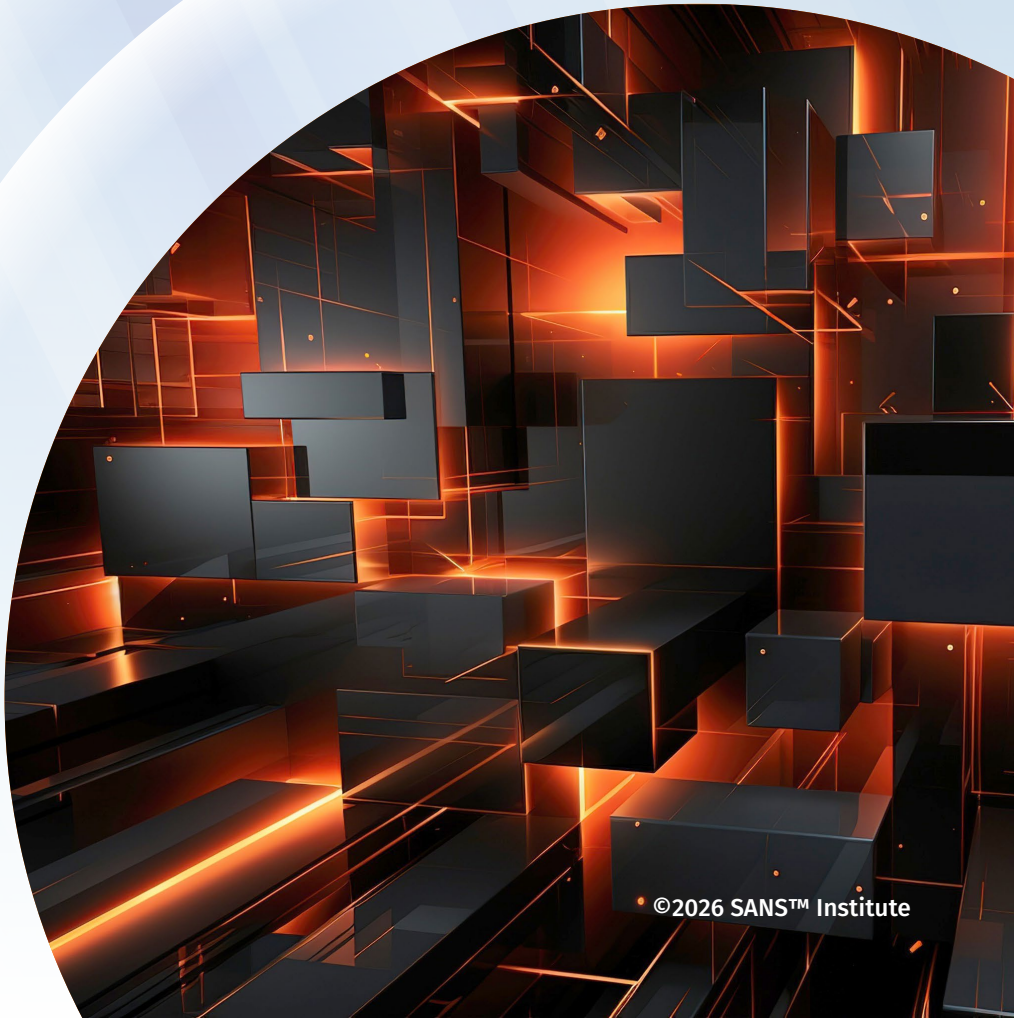
SANS | Research Program

PRODUCT REVIEW

The Next Generation of Data Security

Written by **Dave Shackleford**

April 2026



Lightbeam

Introduction: Data Security Challenges Today

As enterprises accelerate cloud adoption and SaaS usage, sensitive data sprawl has quietly become one of the most consequential security challenges facing organizations today. Critical data (e.g., PII, PHI, financial records, intellectual property) no longer lives neatly inside a handful of databases or well-defined applications. Instead, it proliferates across cloud object storage, analytics platforms, SaaS tools, collaboration environments, backups, and unmanaged repositories, often without the awareness of security teams. This "shadow data" emerges organically as teams move fast, replicate datasets, and grant access for convenience, leaving organizations exposed to risks they cannot see, measure, or govern using traditional perimeter or application-centric controls.

This challenge has driven the rise of data security posture management (DSPM) as a foundational control layer for modern data security programs. DSPM focuses on continuously discovering, classifying, and contextualizing sensitive data across cloud and SaaS environments, answering questions that security teams increasingly struggle with:

- Where does our sensitive data actually live?
- What type of data is it?
- Who or what can access it?
- Whose data is it?

Unlike legacy data loss prevention (DLP) approaches that rely on predefined paths or inline inspection, DSPM platforms analyze data at rest and in use, surfacing previously unknown data stores and highlighting risks introduced through misclassification, duplication, or excessive exposure. This visibility is essential not only for security teams but also for privacy and compliance stakeholders tasked with meeting regulatory obligations across dynamic, multicloud environments.

However, data risk is rarely driven by data alone—it is driven by identity and access. Overprivileged users, service accounts, APIs, AI agents, and automated workloads often have far broader access to sensitive data than intended, turning routine identity misconfigurations into breach-scale events. Therefore, effective data security requires tight integration between data classification and access governance, mapping sensitive data to the identities, roles, and privileges that can interact with it. This identity-aware view is especially critical during incident detection and response. When a credential is compromised or suspicious behavior is detected, security teams must quickly understand what data that identity could access, what was touched, and what regulatory or business impact may follow.

Many organizations today are looking for platforms that unify DSPM with identity context and access governance to enable faster triage, more precise containment, and more confident response, transforming data security from a reactive compliance exercise into an operational security capability. SANS had the opportunity to review the Lightbeam platform, which offers a robust set of identity-centric capabilities addressing data security, access governance, privacy compliance, AI security, and cybersecurity incident detection and response. The Lightbeam team provisioned a test account for us to log in and review a variety of capabilities with a broad set of data sources and environments provisioned. Lightbeam goes beyond visibility-only DSPM by linking sensitive data findings to identities and effective access paths, then reducing risk through direct remediation actions such as access revocation, redaction, account suspension, and policy-driven controls from the same workflow.

Data Security Posture Management (DSPM)

We first reviewed the main dashboard within the platform. (There are also specific dashboards for privacy and other targeted capabilities in the platform, which we cover later.) This customizable interface has a broad array of information readily presented to analysts, including data sources covered and risks detected in them, live data discovery, the number of different identities detected across the covered environments, as well as policy status and alerts (see Figure 1).

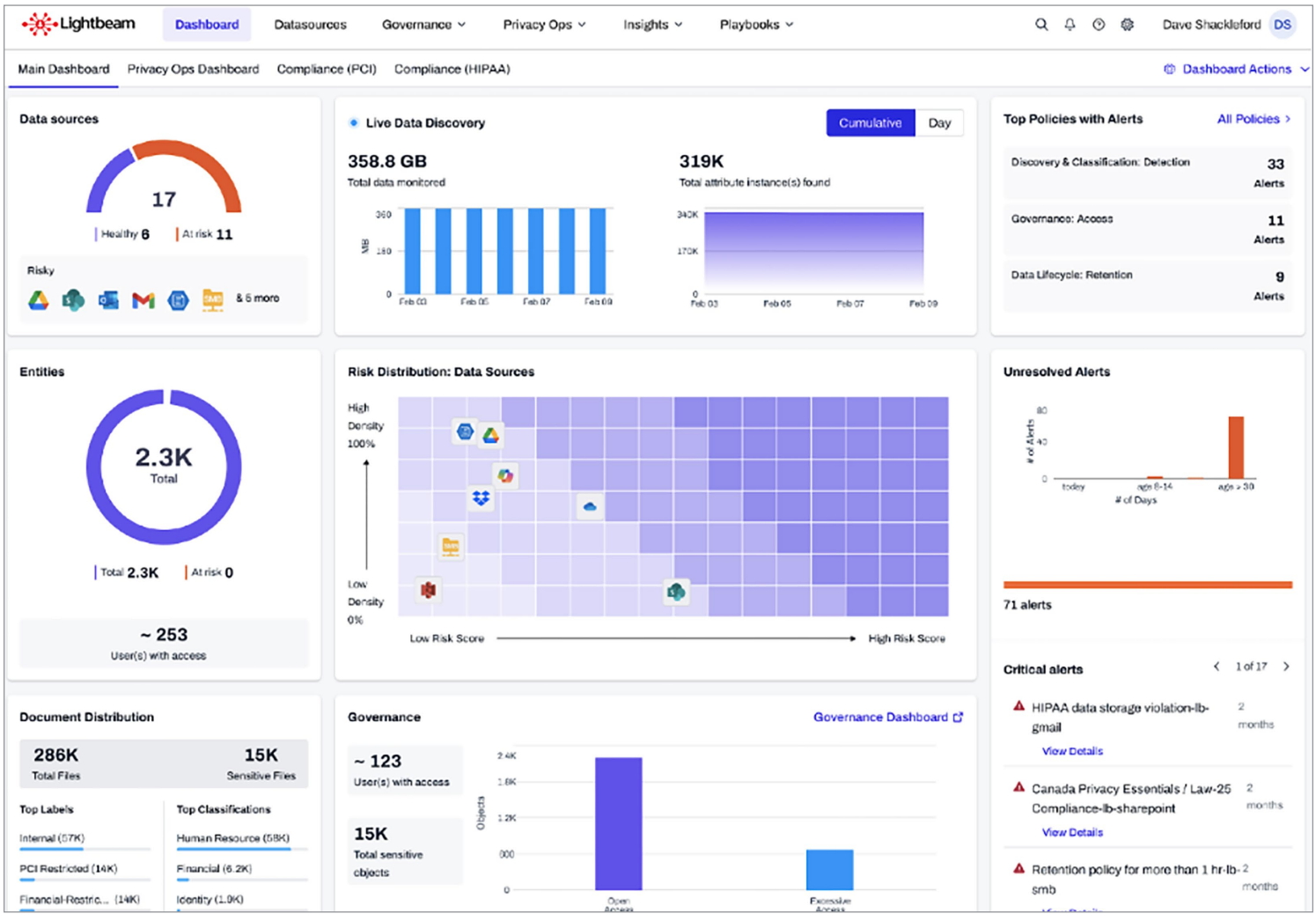


Figure 1. Lightbeam Main Dashboard

We then browsed to the primary "Datasources" tab in the console. A wide variety of data repositories and services were provisioned, ranging from SaaS storage like OneDrive and Dropbox to storage nodes in IaaS services like Azure Blob, AWS S3, and others. Generic SMB servers and data sources like email and collaboration tools (e.g., Outlook, Slack, SharePoint, Jira) also were plugged in, as shown in Figure 2.

