

Proactive Information Governance

By **Rene Laurens**

Information governance (IG) is how organizations tackle growing data volumes — identifying what's important, what isn't, and what to do with it all.

And it's top of mind. In the Information Governance Initiative's "2014 IGI Annual Report," 73% of practitioners and 81% of providers said they're planning to update their policies and procedures for IG in the next year. See the Report at <http://bit.ly/1oRyc0h>.

A recent white paper from that same group, entitled "The Role of Remediation in Information Governance," underscores the benefits of identifying, sorting, and appropriately preserving or deleting the data your organization may be struggling to keep up with. This is known as data remediation, and it's a logical first step toward information governance: a proactive approach to learning from and handling your business's growing data volumes responsibly. The white paper is available at <http://bit.ly/1pVHWA> (registration required).

Data remediation, at a basic level, brings smarter organization to information. Though deletion of irrelevant data

can and should be a part of the process, it's not just about culling. Data remediation supports migration, preservation, and maintenance initiatives by helping teams find the data that matters.

WHERE E-DISCOVERY FITS IN

As the first step in the EDRM, information governance is closely tied to e-discovery. Proper IG means less junk to collect, process, and review when litigation or investigation arises — and that means significant time and cost savings for legal teams, as well as faster insight into the stories your data is telling. But do these new IG workflows require you to expand your budget for software and training?

Fortunately not. Because e-discovery and IG are so closely tied, many of the tools we use to perform e-discovery can also be used to take inventory of and analyze data for a remediation project. Let's take a peek at what those workflows might look like in the e-discovery software you already know, helping you sort key business data into three actionable buckets:

1. Data we understand well enough to defensibly delete.
2. Data we understand well enough to retain.
3. Data we don't have enough information about to make a decision either way.

To reach sufficient understanding, IG teams should consider not just the content of the data, but also its context. This requires some visibility into the data itself, but that visibility can be achieved by degrees. In some cases,

the location and metadata alone can be enough to confidently make some initial decisions on particular data, such as a terminated employee's personal audio or video files.

The IGI's data remediation white paper does a great job of helping you get into the mindset of creating IG policies and procedures, and asking the right questions to determine the safest way to categorize and delete data for your own organization. Be sure to check that out for more detail, but — for now — let's jump to the workflow.

