



Relativity aiR for Privilege: A Step-by-Step Guide to AI that Solves the Privilege Problem

Contents

Introduction: The challenges of traditional privilege review 3

Artificial intelligence for privilege detection 4

How technology-assisted review (TAR) falls short 4

Relativity aiR for Privilege changes the game 5

How aiR for Privilege works 6

Setting up a privilege review project..... 6

aiR for Privilege pipeline 6

Performing annotations 8

Better results, with generative-AI powered context 9

Streamlining privilege log creation..... 11

The benefits of aiR for Privilege 12

Conclusion 14

Introduction: The challenges of traditional privilege review

Privilege review has always been a labor-intensive and nuanced task. But now it's harder than ever. As former U.S. Magistrate Judge Andrew J. Peck explained during the Legalweek 2024 Judges Debate:

“When I got on the bench in 1995, the privilege logs in a typical case [were] two to three pages, maybe 50-100 entries. Now the privilege logs are like little novels, and there may be 10,000 or more entries. That is very expensive and is often useless to the other side in figuring out what is or isn't privileged.”

Identifying privileged documents in a sea of data requires more than spotting attorneys' names or legal jargon. It demands a deep understanding of relationships, roles, and context clues for communications. Even with robust search terms reports and predefined privilege screens, the process is far from straightforward.

In many cases, privilege determinations hinge on factors that are not readily apparent, such as whether an individual is acting in a legal or business capacity or whether a seemingly neutral domain belongs to a legal counsel. Incomplete and outdated privilege screens compound these issues, as organizations often lack comprehensive lists of individuals and entities that might confer privilege.

At the same time, the stakes of privilege review have never been higher. Mistakes can lead to the inadvertent disclosure of privileged information, resulting in legal disputes, reputational damage, and costly clawbacks. Overclassification—flagging more documents than necessary as privileged—may seem like the safer route, but it carries serious downsides. Excessively marking documents as privileged can force repeated re-reviews, driving up costs

and extending discovery timelines. This cycle of unnecessary review slows the litigation process and burdens legal teams with additional work that adds little substantive value to case strategy.

Beyond cost and time, overclassification introduces significant legal risk. Producing overly broad privilege logs may inadvertently expose patterns in communications or highlight documents that, upon deeper inspection, are deemed nonprivileged. This risks broader waiver of privilege, potentially opening the door for opposing parties to challenge privilege claims across entire document categories. The resulting legal battles may jeopardize sensitive material and erode client trust.

These challenges underscore the need for a more advanced, scalable approach—one that combines human expertise with cutting-edge technology. They call for Relativity aiR for Privilege.

Artificial intelligence for privilege detection

Limitations of search terms

Privilege review demands legal expertise plus a nuanced understanding of the relationships, roles, and context in a data set. A persistent challenge in this process stems from the limitations of tools like search terms reports, which, while helpful, often fall short. These tools rely heavily on predefined and incomplete lists of names and terms that can be overly broad, sweeping in irrelevant documents while missing subtler language and indirect references that could indicate privileged content.

The rigid nature of search terms leaves critical gaps—either missing key documents altogether or surfacing them late in the review process, forcing re-review and adding significant cost and delay.

Nuanced communications that reference legal advice without explicitly using flagged terms or that involve attorneys indirectly are easily overlooked. Additionally, documents may include references to legal discussions under ambiguous or business-like phrasing, further complicating accurate identification.

These gaps and inconsistencies slow down review and require retraining and recalibration of reviewers midstream as new information emerges. The result is an inefficient, often frustrating process that strains legal teams and increases the risk of inadvertent disclosures or broader privilege waivers. Addressing these issues requires a solution capable of analyzing the data with the same level of nuance and contextual awareness that experienced reviewers apply.

How technology-assisted review (TAR) falls short

AI can address these complex problems by augmenting human expertise with advanced AI-driven tools designed specifically for the privilege review process. But initial applications of AI fell short of expectations. For example, technology-assisted review (TAR) is very good at identifying responsive documents based on relevance but not at detecting potentially privileged material. The fundamental flaw lies in TAR's heavily weighted focus on the textual content within a document to determine a classification, while overlooking the contextual factors critical to privilege, such as the roles, relationships, and intent of the individuals involved in the communication.