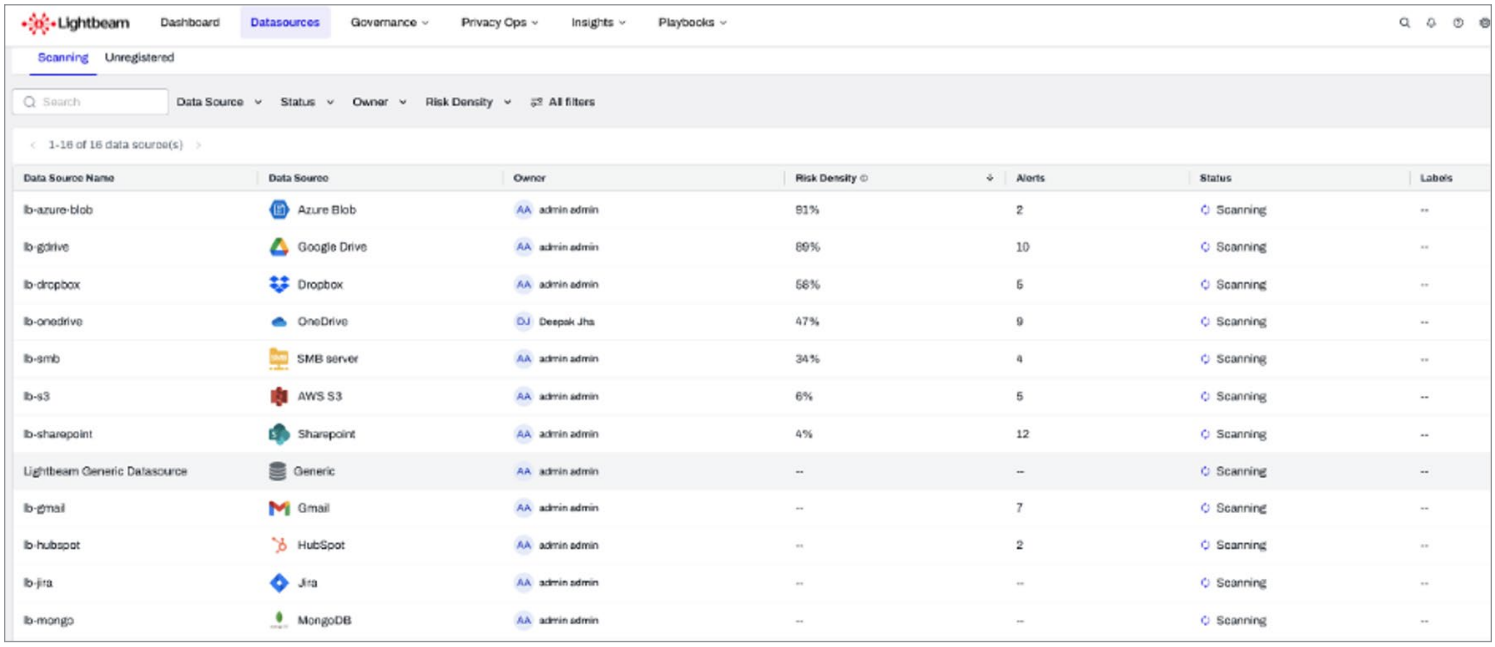


Figure 2. Lightbeam Connected Data Sources

All these services had a status of Scanning to indicate they were being monitored for sensitive data like PII and PHI. When sensitive data was discovered, Lightbeam presented a metric in the dashboard called Risk Density, which shows the percentage of total content in the data source that is considered sensitive. An Azure Blob node in our test environment had a high Risk Density score of 91%, so we clicked in to access a new data source dashboard that breaks down how many sensitive files were noted, the kinds of data detected, and any risk and alerting information about the sensitive data and files (see Figure 3).

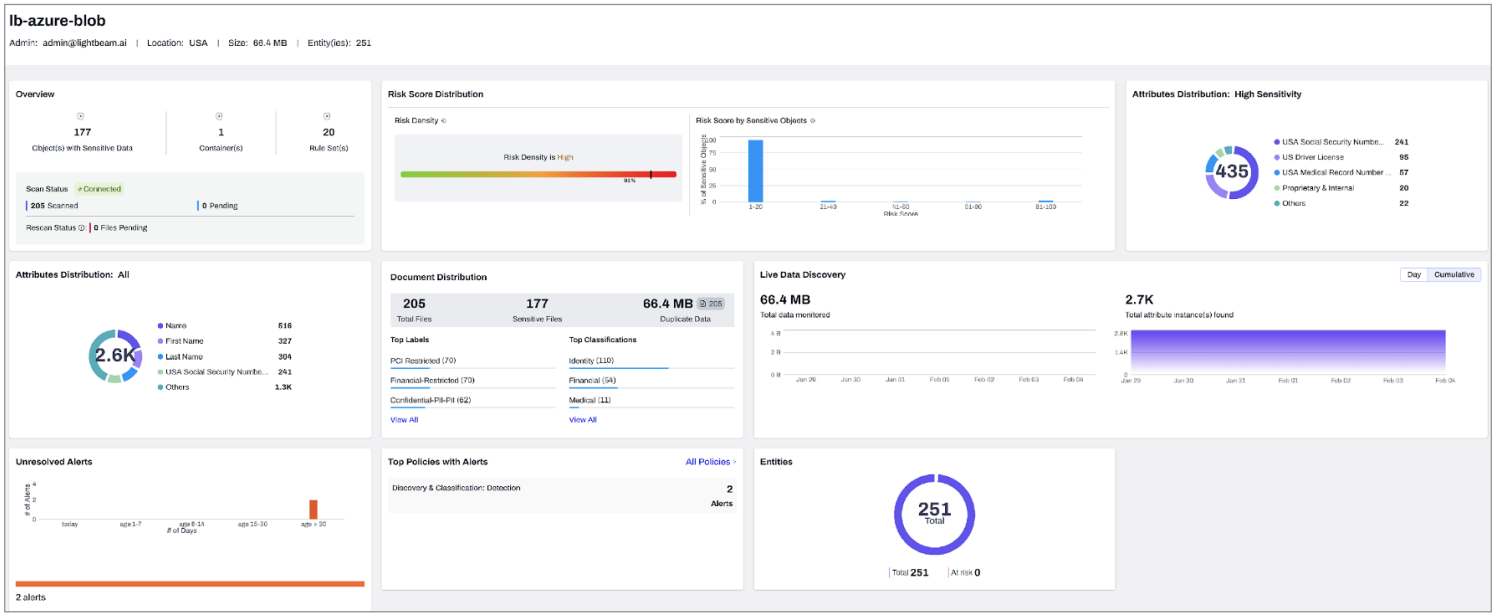


Figure 3. Risk Density Dashboard for Sensitive Data

Being able to scan and discover sensitive data in data sources is a huge part of any DSPM solution, and differentiation of data types (such as PII) can help pinpoint compliance and exposure risks quickly. Lightbeam also can correlate identified data across various types of data sources. We dug into some of the Lightbeam Insights profiles driven by AI, which can identify and then aggregate data attributes to categories. The Lightbeam Attribute Management Insight dashboard is shown in Figure 4, with a small subset of the hundreds of unique data patterns Lightbeam can discover.

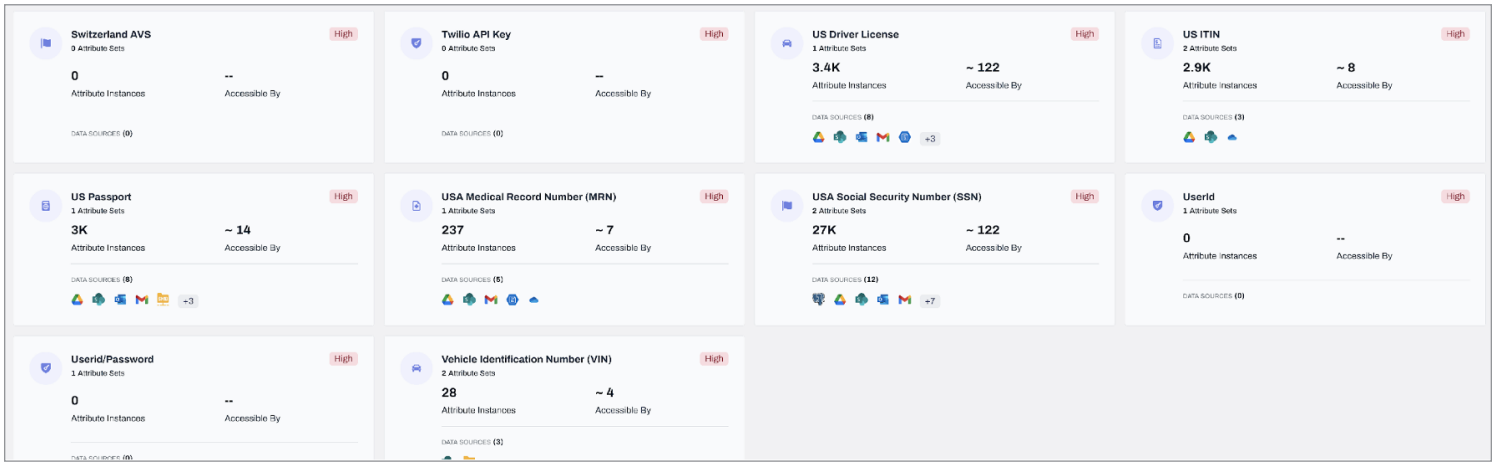


Figure 4. Attribute Insight Categories

When we selected a category (in this case, USA Social Security Number), we were brought to a new dashboard with all Data Sources where this data type was detected. This allowed us to quickly evaluate how much of this data type was present in each data source, as well as how accessible the data was (see Figure 5).

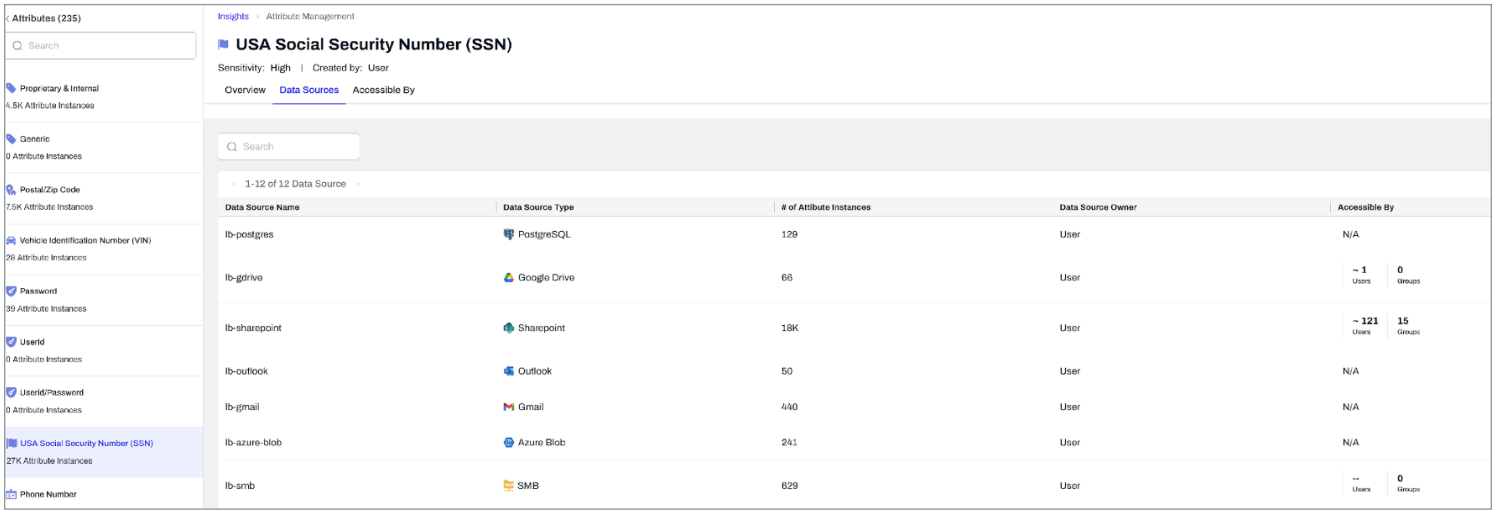


Figure 5. Attribute Data Sources Dashboard

We selected one of the data repositories (OneDrive) and found a sensitive PNG file with masked salary and personal information. Lightbeam can access the full content of files and render them in active memory (nothing is saved), allowing analysts to determine whether the content is, in fact, sensitive and what other attributes and information in the file may be relevant (see Figure 6).