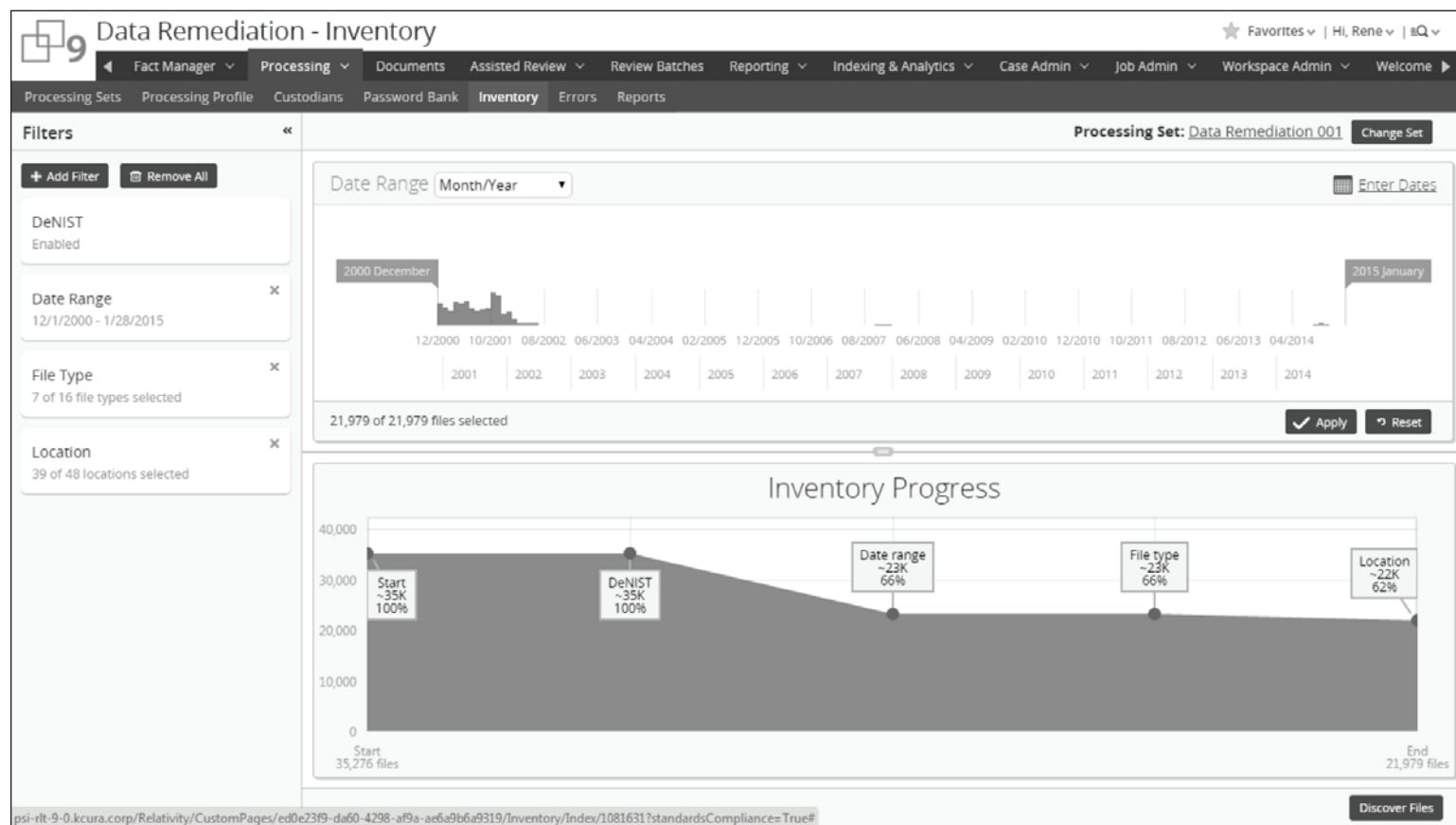
COLLECTION

Performing a targeted collection is a productive first step in a data remediation workflow. Begin by identifying a data source with information you'd like to analyze for a remediation project. Software that supports remote collections means the hardware doesn't need to be onsite for you to access it.

Once you've set up the collection, you can get a full picture of what's on your data source. Examine the full list of files so you can check out high-level metadata and folder structures to quickly make decisions about what needs further analysis, and what can be immediately set aside.

Once you've made those decisions and collected the remaining data, generate a detailed report of which documents were taken and which were left behind. Share this with your IT team to determine what data can be deleted

Rene Laurens is a product specialist at kCura, developers of the e-discovery software Relativity. Rene spends much of his time assisting clients by optimizing workflows. Prior to kCura, Laurens served as a senior litigation support analyst and Relativity administrator in a law firm environment.



first, and let them know what you're still working on.

PROCESSING

Once you've collected your target data, process it for review. The processing step itself is another opportunity to make some initial decisions based on metadata alone.

As you inventory the information, you can perform some simple filtering to further narrow your data. For example, if you have a good date range in mind, or if you know certain file types or e-mail domains can be defensibly deleted, set up some quick filters to exclude those documents from your set. That way, they won't be fully processed, and you'll have less data to analyze and review. See the figure on page 3.

Another detailed report can be pulled following this step; share this one with IT, too.

NEXT STEPS

Pre-collection analytics and inventorying your data should take you a long way in

remediating what you have — they're simple ways to start organizing your data into those three buckets: keep, delete, and unsure. Anything you've excluded so far will fit into the first bucket — it's safe to delete.

The data remaining should, then, fit into either the "retain it" or "we're not sure" buckets. At this stage, you can use your favorite searching strategies to dig through the data and identify any information that's important to retain — or quickly knock out anything that isn't.

Alternatively, text analytics can be a great way to sort this data quickly. You can let your software cluster the data into related groupings automatically, helping you identify any themes. With more advanced workflows like categorization and computer-assisted review, you can tag a small set of data you know is relevant or irrelevant, then amplify those decisions across the rest of the data set.

Customized tags, saved searches, and batching can help sort your remaining data into the three buckets. Once you're satisfied with your work, use reports

once again to share your decisions with IT and move forward with your deletion and migration efforts.

CONCLUSION

By implementing workflows your team is comfortable with, you can turn big data into smart data and act on your evolving information governance policies. Take advantage of technology that helps you stay ahead of the gigabytes of data you're adding to your storage requirements every year, and take more control of the knowledge that data is offering your team. If you want your team to practice before getting started with your own live data, you can try an example data remediation project using the playbook: <http://kcura.com/relativity/igresources>. From that page, you'll also gain access to both of the IGI reports referenced in this article.