Applying TAR to privilege review is like forcing a square peg into a round hole. TAR's algorithms are

trained to classify documents based on patterns and similarities in content, but privilege determinations frequently hinge on who is communicating, not just what is being said. For example, a draft document sent between two business executives may be considered nonprivileged. The same draft, however, shared with or by in-house counsel as part of legal advice, could confer privilege, despite the content being identical. This distinction escapes TAR, as its focus is limited to surface-level textual analysis rather than the broader context of the communication's participants and their legal or business roles.

Additionally, TAR's performance is heavily influenced by data set richness (number of positive examples in a data set). In responsiveness review, where roughly 30% of documents may be relevant, TAR rapidly

builds a reliable model by identifying patterns across numerous positive examples. In privilege review, however, privileged documents often represent a much smaller fraction—typically around five percent of a data set. This lower richness leads to fewer positive training examples, resulting in less accurate models and inconsistent predictions. The scarcity of privileged documents further compounds the issue, making it difficult for TAR to refine its understanding and deliver consistent results across a matter.

Another inherent limitation of TAR is its tendency to amplify errors through iterative learning. As documents are coded and the system refines its classification model, early misclassifications can skew the entire review process, leading to broader errors downstream. This risk is magnified in privilege review, where even a small mistake—such as misidentifying a privileged email chain—can result in costly disclosure or force large-scale re-review efforts.

Relativity aiR for Privilege changes the game

While past AI solutions may not have met privilege review needs, aiR for Privilege sets a new standard. An early version of the technology was able to identify potentially privileged documents by combining advanced machine learning algorithms with social network graphs. This solution delivered strong results in identifying privilege-conferring individuals and legal advice and content. However, Relativity recognized that the new wave of generative AI technology offered even greater potential. As Nathan Reff, manager of applied science at Relativity, explained:

“We conducted months of testing to understand how to best use AI to tackle complex privilege decisions. Originally, we evaluated if generative AI could accurately make privilege calls on its own, but its capabilities were limited. When we instead tested the use of other machine learning and AI technologies in conjunction with generative AI, the results were fantastic. This combined approach was needed to meet the standards for this high-stakes review.”

Now, aiR for Privilege’s approach is multifaceted, combining GPT models with other large language models, machine learning technology, and social network graphs. This unique combination of AI technologies is far better at understanding the nuance of privilege than generative AI by itself. Together, these technologies reduce the upfront setup time and requirement of client prompting,

yet deliver extremely high recall and precision to support decisions with enhanced reasoning. This confirmation process, called Annotations, provides complete transparency and control to the user so they can see how the AI is interpreting the data, each step of the way.

How aiR for Privilege works

Unlike traditional privilege review methods that rely heavily on manual labor and basic keyword searches, aiR for Privilege integrates AI with a structured workflow that extracts all the context needed to make privilege decisions up front, without users looking at a single document. This differs from legacy approaches that collect some context but still require review of each document, a painful process that often results in material being missed.

aiR for Privilege does this through a sophisticated AI pipeline designed to address the complexities of privilege determinations with precision and efficiency. From initial setup to final privilege log creation, each part of the process integrates advanced technology and user-driven oversight to deliver accurate, defensible results that balance the power of AI with the expertise of project teams.

This structured approach mirrors the process that legal teams typically take during privilege review. Each phase builds on the last, enriching the data set with contextual information that enables the AI to make informed decisions.

The first thing to do is simply set up a project ...

Setting up a privilege review project

When beginning a privilege review project in aiR for Privilege, the first item is aligning the application with the matter or case specifics. Here, users map the relevant metadata fields and decide whether to use aiR for Privilege's Client Brain feature, a knowledge repository that stores information from past projects to create consistency across reviews from the same client. We discuss the Client Brain in more detail in the "Performing Annotations" section below.