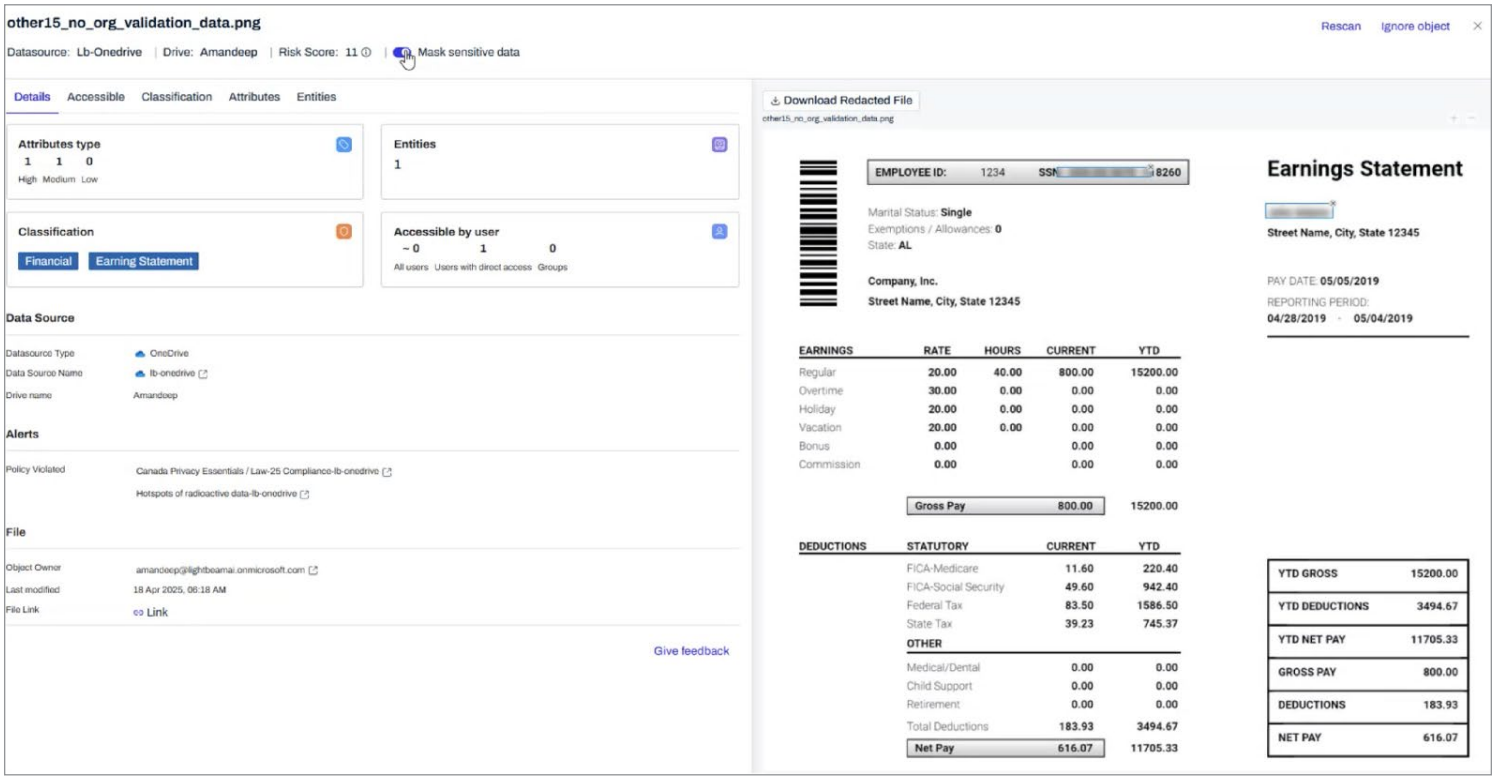


Figure 6. Sensitive File Rendering in Lightbeam

Through this lens, we can see who has access to the file, where it lives, its classification, what entity or identity the data is associated with, and more. Lightbeam AI will automatically apply labels and connect this data to the specific employee the earnings statement belongs to, as shown in Figure 7. (Although we tracked this file through a Social Security Number [SSN], additional personal and financial data labels were applied automatically).

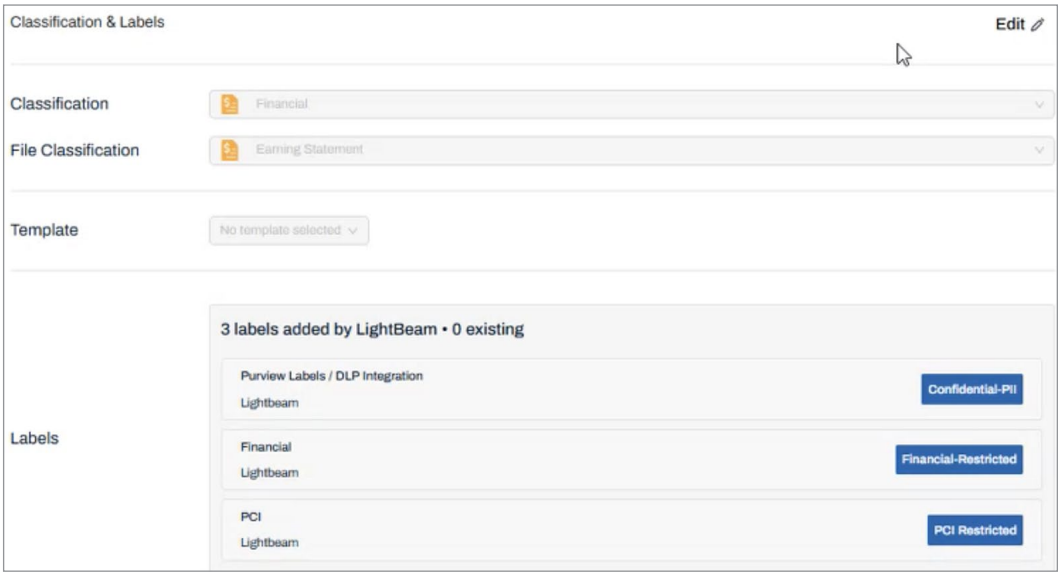


Figure 7. Lightbeam Labeling and Classification

To help train AI and machine learning models in Lightbeam, you also can easily create new templates with just a few clicks to designate new data patterns and classifications, as shown in Figure 8.

In these records, you also can note any discovered entities. In this SSN discovery, we noted a human entity for John Adams, who we were then able to evaluate in terms of user and group access to the entity data, other attributes associated with the entity (just the name in this case), and other data sources where this entity was appearing. See Figure 9.

