

## A Case Of Déjà Review: How DLS Discovery Eliminated 50 Percent Of Their Data Set From Review



*Wilks, Lukoff, and Bracegirdle LLC was working on a lawsuit involving more than 600,000 documents when the U.S. government became involved in the matter. Under a separate issue, the government subsequently renumbered and reproduced more than 300,000 documents also pertaining to the class-action lawsuit. The firm hired DLS Discovery to identify duplicate records among the related matters and saved their team from repeating their first-pass review on the duplicates.*

### DLS Discovery

Headquartered in Wilmington, Delaware, DLS Discovery is a Relativity Premium Hosting Partner that has provided professional litigation support services for more than 15 years. DLS Discovery offers a wide range of services, including on-site data collection, ESI processing, review hosting, records management, trial support, bankruptcy claims administration, document retrieval, and a full courier service.

Andrew McClary, DLS Discovery's ESI and discovery services manager, has worked in litigation support since 2009. Andrew leads global e-discovery operations and oversees project management in DLS Discovery's ESI and Forensics team.

---

### Background

Wilks, Lukoff, and Bracegirdle, LLC handled a class-action lawsuit consisting of more than 600,000 records. Throughout ten months, the records were produced to the firm, and first-pass reviewers worked around the clock to categorize incoming production documents to sort out irrelevant data.

As the case developed, the U.S. Government became involved in the matter under a separate issue and subpoenaed a copy of the productions from the various producing parties. The government renumbered about 300,000 of the 600,000 records with unique identifiers and reproduced them to the law firm under the related matter. No cross-reference was available, and there was insufficient metadata to de-duplicate between the reviewed data sets. Wilks, Lukoff, and Bracegirdle engaged DLS Discovery to help them organize and de-duplicate the data.

DLS Discovery's challenge was not just to identify the duplicates. They also needed to link the first-pass issue coding the firm's team had completed throughout the 10 months prior to the duplicate production to eliminate the expense of reviewing the same documents again.

"Both parties knew the data set well and agreed there was a high number of textually similar documents," said Andrew. "To stay within the projected timeframe and budget, we knew near-duplicate identification in Relativity Analytics would be the most efficient tool for the job."

**“ To stay within the projected timeframe and budget, we knew near-duplicate identification in Relativity Analytics would be the most efficient tool for the job. ”**



## Relativity Analytics in Action

Within two days, DLS loaded all the data—nearly 1 million documents—into Relativity and used Analytics to identify near-duplicates. They set the *minimum similarity percentage* to 99 percent to ensure Analytics would only mark a document as a duplicate if it matched at least 99 percent of the *principal*, or example, document.

The team ran near-duplicate identification overnight and by morning had their results—Relativity identified nearly 200,000 duplicate groups containing about 500,000 documents of the total population.

DLS next determined if any of the issue coding decisions from the firm's first-pass review—performed months ago—could be replicated to the government's production to give the team a better idea of the issues coded on each record.

"If we had five records with the same duplicate group ID, then all five of those records were near duplicates," said Andrew. "So, if one of those records had already been issue coded and the remaining four had not, we replicated that coding decision to those four documents."

They then used random sampling to QC the workflow.

**In the end, DLS eliminated more than 50 percent of the total data set from review** and successfully replicated coding from the first-pass review to approximately 153,000 documents in the government agency's production of duplicates. **In total, this saved the firm an estimated 1,150 hours of review time.**

"We proved the law firm's observation that a large portion of the records in the second production set from the U.S. Government were duplicates of the first production set," said Andrew. "Relativity Analytics eliminated the substantial time and cost associated with reviewing the same records a second time."

**“Relativity Analytics eliminated the substantial time and cost associated with reviewing the same records a second time.”**

Key Project Stats	
Total documents analyzed by near-duplicate detection in Relativity Analytics	900,000
Minimum similarity percentage	99 percent
Total documents identified as duplicates	500,000
Review hours saved	1,150



231 South LaSalle Street, 8th Floor, Chicago, IL 60604  
T: 312.263.1177 • F: 312.263.4351  
sales@kcure.com • www.kcure.com



20 Montchanin Rd, Suite 85 Greenville, DE 19807  
T: 302.888.2060  
edd@dlsdiscovery.net • www.dlsdiscovery.net