During aiR for Privilege's setup phase, users can also input any attorneys or law firms, giving the AI a foundational understanding of privilege-conferring individuals and organizations. Doing so only requires basic details, such as names, email addresses, and the roles these individuals play in the matter. Any pre-existing lists of known attorneys or privilege

screen search terms can be imported easily, saving time and effort. While optional, these inputs accelerate the AI's learning process, allowing it to refine its predictions from the outset.

aiR for Privilege pipeline

Following setup, the workflow transitions into eight structured pipeline steps. Each step helps aiR for Privilege learn valuable information by combining organizational data, human subject-matter expertise, and powerful AI technology through a straightforward and consistent process. The steps build upon each other to organize contextual data and help aiR for Privilege understand what constitutes privilege in the specific matter. Across each step, aiR for Privilege makes predictions about the data while also providing users with the opportunity to confirm or deny them. This confirmation process, called Annotations, provides complete transparency and control to the user so they can see how the AI is interpreting the data, each step of the way.

Order	Actions	Status	Pipeline Step	Description
1		Completed	Prepare Project	Prepare Project organizes data for AI analysis. Upon the successful completion of this step, you will be charged for the number of document units included.
2		Completed	Scrub Disclaimers	Disclaimers often include legal language that can be misinterpreted by our AI as legal advice. The Disclaimer Scrubbing operation identifies all potential disclaimers and gives you the ability to confirm or deny the disclaimers for removal.
3		Completed	Identify Spammers	Email addresses that send spam content generate noise that can wrongfully sway analysis for privilege. The Spammer Identification operation will find all email addresses that appear to be sending spam and give you the ability to confirm or deny the email addresses so they can be ignored for future analysis.
4		Completed	Classify Domains	Understanding what email address domains are owned by law firms, public, or in-house help our AI make more effective decisions about who as a Lawyer and their role in the matter. Our AI makes predictions on these email address domains, but having a human confirm or deny the prediction helps improve accuracy.
5		Completed	Match Equivalent Domains	Identifying equivalent domains helps our AI determine which aliases belong to the same entity. The Equivalent Domains operation matches domains that appear to be equivalent and gives you the ability to confirm or deny these matches.
6	Run	Ready to Run	Validate Attorneys	Identifying attorneys and understanding their role in a matter is essential to determining privilege. The Validate Attorneys operation makes predictions about these entities' roles and the aliases that belong to it and gives you the ability to confirm or correct these predictions.
7		Pending	Confirm Privilege Status	Clients often have context on the matter that doesn't always exist within the documents being analyzed. The Confirm Privilege Status step provides an opportunity for a client to provide final privilege conferring, neutral, or breaking statuses.
8		Pending	Populate Privilege Results	Populate Privilege Results performs final AI analysis using contextual information derived from prior pipeline steps and delivers results on the aiR for Privilege (Priv) fields on the documents analyzed.

Cleansing data and reducing the noise

To begin, aiR for Privilege must appropriately structure the data that will be analyzed. The first pipeline step, **Prepare Project**, indexes and organizes the documents in the matter and retrieves any information in the Client Brain or any of the populated known items entered during the setup phase. This process prepares the data for the rest of the pipeline and analysis.

Next, aiR for Privilege cleanses the data set by identifying and suppressing irrelevant content. It is very important to exclude this material prior to analysis as it can undermine the accuracy of privilege determinations. In particular, aiR for Privilege **Scrubs Disclaimers** and **Identifies Spammers** to avoid skewing results with this extraneous content that could confuse the AI. Disclaimers are filled with boilerplate legal jargon that should be ignored. Spam emails are often sent to numerous individuals across an organization, which can create arbitrary connections between people and impact the accuracy of the social network graphs that the AI uses to make predictions. Suppressing this material removes noise that can distort the system's understanding of key relationships and ensures that the AI is instead focused only on meaningful content.

Mapping domains across communications

aiR for Privilege then learns who is involved in communications. In the next pipeline steps, **Classify Domains** and **Match Equivalent Domains**, aiR for Privilege identifies and categorizes domains, distinguishing between law firms, clients, and third parties, and then also recognizes equivalent domains, such as subsidiaries or alternate domains used by the same individual. This ensures that communications will be correctly associated with a single individual even if that person has multiple email addresses.

