

How Modular Workflows Counter Risks and Costs

By Rich Turner, Vice President, Marketing, Content Analyst Company and Bill Johnson, President and CEO, TCDI

21st century businesses run on knowledge management workflows. Whether they have a formal ERP or ECM solution or simply rely on email, today's companies run their businesses largely through electronic documents.

Litigation is also a fact of life in business, and this is where the trouble begins. By and large, legal discovery is patterned after historic paper collection. Traditional mail often contained advertising or personal materials, but these never made it beyond the "circular file"—in essence, they were pre-culled before they ever made it into the workplace. The physical act of collecting paper—typically by paralegals with some knowledge of what they were looking for—provided a further opportunity to skip clearly non-relevant materials. Once everything was collected, it was more a matter of simply processing and reviewing everything you had.

The world of electronic data has a lot more "noise" than was ever present in paper. There is little "pre-culling," lots of repetition and massive volumes of information, stored everywhere by everybody. All of this information is potentially discoverable when litigation occurs.

In much the same way that companies began tackling the challenges of electronic knowledge management a decade ago with workflow solutions, the challenge of managing the risks and rising costs of e-discovery are being met with innovative workflow solutions.

The Growing Challenges of ESI

The first forms of e-discovery workflow for processing electronically stored information (ESI) were simple data flows: collected data was fed into one end of the pipe, and database records came out the other end to be reviewed. Simple filtering techniques such as keyword and date range helped cull down the amount of material needing to be reviewed.

As the amount of ESI has grown, so has the complexity. ESI contains rich metadata—so email threading is now an important way to enhance discovery. ESI can also be extraordinarily complex; comprehensive storage and knowledge management workflows mean that ESI can contain multiple languages, numerous document and data formats and hidden information. Even simple keyword techniques are

coming under intense scrutiny by lawyers familiar with active and possible future litigation along with representative documents, often requiring iterative cycles of filtering and analysis.

Further complicating the "simple data flow" challenge are new advanced technologies. Data integrity can be compounded by the misapplication of these tools, or applying them at the wrong time. For example, aggressive culling of data early in the process can undermine the success of email threading technologies.

In response to the growing issue of e-discovery workflows, companies are finding solutions in scalable, modular processing platforms that use an integrated workflow engine. These platforms support complex decision points based on the data being processed, not simply a linear "one-size-fits-all" data flow. Their modularity readily accommodates voluminous datasets, complex customization requirements, multiple data formats and languages. Modular workflows that can scale are well-suited to organizations with large amounts of data, and they are less disruptive than rigid, fixed solutions.

This integrated workflow engine also presents opportunities to use advanced technologies in deciding the processing strategy for any single piece of data while maintaining an auditable workflow for defensibility. These decisions are no longer binary (to process or not to process), but can be a complex diversion of data to different processing buckets and strategies.

Enter Analytics

Advanced analytics—one of the most significant advances in e-discovery technologies—is well suited to integration into such modular workflow solutions. Analytics provide a variety of tools which can be applied to the e-discovery processing workflow to intelligently reduce the amount of information being presented for review. Technologies such as conceptual categorization, conceptual clustering, near-duplicate identification and concept search can be used in conjunction with more traditional methods of data organization (keyword filtering, date and custodian sorting, email threading) to enhance the processing workflow. The successful application of such

technologies depends on the careful analysis of the types of data being reviewed, a clear understanding of what each technology does and is capable of doing and a determination of the best technologies available to achieve gains in productivity on the part of the reviewers. While the workflow engine can guide the selection and application of a whole suite of advanced tools, there is still no substitute for human reasoning and judgment in the analysis of data. It is the combination of a modular, integrated workflow, advanced tools such as analytics and human input that develops the most efficient and accurate way to manage e-discovery processing and review.

The modular, integrated workflow is also key to what the future will hold. Progressive markets are always pushing the convergence of technologies: analytical software such as sentiment, semantic and statistical analysis—historically divergent—thrive together, providing new relationships in a modular hybrid workflow. As these cutting- and bleeding-edge technologies behind analytics gain widespread use and acceptance, they will become the central technologies behind e-discovery processing.

E-discovery platforms in the form of processing appliances will feature analytics and workflow as the underpinning of advanced techniques to filter and organize data as standard processing "out-of-the-box." The ultimate payback for the technology, however, will occur when enterprisewide data can be organized using the product of analytics during the collection phase of the workflow. This will require the ability to analyze documents "in-place" in the enterprise, whether the data exists in file stores, enterprise content management systems or email archives.

Limiting the data that enters the e-discovery processing workflow in a defensible manner holds the greatest promise for cost control—and the modular workflow described here is the most likely enabler for controlling that data. Just as people learned with ERP solutions—that the human element still meant the difference between success and failure—e-discovery will always require attorneys to remain the "masters of the data." Only they can understand how the data relates to the case and can make the qualitative judgments on data relevance. ■

Rich Turner is vice president of marketing for Content Analyst Company. Bill Johnson is president and CEO of TCDI. Content Analyst Company is a provider of advanced search and document analytics software to e-discovery providers and the public sector; headquartered in Reston, VA, they can be reached at 1-888-349-9442, or info@contentanalyst.com. TCDI serves the litigation technology needs of large corporations and law firms and can be reached at 888-823-2880 or tcdiinfo@tcdi.com.