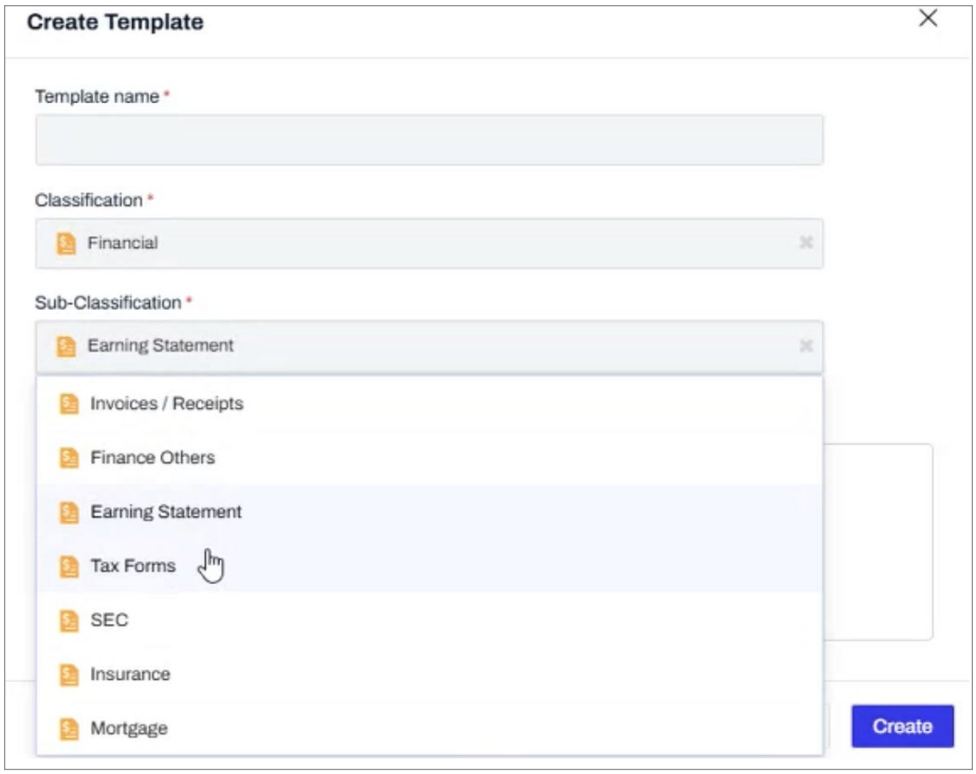


Figure 8. Creating New Classification Templates

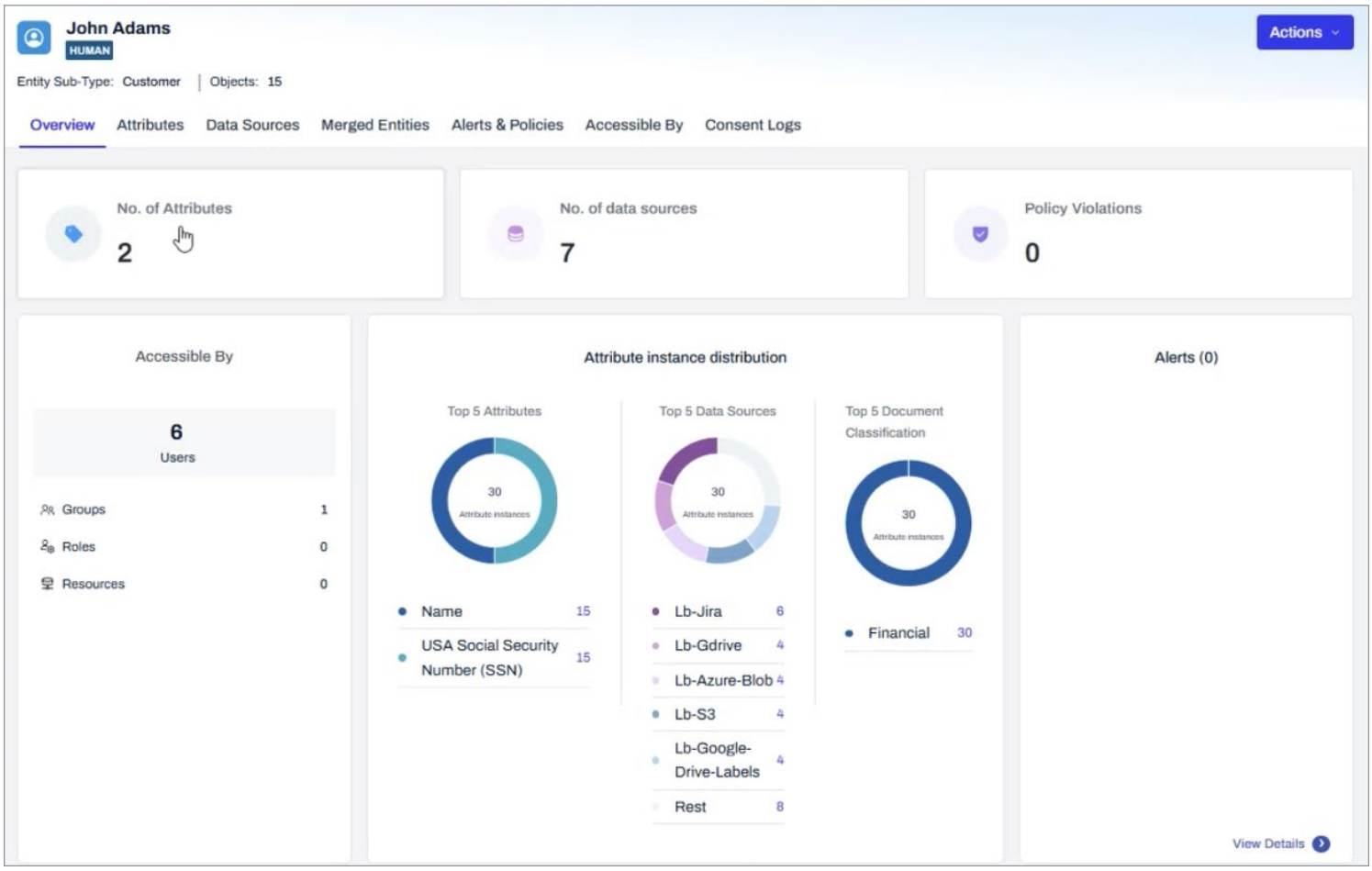


Figure 9. Entity Detail and Data Source Discovery

We then looked into the "Merged Entities" category to find other locations where the John Adams entity was detected, along with any details about files and data sources, as well as additional attributes that we could potentially search on or validate (see Figure 10).

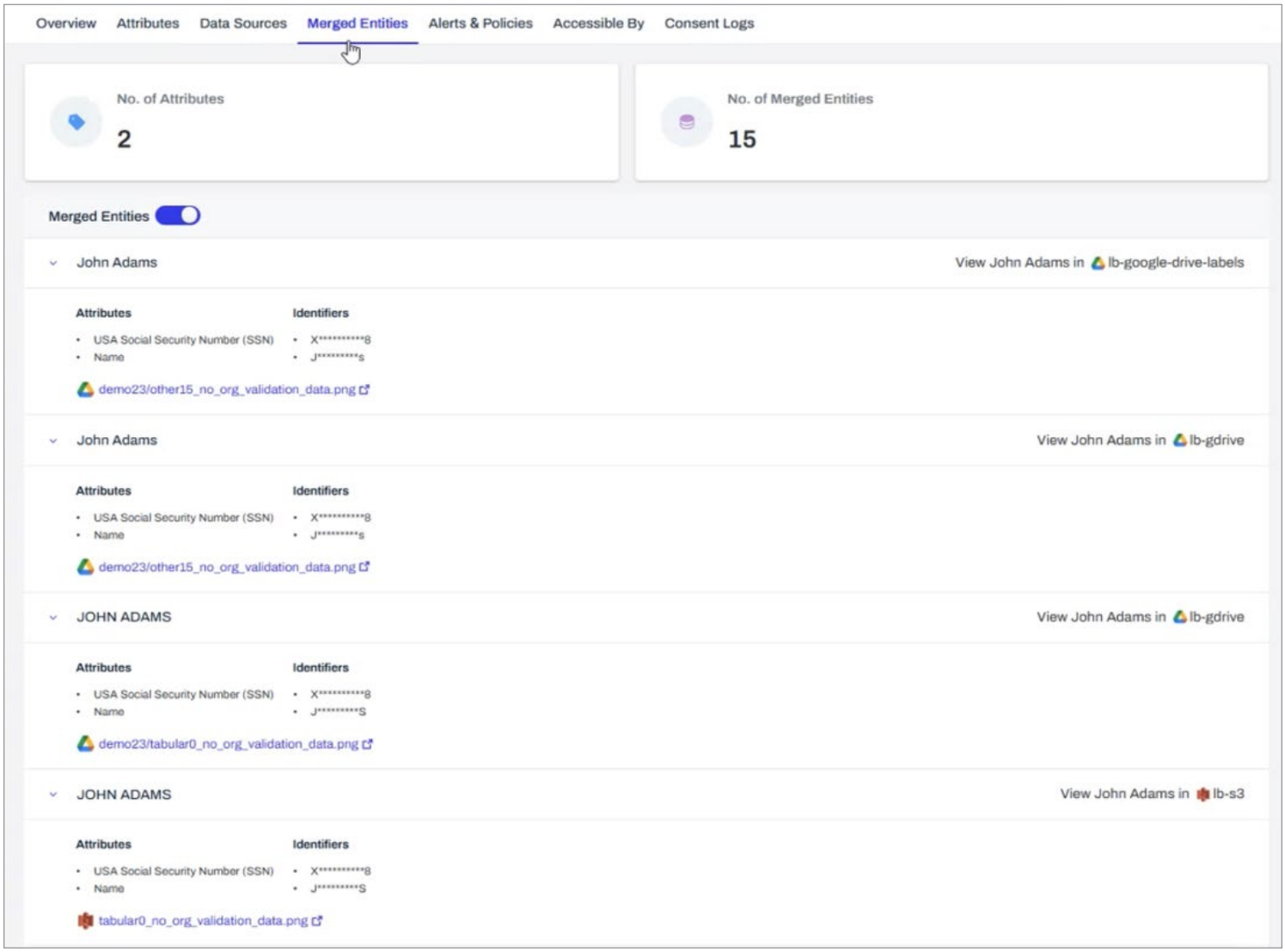


Figure 10. Merged Entities

The data discovery and classification capabilities in Lightbeam were impressive and detailed, and they allowed us to quickly discover sensitive data and to find other data sources and possible exposure risks as well.

With the rapid growth in AI applications and services across many organizations, DSPM tools are focusing more on common AI use cases seen in the enterprise. For example, Lightbeam can detect and alert on sensitive data being transmitted through Microsoft Copilot, as shown in Figure 11.

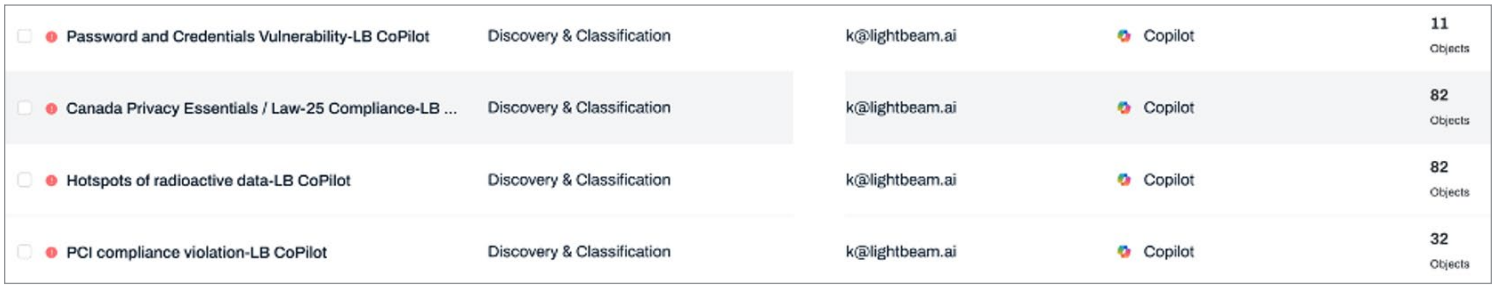


Figure 11. DSPM Alerting on Copilot

Lightbeam can track the AI prompts and responses, and easily allow you to track the users who are querying Copilot (see Figure 12).

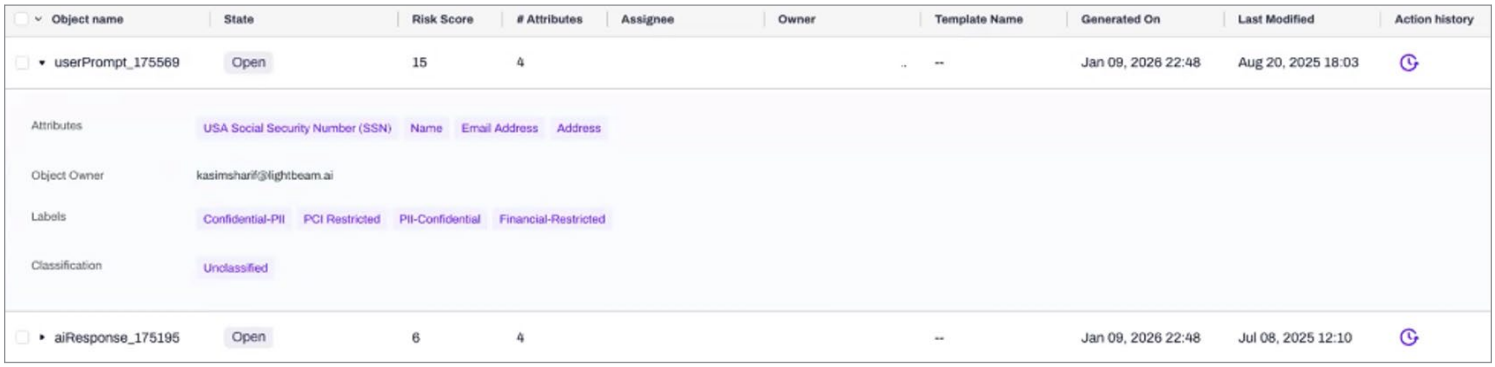


Figure 12. Exposed Copilot Requests and Responses

Lightbeam's integration with identity stores allows you to track the users and see what department they're in, what groups they're members of, and more. In addition, any specific alert can have a variety of actions applied to the event, including changing the state to a number of status tags such as "Approved" or "On Hold." You also can track down which users are querying Copilot for sensitive content by looking at the Copilot data sources tracked, with the "Prompts" category highlighted, as shown in Figure 13.