Establishing legal roles

Now aiR for Privilege is ready to determine the role of everyone in the data set, including whether they are likely to be lawyers. In step six, **Validate Attorneys**, aiR for Privilege assigns each communicating email address (Alias) to an individual owner (Entity) while also predicting whether each person is in a legal role. This process creates a comprehensive map of participants and their roles within a specific project. Then in step seven, **Confirm Privilege Status**, individuals are classified as privilege conferring, neutral, or privilege breaking based on their involvement in the matter at hand. This allows the AI to understand the status of every person in the matter

prior to reviewing documents. It also helps to differentiate between who may create or break privilege at a very granular level, which is essential to making accurate calls.

Performing annotations

As mentioned above, aiR for Privilege pauses across each pipeline step to allow subject-matter experts to review the AI's findings. Users can confirm or override classifications for attorneys, law firms, and domains, ensuring the contextual data is accurate and aligned with the specifics of the case. By doing so, users can apply their knowledge across the aiR for Privilege workflow, see how the technology is classifying data, and build trust in the process.

aiR for Privilege has the ability to learn from annotations to promote greater efficiency, more consistency, and less setup time for future projects from the same client. This occurs if the user elects to leverage the Client Brain at the time of project setup, confirming that they want aiR for Privilege to store contextual information collected during the annotation process. The information is stored locally and securely, ensuring that no data is shared between clients or used outside the designated environment. Furthermore, data stored within the Client Brain is not used to train the underlying GPT models that power aiR for Privilege.

Done

BigThorium

aiR for Privilege

Annotation (Priv)

Entities (Priv)

Search

CTRL+V

Star

Speaker

KD

Projects (Priv)

Annotation (Priv)

Setup (Priv)

Entity (Priv) Layout

Save

Save and Previous

Save and Next

Save and Back

Cancel

10 of 10

Entity Annotation

Annotator Decision

☐ Legal Role

☐ Not Legal Role

☐ Spammer

Prediction

Legal Role

Final Privilege Status

☐ Privilege Conferring

☐ Privilege Neutral

☐ Privilege Breaking

☐ Unknown

Cluster

897432

Source

Newly Discovered

Aliases

Move

Remove

1 - 4 of 4

10 per page

Name	Sent	Received	Email Signature
☐ eroberts@bigthorium.com	62	102	Emily Roberts Attorney
☐ erob@bigthorium.com	24	2	Emily R SNGKX Law Attorney
☐ emilyr@bigthorium.com	3	1	Emily R SNGKX Law
☐ emily@bigthorium.com	1	0	Emily Roberts

Entity Information

Full Name

Emily Roberts

First Name

Emily

Middle Name

Last Name

Roberts

Maiden Name

Nicknames (Comma Delimited)

Job Category

Job Title

Attorney

LinkedIn URL

Sent

Received

1 - 9 of 9

10 per page

The Client Brain is one component of the [aiR for Privilege Brain](#). The Brain has two levels, each of which store and use information differently:

1. Client Brain:

- Securely stores client-specific annotator decisions that can be applied across workspaces or different matters for the same client. For example, once the system has identified a particular domain as a law firm or individual as having a legal role, it can apply that knowledge to future matters from the same client, reducing the need for repetitive human annotations.
- Does not store privilege statuses of entities (i.e., whether they are privilege conferring, neutral, or privilege breaking) since this can change from matter to matter.

2. Matter Brain:

- Securely stores matter-specific information to be applied, if relevant, to different projects in the same workspace.
- Does store privilege statuses of domains and entities. For example, if a specific lawyer is determined to be opposing counsel and a privilege breaker in a matter, they will be treated the same across projects in the same workspace, as this classification remains the same throughout the matter.

By leveraging layered learning in the Brain, aiR for Privilege reduces the need for repeated manual input, driving efficiency, reducing the efforts to set up projects across the same client, and improving accuracy over time.

