be redacted for multiple reasons.

The find and redact feature fulfills these needs. When you search for keywords—or more complex patterns such as credit card numbers or personally identifiable information—the software generates a list of hits, or results to that search, so you can apply redactions accordingly. Once you've searched a term, you'll see the context text around it. From here, you can select individual excerpts you want to redact, then label them with a color.

These colors will allow you to categorize and understand why something was redacted. If there are two reasons why an excerpt should be redacted, you can select it multiple times.

In the New Information Age, you also need a solution for redacting native Microsoft Excel files. With Nuix, processing prepares the Excel spreadsheet for redaction, including the exposure of data and the extraction of comments. You can then select cells, rows, columns, or worksheets for redaction—or use the option to redact everything but a selection.

Copyright ©2019 Nuix. All rights reserved.

CHAPTER 3:
Smart Review for the Everyday

Because of the effort involved in applying analytics with traditional tools, legal teams typically only used analytics for large and complex matters. The reality is—even for your small cases—intuitive and easy-to-use analysis and visualization workflows will significantly decrease the time needed to review documents, find the right information, and inform case strategy.

Smart Review enables industry-best models for you to understand concepts at the assignment and case level, drill into patterns, group similar documents, perform quality review, and take advantage of predictive coding and Continuous Active Learning. As a result, this model of efficient review helps cut the cost of an expensive process.

The average eDiscovery review costs around $18,000 for every 1GB of information, taking up about 70% of the total costs of the discovery process.[5] However, an eDiscovery solution that delivers Total Data Intelligence, such as Nuix, can substantially reduce the volume of irrelevant documents going into downstream litigation—by as much as 75% in some cases—and drastically reduce the costs of legal review.

Spending a bit of time planning what exactly you need to review, and how to review it, makes a world of difference to improve review efficiency, even in the smallest cases.

6 ways Smart Review can improve your review process

1. Delivers higher throughput and greater certainty around project timeframes.
2. Minimizes errors by standardizing settings and providing a clear set of steps to follow at each stage of the process.
3. Makes outputs more defensible by documenting each step taken to achieve the result.
4. Enables firms to share work between teams and offices by synchronizing the processes they follow.
5. Reduces staffing and training costs by empowering analysts and technicians to complete complex eDiscovery tasks.
6. Lightens the load of repetitive grunt work on eDiscovery experts, freeing them to deliver higher value up the chain.

[5] Nicholas M. Pace, Laura Zakaras, *Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery*, RAND Institute for Civil Justice, 2012

Copyright ©2019 Nuix. All rights reserved.

CHAPTER 4:
Smart Review in Action

Now that we've covered what Smart Review is, we can begin to understand how to put it into practice. The concepts of Smart Review, such as visualization and analytics, can be used to speed up eDiscovery, reduce the costs of review, and increase productivity of your legal team. Here are several ways you can leverage Smart Review.

SUPPORT EARLY CASE ASSESSMENTS

Early case assessment helps establish the landscape of data sources and types your legal team will traverse in a pending legal matter or investigation. It aims to facilitate informed decision making during the early stage of an eDiscovery matter. It can also help facilitate investigation of data that represents risks or rewards for the parties involved.[6]

During early case assessment with Nuix, you can use data cubes to identify important documents that meet particular combinations of criteria. This means complex searches become simpler, because you can select the data you want to compare and can then manipulate the data to see which documents meet the criteria. Unlike traditional two-dimensional spreadsheets, you can create and view a cube that has multiple dimensions. Each row or column level is more granular than the one above it. This allows reviewers to drill into hierarchical data elements and focus attention on critical documents during early case assessment. It also allows you to evaluate the efficacy of data collections or test and sample keywords.

THE TOP LITIGATION STRATEGY FOR LEGAL DEPARTMENTS IN THE U.S. INVOLVED WITH THE EDISCOVERY INDUSTRY IS EARLY CASE ASSESSMENT, OR ECA.

76% of insurance carriers say that ECA was their primary litigation cost containment strategy.

98% of professionals using ECA state that it is an effective approach.[7]

IDENTIFY AND PRIORITIZING KEY TERMS

As previously discussed, teams and projects often use code words or jargon that are unfamiliar to outsiders. People who conduct illegal or shady activities use coded language to hide their wrongdoing, or make it less obvious. Analytics can help you understand how well your keywords fit the language used in an organization to highlight terms you'd never considered—and rule out ones you thought were promising but are actually too broad or irrelevant to the case. As a result, you'll discover the best key terms to use to cull documents for review.

