Rampiva

Cross-Case Analytics

Detect critical patterns in real-time

Driving Change Whitepaper Series

November 9, 2021

Disclaimer

This document is provided “as-is.” Information and views expressed in this document, may change without notice. You bear the risk of using it. This document does not provide you with any legal rights to any intellectual property in any Rampiva product. You may copy and use this document for your internal reference purposes.

Table of Contents

What is a Driving Change Whitepaper?

Our mission is to make data analytics more accessible. In part, this means taking the lessons we learn from current clients and using them to help new buyers plan their journey. Rampiva's Driving Change Whitepaper Series is designed to help leaders understand how we are thinking about the market, specific use cases, and inventive applications of technology.

Each paper combines our big picture perspective with practical guidance, observed metrics from our benchmarking and links to supporting technical material. We want readers to come away with an understanding of how the Rampiva team thinks, communicates, and works. Hopefully, they will also get insights into the way other teams have solved the problem they are exploring now.

The Current State

The Nuix data processing engine is a great platform – it offers performance and stability across thousands of file, image and container types. It is the backbone of many eDiscovery and digital investigations teams, and increasingly used to identify sensitive, unsecured or out-of-policy data for governance, risk and compliance teams. One of the reasons teams process data in Nuix is to be able to analyze that data for facts and patterns using:

- Workstation: This Nuix Case Interface allows users to group, subset, search, tag and view items as well as wordlists, tag-hit counts and more. This helps identify key facts, test keywords, investigate anomalies and even fix corrupted, encrypted and other problematic data.
- Investigate: This Nuix Case Interface introduces features that help subject-matter experts review data, such as assigns user-specific item sets, enhancing redaction and tagging capabilities, and great graphical representation of item concentrations and patterns.
- Case-Level Reporting: Teams can also report on that data by generating an Excel spreadsheet or exporting metadata and content to a third-party platform.
- Compounding Cases: Nuix Case databases can be joined together to allow users to analyze the data together. This can be helpful when managing large projects or data from different events.
- Elastic Search: Users can write Nuix Cases to ElasticSearch. This makes it easier for teams to track metrics across multiple Cases but requires a special hardware architecture and administration.

Rampiva's Benchmarking Program

To help us understand the impact our clients experience, Rampiva runs a voluntary and anonymized Benchmarking Program. This allows us to track productivity, performance, and trends.

13,480
Cases

355,300
Sessions

346 TB
Loaded

123 TB
Exported

Rampiva Benchmarking Data 2018 - 2021

A New Approach with Rampiva Dashboards

Rampiva Automate aggregates productivity and performance data for all Nuix Cases processed through the platform. This allows users to perform real-time analysis on the health of their environment and the effectiveness of their team.

It also provides the framework and infrastructure for aggregating item-level detail across Nuix Cases.

Simple Approach

For teams that want to aggregate item-level detail, such as custodian names, item date, tag hits, duplication rates, or the concentration of different entities, simply add a step near the end of select Workflows: Scan Nuix Case Statistics. This step will pull these details into Rampiva Scheduler so that users can measure them in a business analytics platform, such as PowerBI. Since Rampiva Automate already provides Client, Matter, and Case-level analytics, teams are able to then compare patterns across Cases for specific Clients and Matters, or across all Cases.

Collecting this information typically adds less than a minute to each Workflow and does not change the architecture or hardware required. And, because the information is stored separately from the Nuix Case, users do not have to check-out Workers, compound Cases, wait for Workstation to open, or wait for other processing jobs or review activity that is happening in that Case to finish. Just be sure to include this step on any recurring jobs that look for manual work in Cases.

It's a simple and easy way to give teams a new level of insight into their Cases.

When a new matter comes up that requires analyzing data from a Custodian or data source, Rampiva clients can quickly check to see whether any of their historical Cases already have some data from that Custodian or data source. If so, the team can open the exact Nuix Case(s) and quickly search and review the material to see if there's any information that would inform the legal strategy for the matter. There may be critical facts or known issues that will impact the cost of satisfying the new discovery obligation.

Advanced Approach

More advanced teams may want to explore cross-case analytics even further – without having to open the underlying Nuix Cases. This can accelerate response time and even enable automatic escalation in the event the detected pattern indicates a significant or urgent finding.

These teams can use Rampiva's External Commands feature, open API, and webhooks to push custom metadata from the Nuix Case into another database (such as SQL). These databases can support much larger datasets, allowing the team to be even more granular about the level of detail collected. While it might not be feasible to provide item-by-item details across all Cases, some simple aggregation can break this bottleneck. Measuring hit counts within item-month for Custodians and other Properties adds enormous flexibility and value without requiring millions of entries for each Case. Similarly, limiting the practice to specific custodians or items from specific data stores, or from specific data ranges, can also help focus the program without sacrificing impact. Risk-scoring and conditional nested logic let teams enrich datasets without overloading individual Workflows or creating friction in the day-to-day processing pipeline.