Better results, with generative-AI powered context

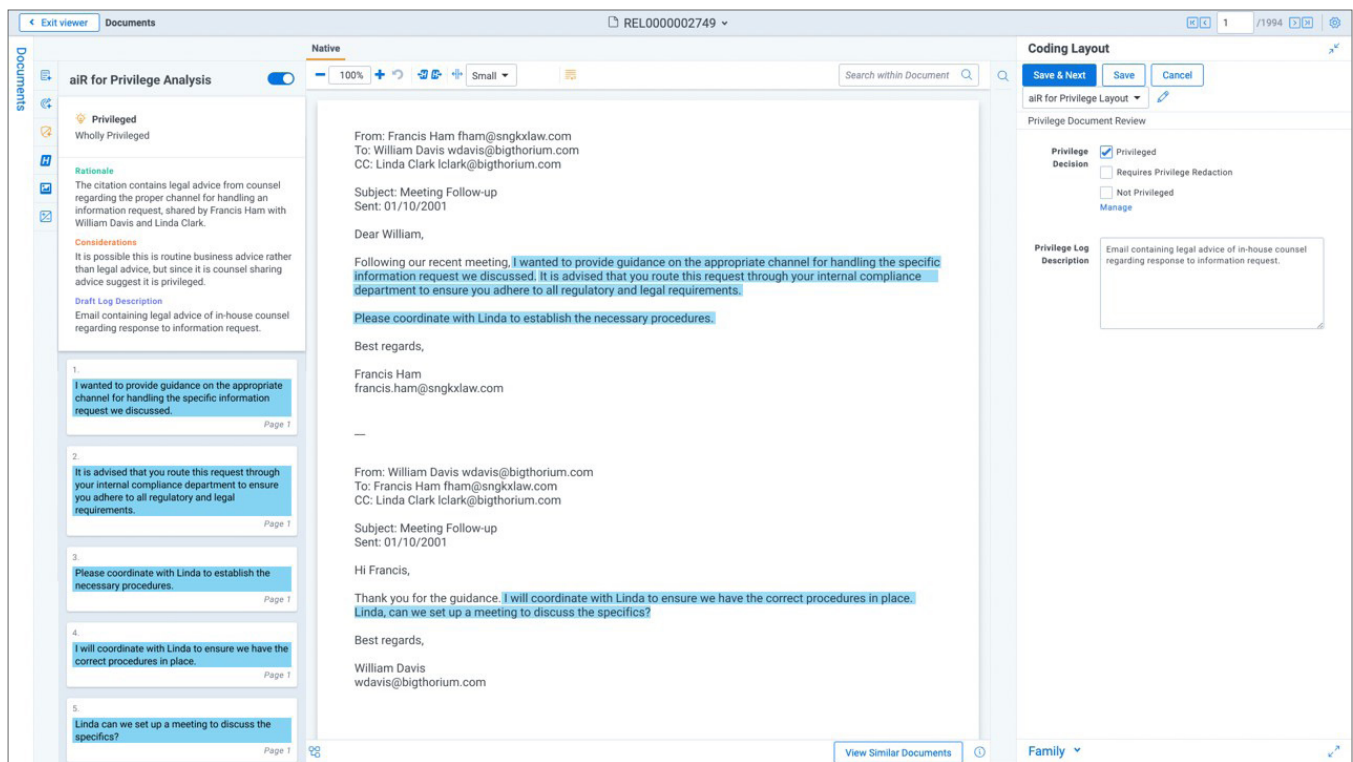
The final step in the pipeline is **Populate Privilege Results**, which incorporates both the decisions made in the prior steps and generative AI to make privilege determinations across the document population. All the information collected throughout the pipeline steps is used to craft unique, document-specific prompts that include context regarding the roles of the individuals and organizations communicating or

mentioned in the document. The context is reflected in 14 different results fields that can be used to gather insight and understand the calls, assisting in quality control review. The prompt is then sent to generative AI large language models (LLMs), providing detailed instructions on how the LLM should analyze every document in the data set. By combining this tailored prompt with generative AI's remarkable ability to analyze nuanced language, aiR for Privilege delivers rapid first-pass privilege calls with unmatched accuracy.

aiR for Privilege delivers detailed rationales, citations, and considerations that clarify why a particular determination was made:

- **Rationale:** A clear, human-readable explanation of why the document was or was not flagged for privilege.
- **Citations:** Snippets of text from the document that support the AI's determination.
- **Considerations:** Counterarguments or areas of uncertainty to guide reviewers in making final decisions. For example, the AI may report that it doesn't understand an aspect of the role, how the language was used, or the project being discussed.

