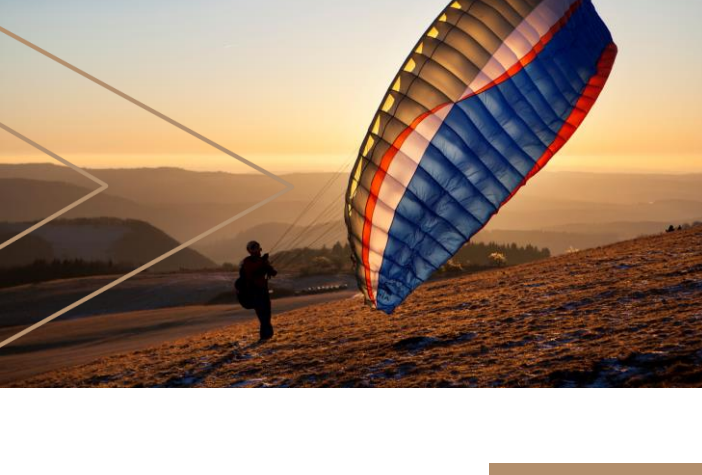


HARBOR

Simply Smart eDiscovery



CASE STUDY

Government Entity Uses Legal Technology to Cull and Review 5TB Collection within Budget

THE PROJECT

Process data from 40+ custodians from a government entity with a limited budget. Provide outside counsel with review database for searching, review, and issue coding. Produce responsive data to the opposing side in the most cost-effective way.

THE CHALLENGE



Reduce the **5TB** that were collected to a manageable size for the review team while keeping costs down.

THE PLAN

The legal team (both in-house and outside counsel) meet for several hours to brainstorm on keywords, names, phrases and all possible terms that may seem relevant to the case. After a couple of weeks of brainstorming, the massive list of 150+ keywords and key phrases is ready.



THE INITIAL RESULTS

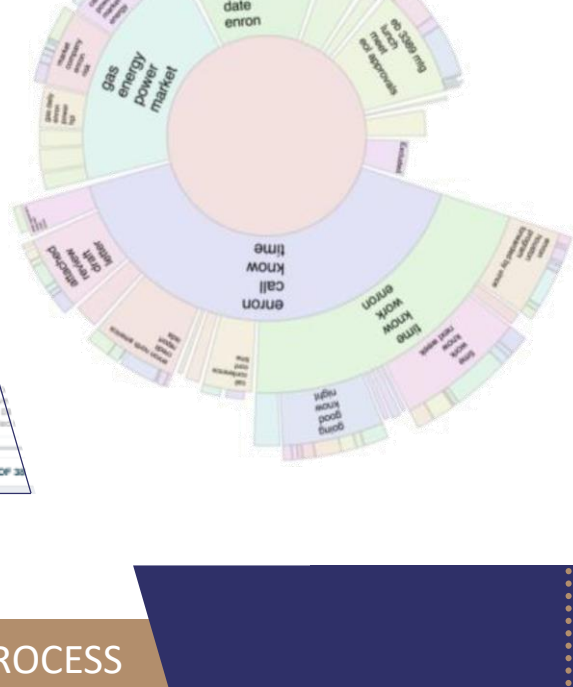
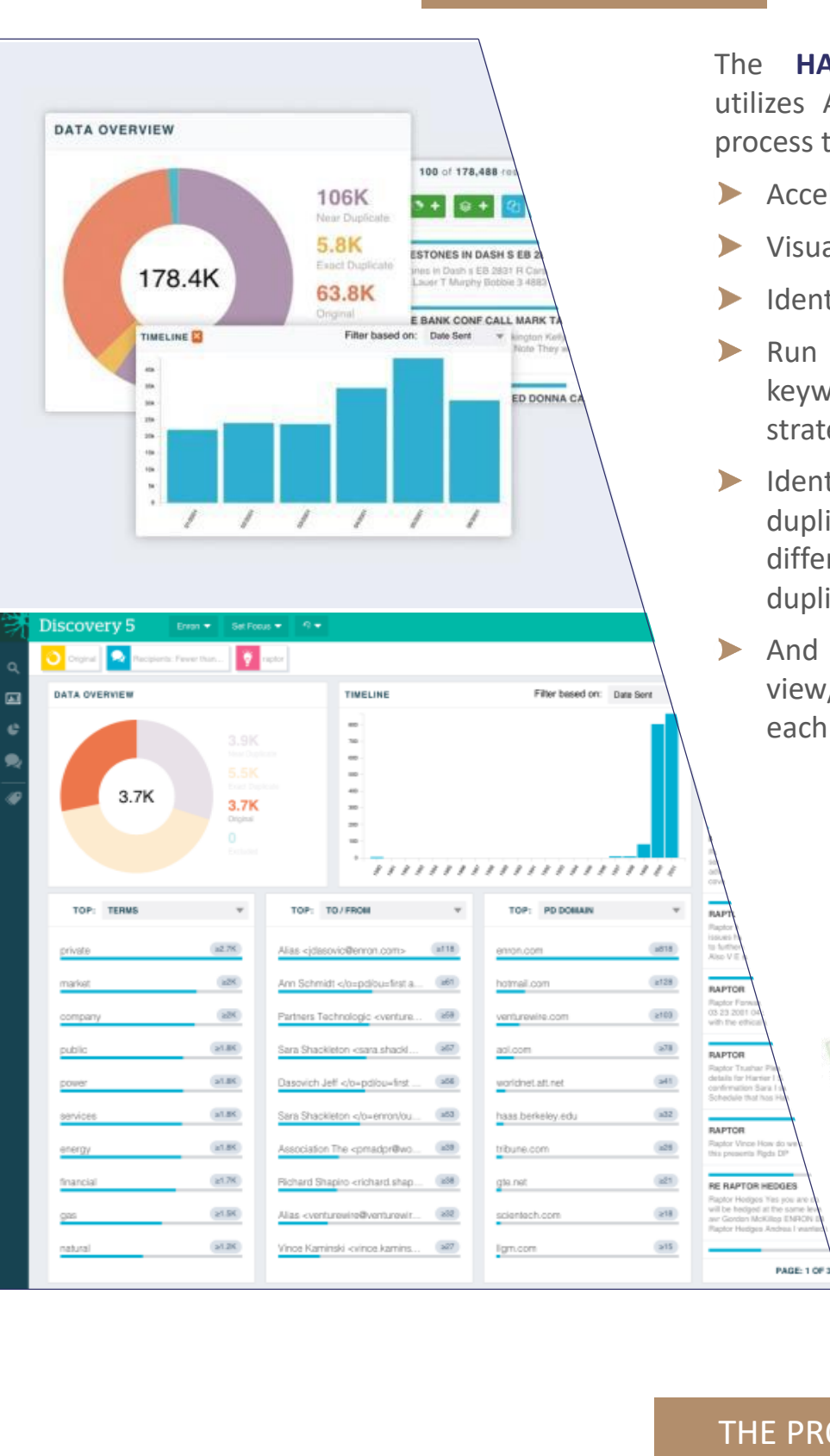
After Processing (DeNISTing, De-Duping, Date Filtering) and applying the keyword search criteria to the 5TB, a total of 750GB of data remained, which fell right at the 15% industry standard of the collected size. The problem was the total number of documents in the set remained overwhelming for the review team.