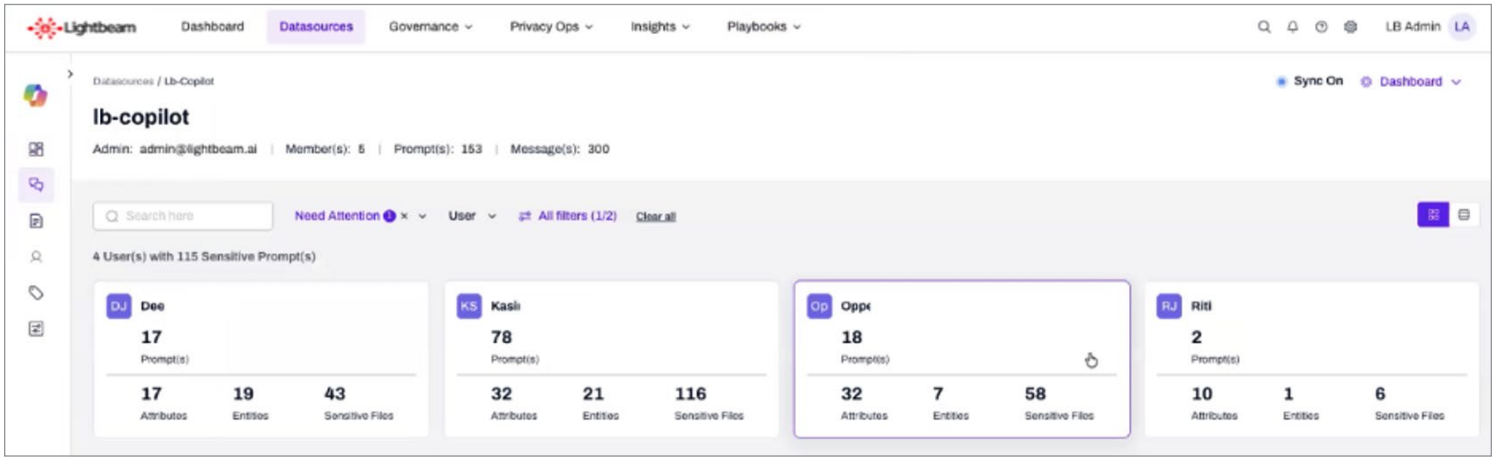


Figure 13. Copilot Prompt Tracking in Lightbeam

By digging into the details of any user's prompts, you can see how the prompt was accessed, what the queries were, and what the actual data returned may look like (depending on permissions).

A key element of any DSPM platform is the ability to define data security policies to help identify and remediate open or excessive access to content stored in data stores of all types. We explored some policies in place within our test environment, and started with a policy flagging open access to a range of storage nodes. We were able to see exposed objects and select objects to then take action on. In Figure 14, you'll see that we selected an exposed object and opted to revoke access to it.

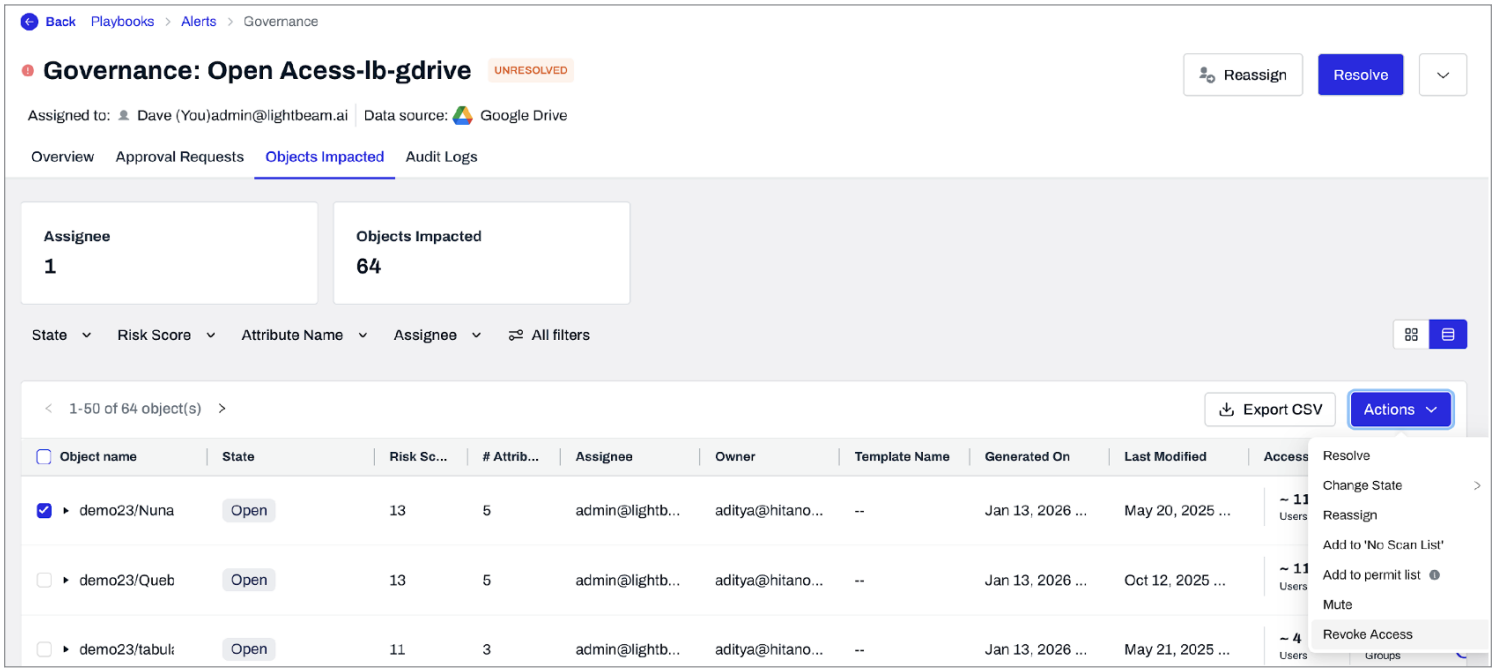


Figure 14. Revoking Access to an Exposed Object

The access revocation was immediately tracked in the audit log (see Figure 15). We executed this action manually for demonstration, but you also can automate it based on AI governance policies.

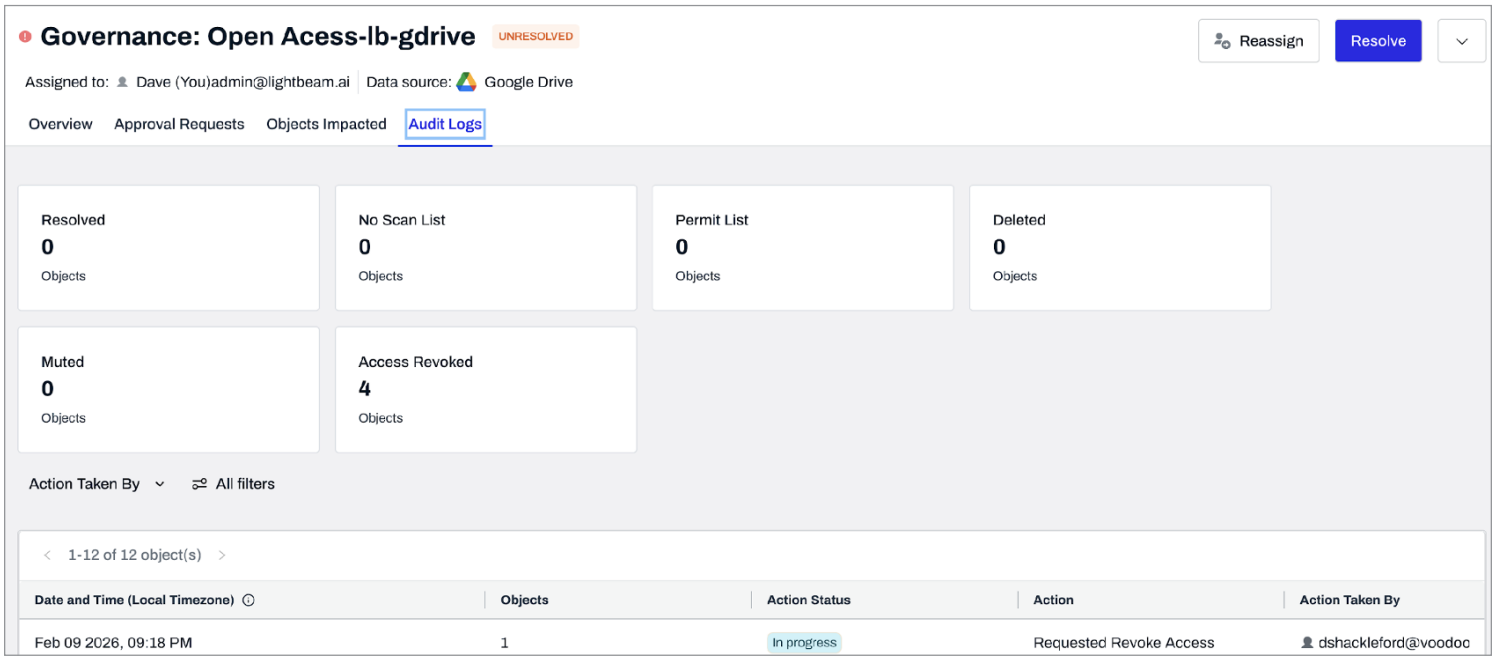


Figure 15. Access Revocation Logging

Finally, we checked out a range of reports available in the platform, and successfully exported multiple reports in CSV and PDF formats for review (see Figure 16).