This level of detail eliminates ambiguity, enabling reviewers to focus their efforts where they are most needed and more easily, and definitively, make final decisions on documents.



aiR for Privilege’s results also include defined privilege prediction categories that provide highly valuable insight. Categories such as **Wholly Privileged, Privilege Redaction, Privilege Individual / No Privileged Content, Borderline, etc.**, highlight the nuance around why a document may or may not be privileged. They can also be used to identify groups of documents

that should be reviewed by senior reviewers or require redaction. This information, coupled with natural language reasoning, greatly expedites QC and second-pass review so that legal teams can work more quickly and reduce costs all while making better decisions.

Streamlining privilege log creation

Creating privilege logs is one of the most labor-intensive tasks in document review. David Cohen, practice group leader of the Records & E-Discovery Practice Group at Reed Smith and member of the EDRM Privilege Log Protocol Project, elaborates on why this process has been so painful in the past:

“In many cases, log entries must be individually tailored to accurately reflect the basis for withholding each privileged document. While automation (including metadata) can help, legal teams often have to draft detailed descriptions manually, a burdensome process that can take multiple weeks and hundreds of hours of valuable lawyer time in large cases. This process is also prone to some inconsistency, as different reviewers may have different writing styles and describe similar documents in different ways.”

aiR for Privilege greatly accelerates the preparation of privilege logs by generating draft privilege log descriptions for all flagged documents. These descriptions are crafted in natural language, mirroring the style and structure of those written by human reviewers. Unlike traditional descriptions built by concatenating fields with set topics (e.g., “contract discussion”), the AI-generated descriptions are nuanced and defensible, reducing the likelihood of challenges. Senior attorneys can review and refine these drafts as needed, ensuring that the logs are accurate. The combination of automation and manual oversight saves time while ensuring that privilege logs meet the highest standards of quality and reliability.

aiR for Privilege was remarkably good at developing privilege log descriptions. We appreciated the ability to edit them, but 50% of the descriptions required no modification at all. The product was particularly good at finding the right balance of detail. It made it clear why the document needed to be withheld without sharing too much information.

MARTHA K. LOUKS

Director of Technology Services



The benefits of aiR for Privilege

aiR for Privilege marks the transformation of privilege review. Automating first-pass privilege review decisions and generating detailed, defensible privilege log descriptions reduces the time, cost, and risk of traditional privilege workflows. Legal teams can focus on high-value tasks, such as quality control and strategic decision-making, rather than getting bogged down in repetitive, manual processes, all while maintaining the peace of mind that only the right documents are going out the door.

Here are just some of the benefits teams experience by revolutionizing their privilege review processes:

Unparalleled, AI-driven accuracy

aiR for Privilege is designed with an unwavering focus on recall—ensuring that no privileged documents are missed. Recall rates consistently exceed 99%, significantly outperforming traditional processes in identifying privilege while reducing the risk of errors. This focus on recall ensures that teams can prioritize the flagged documents without fear of overlooking critical information.

Months of testing on real matter data informed the development of aiR for Privilege to ensure it delivered the accuracy needed for high-stakes privilege review. In collaboration with customers, Relativity's applied science and product teams tested and iterated upon versions of the prompt that is sent to the generative

AI model. Rigorous evaluation and improvements led to the final prompt structure that now helps the AI make the right privilege call for every unique case.

By combining the advanced AI technologies used throughout the pipeline steps with powerful generative AI LLMs, aiR for Privilege can effectively identify legal entities in a corpus of data, classify their roles throughout a matter, and understand how to make nuanced privilege decisions with more accuracy and consistency than human reviewers.

Smarter work, faster results, more savings

Privilege review is often the most expensive phase of document review, requiring intensive manual effort from senior lawyers. aiR for Privilege's rapid analyses and high precision reduce the burden on reviewers by quickly getting to accurate results while minimizing false positives. This streamlines both the initial privilege assessments as well as quality control and second-pass reviews, which are often the most time-consuming and costly stages of document review.