Sample Properties for Cross-Case Analytics

- Custodian Names
- Company Names
- Country Names
- Languages
- Email Addresses
- Websites
- Networks
- Alphanumeric Identifiers

Why Automation is Important

Cross-case analytics is a great example of how Rampiva Automate empowers Nuix clients to do more. There may be teams that go through the manual process of generating reports on search hits for specific custodians within specific data ranges and then aggregating those reports in a separate system for analysis – but that's time-consuming and error-prone process.

Or, teams might be compounding potentially related Cases to look for patterns and facts. Doing so locks those Cases and prevents other work from taking place, puts more work on the plate of the investigators, and can require significant hardware resources to run searches over very large datasets.

Automating a similar process improves the reliability of the outcome, keeps experts focused on analysis instead of preparation and accelerates the entire process.

Data Processing Quality

Cross-case analytics relies on the quality of the data processing. Teams need to be confident that the pattern they detected (or, the lack thereof) is real, and not the result of a data processing error. In many scenarios, there might not be room for quality control checks or re-running jobs. Adding to this challenge are varying skill levels among the team, pressure to enable non-technical users, late nights and tight schedules. Asking users to do one more step in every single Workflow can be seen as an unnecessary distraction when there is a backlog of essential activity.

Unfortunately, manually processing data is a high-risk activity. Each step can require multiple button-clicks, relies on the user to execute exactly the right order of operations, and has no easy way to detect when a mistake was made. Automating the same process ensures that everything is executed exactly the same way – and if a change needs to be made, that the change is applied consistently and by everyone.

This ensures that the outcome is predictable, reliable and trends towards continuous improvement.

Productivity

Building a central repository of item-level detail adds a couple of steps to every Nuix Session. For any single Session, that is probably not particularly burdensome. But, for every single Session, it can add up. Since it is a multi-step process that includes work outside of Nuix, building this repository manually could add 10 or more minutes to ever Session that gets executed.

So, while teams could build this repository manually, it might be hard to sustain – particularly during peak periods.

Fortunately, this same activity in Rampiva has no meaningful impact on the time it takes to complete a Session. It takes the same amount of time to schedule a Job that will run a Workflow that includes these Operations as it does to schedule a Job that will run a Workflow that does not. Since the amount of data being written to the Rampiva database is small, there's also not a noticeable impact on the time Nuix Workers are checked out to run the Job.

Law Firm Case Study

"Leveraging Rampiva, [we were] able to reduce data processing errors by 30%... and reduce the amount of documentation filled out by the team by 95%."

4.5 The average number of Rampiva Operations that can be run in the time it takes to run one manual Operation.

Productivity in Rampiva compared to manual activity during the first six months

One goal of a Cross-Case Analytics program can be to identify potentially relevant information that has already been processed for another matter.

However, it doesn't really help to identify information early unless you can quickly get it to the relevant team for review. This is another reason why teams that could maintain a central repository of item-level details often don't – the time and effort required to act on that information might disrupt or delay their other work.

Rampiva helps in two ways – first, by increasing the efficiency of Nuix Workers and reducing the time it takes to set up, execute and QC each Operation, Rampiva gets projects out the door faster. Second, by tracking performance across projects and providing the Scheduler interface for managing and prioritizing the job queue, Rampiva helps teams forecast completion times.

This means teams can get information to the right team quickly – and still meet their other deadlines with confidence.