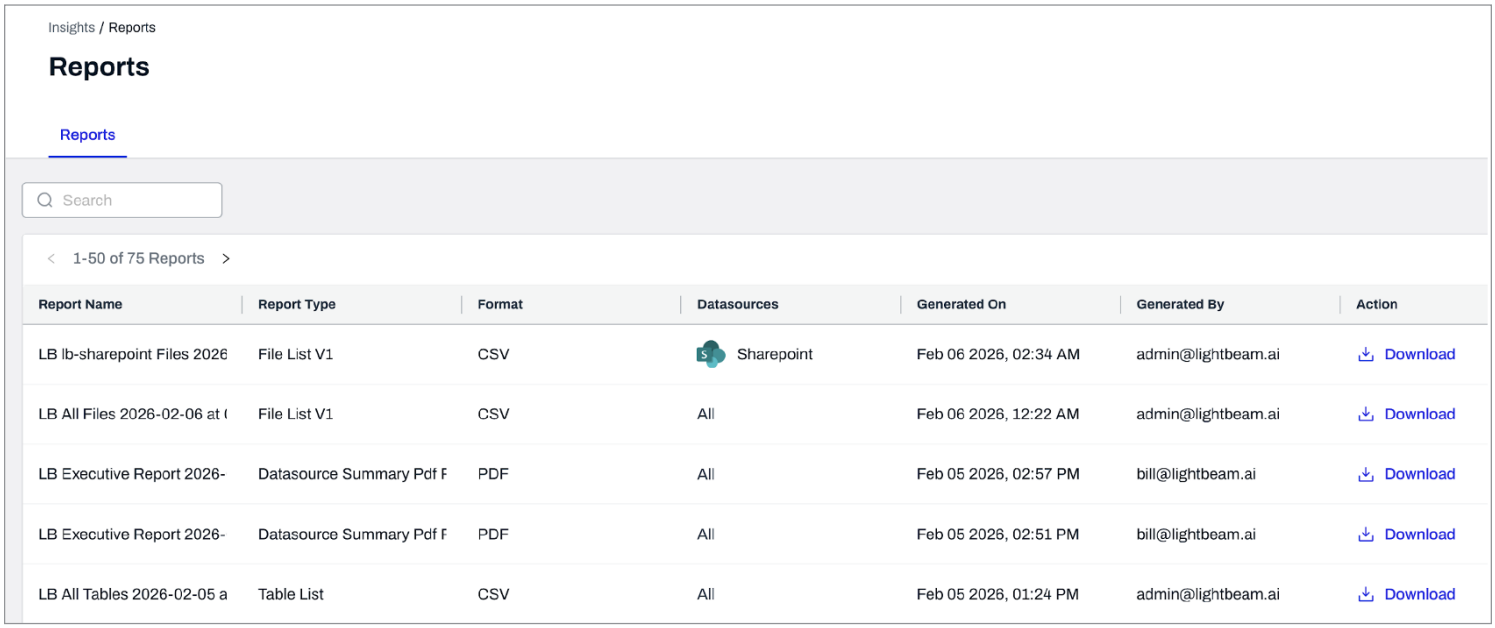


Figure 16. List of Lightbeam Reports Available

Access Governance

One of the strong differentiators of Lightbeam is the alignment with data and identity and access security controls and policies. A key element of the platform is tracking and reporting on users and identities that have access to data sources and objects. In the "Governance" dashboard of the platform, we found a variety of data sources that had open access (see Figure 17).

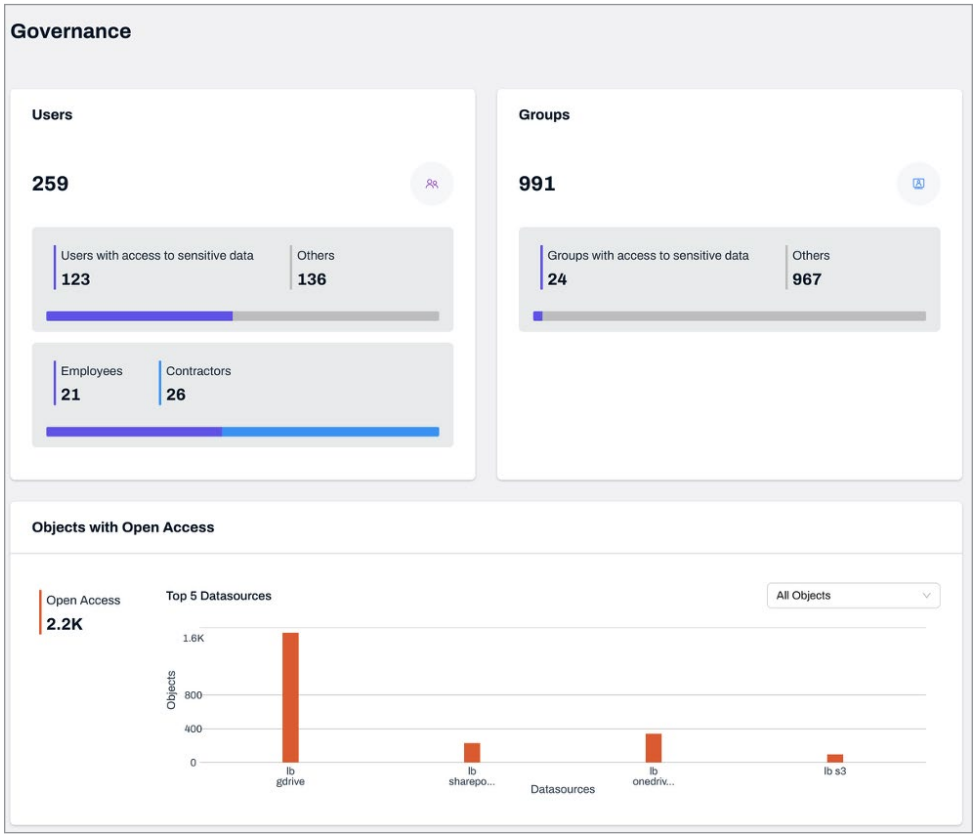


Figure 17. Objects with Open Access

By clicking into any of the objects, we were quickly able to see files that were exposed. For example, Figure 18 shows an AWS S3 bucket with exposed files.

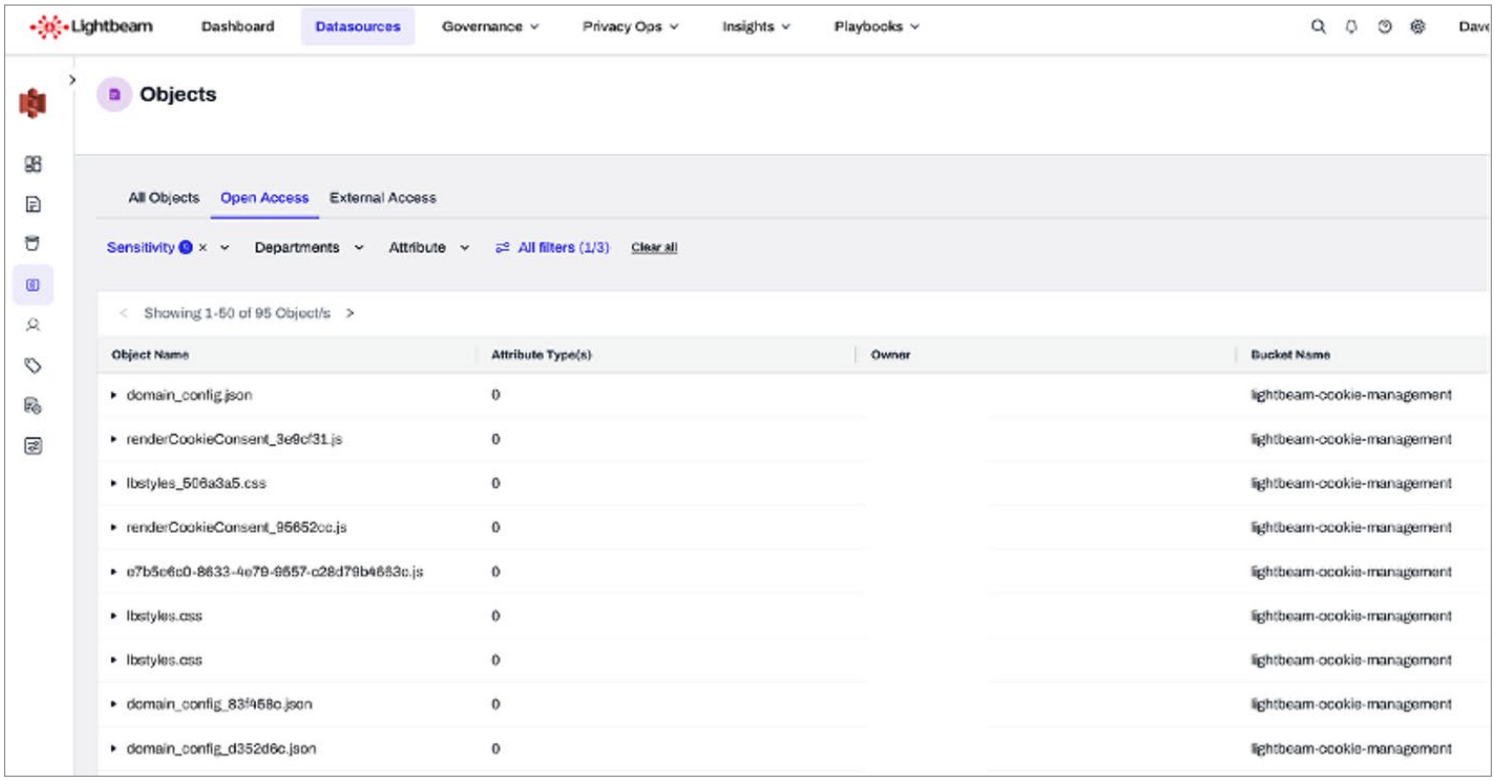


Figure 18. Exposed Files in an S3 Bucket

Lightbeam also offers analysts an "Access Review" governance module that allows you to analyze and report on a specific type of storage drive or folder (e.g., SharePoint, S3) with detailed information on both objects and users with access. Figure 19 shows a breakdown of users with access to a reviewed storage node.