We suggested they employ our workflow.

THE SOLUTION

The **HARBOR ANALYTICS WORKFLOW** utilizes Artificial Intelligence early in the process to:

- Accelerate case understanding,
- Visualize concepts and key patterns,
- Identify and remove irrelevant data,
- Run and employ key term reports and keyword expansion to assist with strategic decisions,
- Identify exact duplicates and near duplicates based on content (this is different than hash value de-duplication),
- And employ email threading to only view/review the most unique email in each email thread.



THE PROCESS

First, we applied the **ANALYTICS INDEX** to the entire 5TB data set.

Then we prepared an **ECA/EDA WORKSPACE** in Relativity with metadata, extracted text and supplemental OCR for all static images that had no inherent text. The analytics allowed the legal team to perform their keyword investigation in an interactive manner, immediately seeing the results of their searches along with expanded keywords from “conceptually” related documents.

The legal team was also able to **REVIEW “CLUSTERS”** of conceptually-similar documents found in the data set, enabling them to quickly identify groups of documents that were not relevant to the investigation.

Additionally, they found clusters of documents that were critical to the case, and which were quickly batched for priority review and coding. Clusters of relevant, conceptually similar documents were batched together so they could be coded by the same reviewers, improving review speed and accuracy. We call this “smart batching.”

Clusters of relevant documents were also used to identify key terms that had not been included in the initial search criteria.

The **ANALYTICS ENGINE** further allowed the team to reduce the review set by identifying “original” documents and “inclusive” emails.

The analytics technology for identifying original versus “exact” and “near” duplicates goes beyond the typical hash value de-duplication and is based purely on the content – or text – in the documents.

EMAIL THREADING allowed the legal team to only review the “most inclusive” email that contained the entire thread of communications going back and forth.

Typically, these methods reduce review costs dramatically as we’ve all seen how much time can be spent reviewing the “same” or “similar” documents multiple times.

THE ULTIMATE DATA SET FOR PROCESSING AND REVIEW TOTALLED 350GB OF DATA.

Once the Review was finalized and all responsive documents were identified, technology was employed to prepare the documents being turned over to the opposing side.



The **HARBOR PRODUCTION CONFLICTS CHECK** are used to ensure that:

- No privileged or confidential documents are produced
- Full families are kept together and if broken for privilege or confidentiality purposes, place holders are used
- All documents that are identified for redaction have been redacted
- Redacted documents are OCR'd and their original text withheld

THE RESULT

The legal team found a massive volume of documents that were false-positive hits on the initial keyword searches and was able to eliminate them from review – greatly reducing the number of documents, and cost, for review.

By improving on the results of the initial keyword search from 750GB to 350GB of data for review, we were able to reduce the client’s data processing, hosting and review costs – the most expensive pieces – **BY OVER 50%!**

WHAT ABOUT YOU?

HOW ARE YOU MANAGING YOUR EDISCOVERY?

Want to learn more about how you can employ technology on your next matter? Contact us for a demo of the **HARBOR ANALYTICS WORKFLOW**.

info@harborlitigation.com

HARBOR

For organizations looking to improve their discovery approaches, processes and methods... Harbor is the easiest eDiscovery company to work with. Our clients benefit from the right technology, deployed at the optimal stage, for the most cost-effective outcomes. Our reliable, service-centric team simply partners with you to provide smart solutions. **Harbor, Simply Smart eDiscovery.**