



# 1 Million Documents & \$1 Million in Cost Savings: Predictive Coding as a Replacement for Human Review

Technology Assisted Review – Case Study & Cost Analysis

Innovative Discovery, LLC

## Purpose

The purpose of this document is to give a high-level summary of the cost savings recognized in a Managed Review project conducted by Innovative Discovery, LLC (“ID”). A technology assisted review (“TAR”) workflow was utilized in order to assess a potential review universe of 1.3 million documents under an extremely aggressive timeline (one month from data ingestion to the production deadline).

*Due to the utilization of this TAR workflow, the client saved \$1 million when compared to a linear review of 1.3 million documents. The case team was so pleased with the results that they will be incorporating this workflow into all future matters.*

## Analytic Tools

The following analytic tools<sup>1</sup> were used in this specific case study:

**Categorization:** A search and organizational method whereby documents are categorized based on an example document or a set of example documents. This method requires user input in that the user must identify the exemplar document(s) before the categorization can take place.

**Predictive Coding:** A TAR process whereby a universe of documents is categorized based on a “training set” of documents, commonly referred to as the “Seed Set”.

*In this specific case study, a Predictive Coding workflow was utilized as a replacement for human review.*

## Cost Savings & Metrics of Linear Review vs. TAR Review

Review Category	Documents per Hour	Documents Reviewed	Hours to Complete Review	Cost to Complete Review
SME Review	60	6,585	109.8	-
Privilege Review (Including Redactions and Privilege Logging)	25	65,913	2,636.5	-
Cost of Analytics	N/A	N/A	536.27 GB	-
Project Management & TAR Consultation	N/A	N/A	111.3	-
<b>TOTAL - TECHNOLOGY ASSISTED REVIEW</b>			<b>2,746.3</b>	<b>\$254,543.09</b>
Linear Review - First Pass	50	1,077,639	21,552.8	-
Linear Review - Privilege Review (Including Redactions and Privilege Logging)	25	219,461	8,778.4	-
<b>TOTAL - LINEAR REVIEW</b>			<b>N/A</b>	<b>\$1,273,910.93</b>
<b>**SAVINGS USING TECHNOLOGY ASSISTED REVIEW</b>				<b>\$1,019,367.84</b>

<sup>1</sup> The analytic tools detailed in this document are available via kCura’s web-based application, Relativity.



## Summary of TAR Workflow

At the beginning of the TAR review, the total number of documents in the database was **1,297,100**. The total number of documents in the Analytics Review Universe was 1,219,404.

The Subject Matter Experts (“SMEs”) reviewed a total of **6,585** documents across the Seed Set, Control Set, and Quality Control Reviews and achieved a target Recall of 82.26%. After the target Recall was achieved, the total number of documents categorized as Responsive was **521,661**.

- **Potential Production Universe** – The number of documents categorized as Responsive + Families was **858,338**
  - **361,066** – The number of documents categorized as NOT Responsive or Unknown that were NOT family members of a document categorized as Responsive
- **TAR Potentially Privileged Universe** – To identify potentially privileged documents, ID ran a privilege screen against the Potential Production Universe. The result was **219,563** documents.
  - **638,775** – The number of documents categorized as Responsive but did NOT hit on a Privilege Screen Term Hit
- **Promoted to HUMAN REVIEW** – The number of documents remaining in the TAR Potentially Privileged Universe AFTER Culling (date restrictions, custodian restrictions privilege footer analysis, etc.) was **65,913**
  - **153,650** – The number of documents removed from TAR Potentially Privileged Universe due to Culling

Once Privilege Term Screening & Culling were completed, the resulting TAR review universe was:

**65,913**      **TAR – Total Documents requiring Human, Privilege Review** (Including Redaction & Privilege Logging)  
**219,563** (TAR Potentially Privileged Universe)  
– **156,650** (Documents removed from TAR Potentially Privileged Universe due to Culling)

**1,153,491**      **TAR – Total Documents NOT requiring Human Review**  
**361,066** (Documents categorized as NOT Responsive or Unknown that were NOT family members of a document categorized as Responsive)  
+ **638,775** (Documents categorized as Responsive but did NOT hit on a Privilege Screen Term Hit)  
+ **153,650** (documents removed from TAR Potentially Privileged Universe due to Culling)

## Summary of Linear Review Workflow

At the beginning of the linear review, the total number of documents in the database was **1,297,100**.

- **Potentially Privileged Universe** - After screening the database that “hit” on Privilege Terms, the number of documents (+ Families) was **731,046**
  - **566,054** – The number of documents NOT in the Potentially Privileged Universe
- **Promoted to HUMAN REVIEW** – The number of documents remaining in the Potentially Privileged Universe AFTER Culling (date restrictions, custodian restrictions privilege footer analysis, etc.) was **219,461**
  - **511,585** - The number of documents removed from the Potentially Privileged Universe due to Culling

Once Privilege Term Screening & Culling were completed, the resulting linear review universes were as follows:

**1,077,639**      **LINEAR – Total Documents requiring First Pass Review**  
**566,054** (Documents NOT in Potentially Privileged Universe)  
+ **511,585** (Documents removed from Potentially Privileged Universe due to Culling)

**219,461**      **LINEAR – Total Documents requiring Privilege Review** (Including Redaction & Privilege Logging)  
**731,046** (Potentially Privileged Universe)  
– **511,585** (Documents removed from the Potentially Privileged Universe due to Culling)