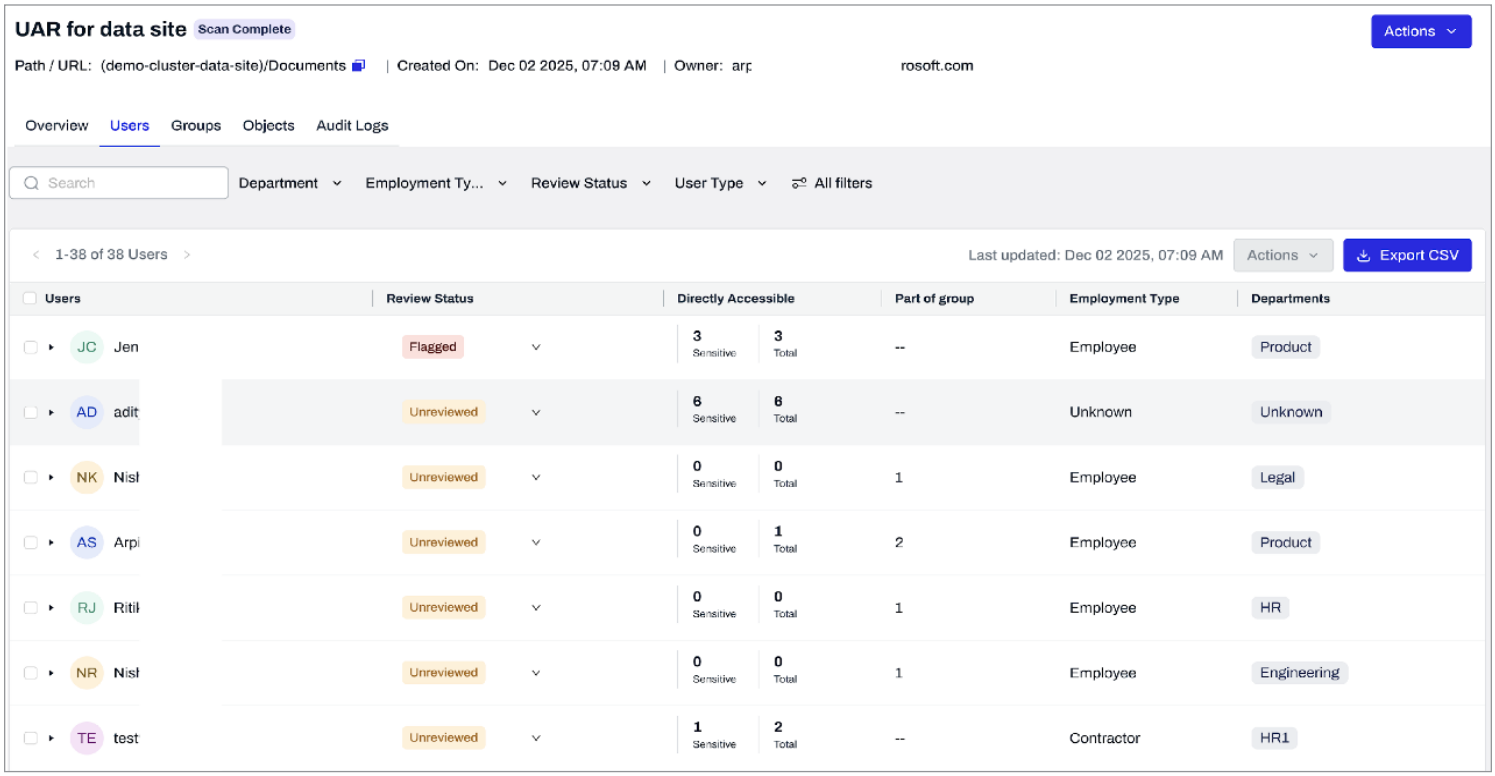


Figure 19. User Access Review for a Storage Object

In addition, each user can be labeled with a review status tag, and a variety of actions can be taken for users such as changing their assigned status or revoking access altogether. Analysts also can easily export all this information in a CSV file for analysis or import into other tools and databases.

We finished up our review of the core data access control and access review controls by assessing the policy engine to control external access and sharing for file types. The policy playbooks were simple to generate, with policy options for controlling external and internal access, data life cycle and discovery, user activity, ransomware indicators, and more. Rules for things like data labels and classification/sensitivity were flexible, and entire groups of file types/extensions also can be chosen. We created a new sample policy to restrict sharing of all text and word processing documents within all data stores (see Figure 20).

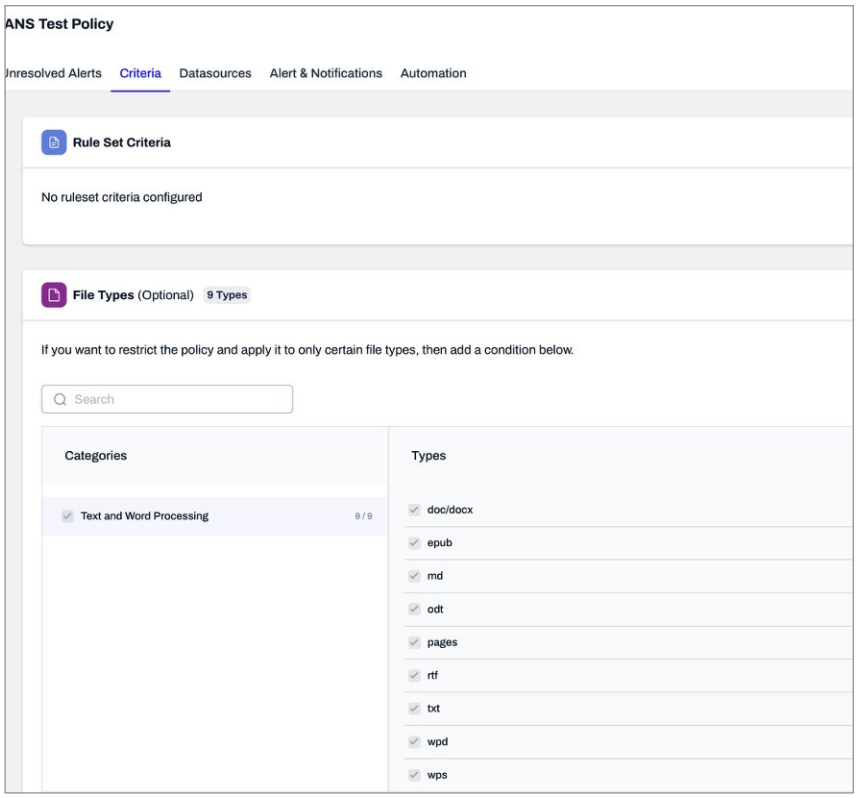


Figure 20. A Sample Document Sharing Policy

Sharing any of these types of files with external users would trigger a critical alert. It's also possible to revoke a share or block access through automation actions.

Data Classification Capabilities

Any DSPM product should have strong data classification and labeling capabilities to help organizations tag and track content across data sources. We first checked to see whether Lightbeam allowed us to create our own document classification labels and types, where "Classifications" in Lightbeam serve as the default tracking and tagging method, providing a hierarchical structure where documents belong to specific categories and sub-categories. In the "Insights" section of the portal, the "Documents" section has options for customers to track documents in data stores by labels (covered shortly), classification, or duplicate discovery in multiple locations. In the Classification section, we quickly created a new classification called "IP" with a description of "Source code" and then applied it to existing documents discovered (see Figure 21).

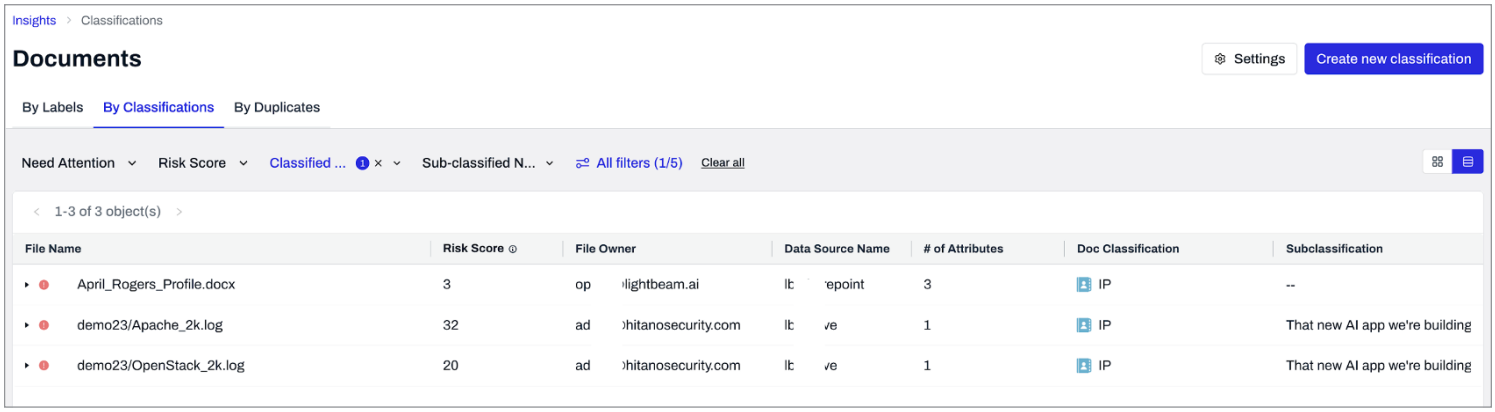


Figure 21. Tagging and Listing Custom Document Classification

The second type of data classification and tracking tool we evaluated was the definition of attribute sets. Lightbeam has a number of predefined sets available for use that coincide with well-known data types (e.g., financial, healthcare, personal data, PII), but we created a new set with a mix of different well-known attributes to scan across all known data stores. We were immediately able to report on where these data types resided (see Figure 22).

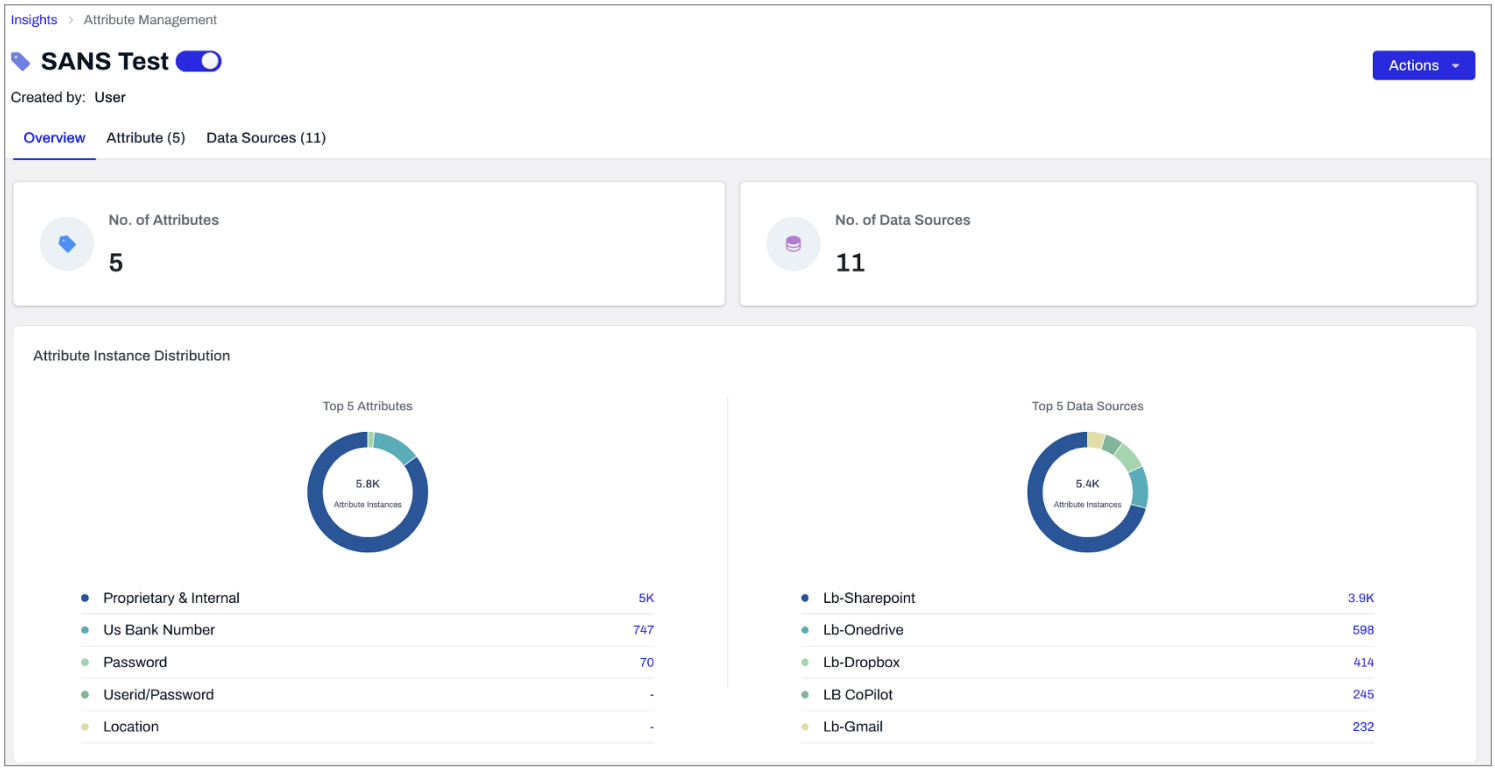


Figure 22. A Test Attribute Set Scan Across Data Sources

Lightbeam also allows you to include some custom and business-specific attributes in policy conditions such as external sharing. The third type of data classification element we explored was the Lightbeam label, which provides an alternative grouping method that can be enabled through user settings for more flexible organization when needed. We created a new label and label set in our review that mapped to a variety of high-sensitivity attributes (detected by Lightbeam) and classification designations (including our test "IP" classification created earlier). See Figure 23 on the next page.