Conditional Logic

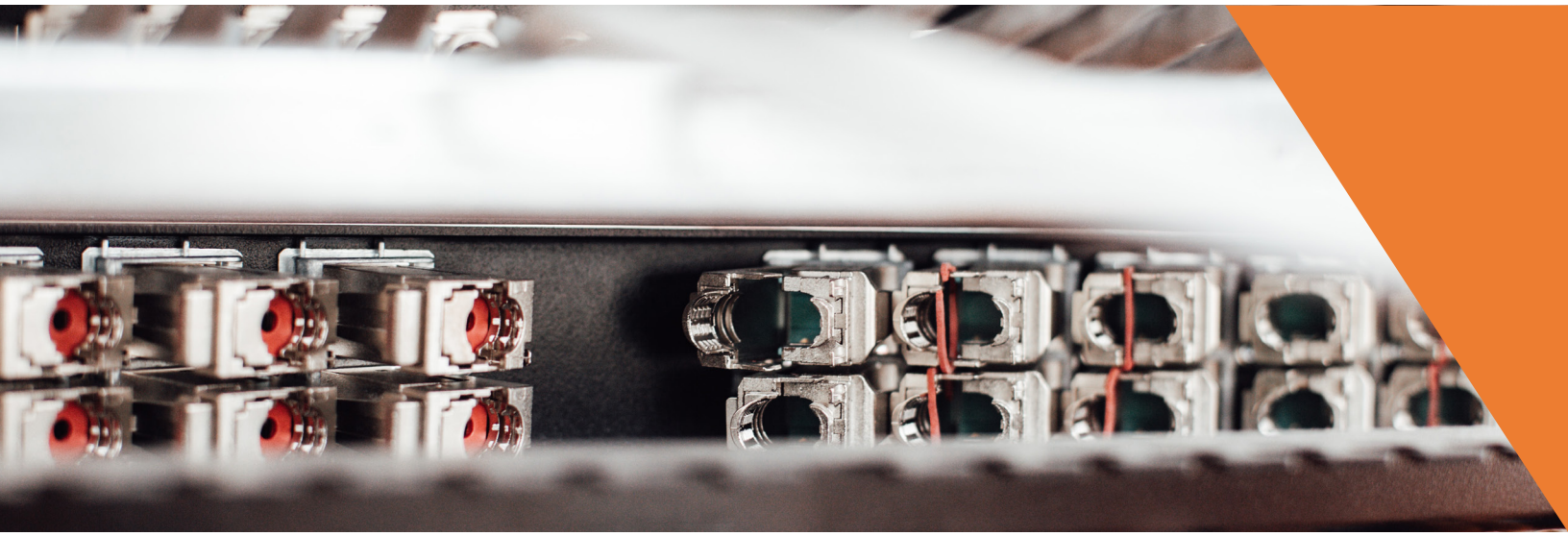
Cross-Case Analytics can be helpful, but one of the big limitations is that the results are rarely conclusive on their own. Instead, they indicate that there is potentially relevant information. Additional activity is needed to determine what the pattern means and if it is relevant to the project.

Rampiva teams can leverage conditional logic to close the gap between the identified pattern and actionable insights. This can include:

- Optimized Tagging: A team that is leveraging cross-case analytics may consider changing their naming conventions or running additional Search & Tag operations specifically to enhance the cross-case analytics.
- Data Enrichment: Adding context to the data already presented in the item can also help. Since this information is often simple metadata, it can sometimes be brought into the Nuix Case or the Rampiva database without overburdening the systems. For example, it may be helpful to maintain a user alias library that combines alphanumeric indicators (such as phone number, email address, employee ID number) along with details about the employee, like title, home office, security clearance, and more.
- Baselines and Thresholds: One of the best ways to tell if a pattern is interesting is to determine if it is meaningfully different from historical or comparable scenarios. Particularly when applied to appropriately enriched data, baselines and thresholds create binary scenarios. For example, emails from when a custodian was in a different role or worked at a different office might not be interesting to a new litigation event. A large number of old, unencrypted Excel files that contain alphanumeric patterns similar to Social Security numbers in a data set that also has a high (as defined in the model) amount of duplication might be a pattern that pointing to the data store being a great target for a data cleanup project.

64.9% **Improvement in Average Time to First Export for teams that adopted Rampiva Automate**

A Client's Journey - 3 Years After Adopting Rampiva



Impact

After understanding how to deliver Cross-Case Analytics with Rampiva Automate's Dashboards and some ways to leverage other features to help optimize the operating environment, it is helpful to explore ways that this program can positively impact operations.

Detection

The first impact is simply the ability to detect patterns that might have otherwise gone unnoticed. Departments that are struggling to deliver projects on time, have limited staff resources, or have a big backlog of projects might not have the time or energy to focus on more than the project immediately in front of them.

By automating a standard, comprehensive and real-time effort to centralize information into a defined Dashboard, teams can focus on simply checking for patterns and acting on them.

Escalation

With more planning, a team can define scenarios in which detected patterns should automatically be escalated to subject matter experts. With Rampiva's Notification feature, specific people or teams can be alerted if an urgent pattern is identified. Supporting reports can be attached along with a hyperlink to the relevant set of items for review in Nuix Investigate.

Accelerated Response

By detecting patterns early and automatically escalating results to subject-matter experts, teams can deal with issues faster. For teams managing litigation, internal complains, security events, or even data recovery for employees, timelines can be tight and speedy insights can help determine strategies and budgets.

Cost Control

Many data processing teams are considered cost-centers and may struggle for budget as a result. Even departments that generate profits or have enthusiastic support from leadership are motivated to minimize marginal costs. Cross-Case Analytics can help do that by giving teams important insights into meaningful data without significantly adding operational delays or new costs.

Rampiva

rampiva.com