The process typically begins with attorney review of a random sample of documents to identify key terms within relevant documents. Then eDiscovery teams use document clustering and text analytics to understand trends and frequency in how terms occur across the document set, to identify additional words that tend to appear in connection with these terms, and to evaluate the effectiveness of specific terms. In doing so, the documents tell us what the keywords should be.

From here, you can set search term families, which are preloaded groups of search terms that can be executed against all documents in the case. With this feature, you can automatically highlight terms in content and in document clusters and it can be used to test search term efficacy.

During this process, you can also use relationships and connections between documents to identify potentially relevant documents that do not contain key terms. This way you gain a powerful way to obtain more effective sets of terms and flagged documents—and eliminate irrelevant materials from your data sets.[8] This saves time and makes the entire review process run more efficiently.

Sample review sets can be used in support of a proportionality argument regarding particular proposed search terms or for seeding in a predictive coding scenario.

[6] Jie Zhang, Garth Landers, *Critical Capabilities for E-Discovery Software*, Gartner, October 2015
[7] LexisNexis, P&C Law Departments Prioritize and Rate Initiatives for Controlling Costs, October 2016
[8] David Houlihan, James Haight, Evaluating the Opportunities Presented by Analytics and Visualization in Legal Review, Blue Hill Research, November 2015

Copyright ©2019 Nuix. All rights reserved.

PRIORITIZE AND ACCELERATE REVIEW

With Smart Review, you can also accelerate the entire review process, with a number of key features. For instance, using analytics, you can replace linear review that's organized by custodian, with topic-based review that's organized according to potential relevance. With document mapping capabilities, you can cluster documents according to timelines, narrative relationships, conceptual similarity, or occurrence of key terms. These clusters help you prioritize those that appear to have the highest likelihood of offering information that is meaningful to the underlying matter.

By performing this analysis at the start of the review process and continually throughout, based on ongoing feedback of reviewers and attorneys, you can keep the highest priority documents at the head of the review queue. This supports in uncovering relevant information early on and accelerates the review process as a whole.

In the case of a review of a received production, once the relevancy of clusters reaches a point of diminishing return, the legal team may opt to discontinue further review, again saving the client on review costs.[9]

QUALITY CONTROL

Analytics capabilities assist in quality control checks of reviewed documents prior to production and can help prevent the inadvertent production of privileged materials. These capabilities play a crucial role in helping quality control reviewers to quickly target potential issues in document sets.

Visually clustering documents coded as responsive *and* protected helps you quickly identify groups of documents that warrant quality control review. This reduces the time required by quality control checks as well as the risks of inadvertent disclosure of protected information.[11]

SMART REVIEW IS A NEW INFORMATION AGE IMPERATIVE

The New Information Age is making it imperative to start applying Smart Review to the eDiscovery process. The days where legal teams could solely rely on linear review methods are behind us. Now, review material can include email documents, as well as data from mobile devices, messaging services, and cloud storage. Your eDiscovery tool must be able to gather and ingest information from a wide range of sources and make that information tell a story.

Nuix offers comprehensive eDiscovery without compromise—one that's fit for the New Information Age. With Nuix, eDiscovery teams can be faster and smarter than ever before and get the right documents to attorneys, the first time.

75% of chief information officers within the US Government say that they are not confident in the quality of their eDiscovery programs.[10]

[9] David Houlihan, James Haight, *Evaluating the Opportunities Presented by Analytics and Visualization in Legal Review*, Blue Hill Research, November 2015
[10] Deloitte, *Ninth Annual Benchmarking Study of Electronic Discovery Practices for Government Agencies*, 2015
[11] David Houlihan, James Haight, *Evaluating the Opportunities Presented by Analytics and Visualization in Legal Review*, Blue Hill Research, November 2015

Copyright ©2019 Nuix. All rights reserved.

In the New Information Age, analytics and visualization tools support Smart Review, making the review process more efficient, effective, and organized than ever before.

An eDiscovery solution for the New Information Age should include the following essential features:

- Document mapper and mines
- Predictive coding
- Templates with conditional coding
- Continuous active learning
- Email chain visualization
- Native Excel redaction
- Social networking view

nuix

Nuix understands the DNA of data at enormous scale. Our software pinpoints the critical information organizations need to anticipate, detect, and act on cybersecurity, risk, and compliance threats. **To learn more visit www.nuix.com**.

NORTH AMERICA

USA: +1 877 470 6849

EMEA

UK: +44 203 934 1600

APAC

Australia: +61 2 8320 9444