Further, privilege reviews become faster with each new matter at the same client, thanks to the system's Client Brain. Building institutional knowledge about a client over time allows review teams to leverage insights from past projects, reducing the need for redundant setup and increasing efficiency.

aiR for Privilege offers a cost-effective and powerful way to protect privilege, providing accurate and efficient results. Considering the significant cost and time required to conduct privilege reviews, aiR for Privilege is poised to significantly outperform linear processes. Its ability to recognize more complex privilege concepts while providing rationales and considerations around its decisions has been crucial for clients' adopting this technology. aiR for Privilege quickly explains its privilege determinations on documents, speeding up the review and QC process and bolstering confidence that it's getting calls right. The solution is an invaluable tool that we encourage all review clients to consider.

CODY HOLIFIELD
Director of Managed Review

cimplifi

By automating repetitive tasks and leveraging past learnings, organizations can cut privilege review time by over 50%. For example, in one law firm's test of aiR for Privilege, the team compared aiR's results against human coding of 4,300 responsive documents. aiR for Privilege took only five hours total to set up, run, and validate, compared to 130 hours needed for manual review, with similarly impressive accuracy.

aiR for Privilege is way more effective than manual review. It only took eight hours to set up the project and get through 4,000 documents—and that includes reading through documentation since this was the first time we used the tool. This will cut privilege review time by more than 50%.

ERIC CHAPMAN

Senior Discovery Attorney



right documents are released and that they are delivering the best possible privilege review for their organization and clients.

By providing detailed rationales and evidence-based citations, aiR for Privilege empowers reviewers to make informed privilege determinations with certainty. This confidence translates into faster decision-making and stronger defensibility of privilege calls. For example, during the domain analysis phase, users can review and confirm or override the AI's classification of domains as law firms, belonging to a client, or third parties. Similarly, predictions related to privilege-conferring roles can be validated and refined by subject matter experts. This interactive approach ensures the defensibility of privilege determinations and puts control over final decisions in users' hands—establishing confidence that only the right information is going out the door.

Reduced organization and client risk

Ensuring sensitive information doesn't slip through the cracks of a privilege review is critical for protecting client confidentiality and maintaining trust. aiR for Privilege excels at identifying even the hardest-to-find privileged content and provides detailed rationales for every decision. This level of transparency gives reviewers peace of mind, knowing that only the

Fortune 100 Telecom Company Cuts Review Time by 80% with Relativity aiR for Privilege

A fortune 100 telecom company's lean e-discovery team was looking to more efficiently handle vast amounts of data in their complex litigations.

The team decided to use aiR for Privilege on a challenging case that consisted of millions of documents and was expected to last quite a while. They started by using the solution on 93,000 documents, knowing that they could leverage aiR for Privilege's learnings to drive even greater efficiencies in future productions related to the matter.

In just one week, a single e-discovery attorney used aiR for Privilege to review the entire document set, delivering impeccable accuracy and catching over 5,000 privileged documents that had been missed by human reviewers. The solution delivered 99% recall, conducted the review 80% faster than their previous manual processes, and saved \$35,000 on the single project.

“aiR for Privilege found over 5,000 privilege documents that contract reviewers had missed and highlighted exactly why they should be protected. It reduces our risk while driving massive savings. You can't beat that.”

aiR for Privilege Customer

Head of eDiscovery, Fortune 100 Corporation

Conclusion

Privilege review is one of the most critical—and complex—aspects of legal work, demanding a balance of accuracy, efficiency, and defensibility. Relativity aiR for Privilege meets these demands by leveraging the latest advancements in AI and contextual analysis. With its structured pipeline, transparent decision-making, and powerful automation features, aiR for Privilege empowers legal teams to conduct privilege reviews with unmatched speed and confidence.

As organizations continue to face growing volumes of electronic data and tighter production deadlines, tools like aiR for Privilege represent the future of legal technology. By combining human expertise with AI-driven insights, Relativity helps legal teams protect their most sensitive information while conserving time and resources.

**Take the first step toward faster,
more accurate privilege reviews**

Sign up today to see aiR
for Privilege in action.

[Request a demo](#)



231 South LaSalle Street | 20th Floor
Chicago, Illinois 60604
+1 (312) 263-1177 | relativity.com