Label Sets can be assigned a number of different label designations (often to differentiate priority and sensitivity), and all data in discovered data sources can be scanned to identify these data categories and types. Labels also can integrate with other major technology provider services like Microsoft Purview and Google Labels. We evaluated test integrations with both Purview and Google that were assigned and tracked in Microsoft OneDrive, SharePoint, and Google Drive.

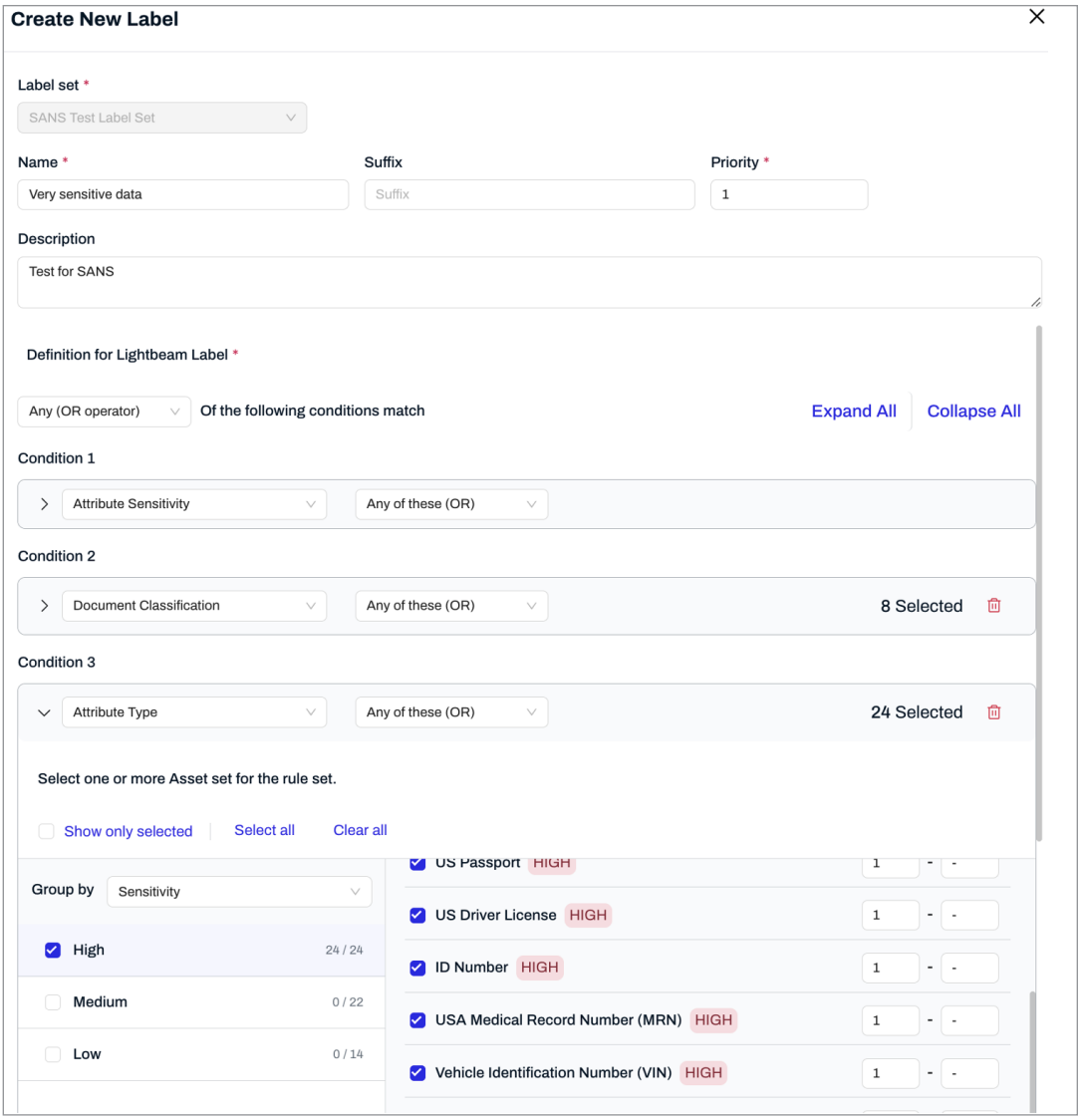


Figure 23. Creating New Labels and Label Sets

Data Retention and Minimization

Another key aspect of data security management is ensuring that data is retained for required periods of time (per internal policy and/or regulatory requirements) and kept to a minimum where needed and as applicable for privacy and legal reasons. Enforcing data retention and minimization policies has the added benefits of reducing the potential impact of data breaches and reducing storage costs. We created a test data retention policy that easily deletes data of specific types after a specified time period. The summary of this policy is shown in Figure 24, and we tested it to see if files were archived and alerts were sent out.

Retention policies are very flexible, and they can be configured based on extracted content from files, such as contract expiration, as well as specific labels and folders within data sources. Duplicate designations also can be applied to retention and archival policies.

Insider Threat, Breach, and Ransomware

DSPM solutions today also need to assist security teams with incident detection and response efforts during breaches and investigations. We reviewed some of Lightbeam's capabilities in this area, starting with general breach and ransomware detection, alerting, and automated response options. Given the prevalence of ransomware activity today, Lightbeam offers a governance policy type of "Ransomware Activity" that can be tailored to any organization's data sources. We created a test policy for all data sources in our lab environment that detected more than 10 encrypted file actions by any user account in several data sources that would suspend any user access immediately (see Figure 25).

Detailed audit logs are also available in the platform to help in analyzing incidents and possible breaches.

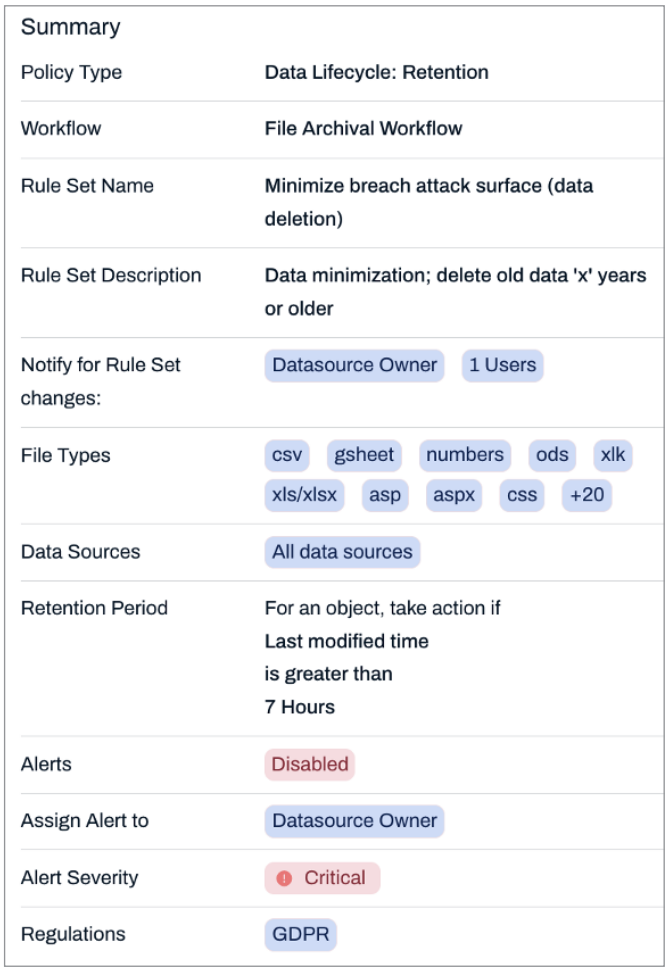


Figure 24. Sample Data Retention Policy with Alerting and Actions

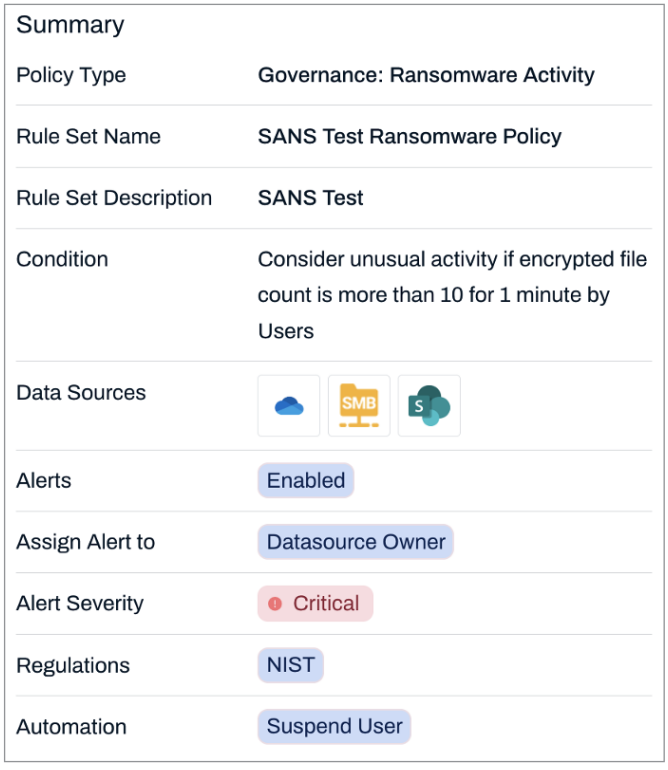


Figure 25. Ransomware User Suspension Policy

We reviewed some of the Lightbeam insider threat detection and response functions as well. For each data source, there are activity logs that show what a user has done over time and what files and data-related activities have been observed, as shown in Figure 26.

Users with unusual patterns of access or suspicious behaviors can be tracked and monitored more closely, as warranted.

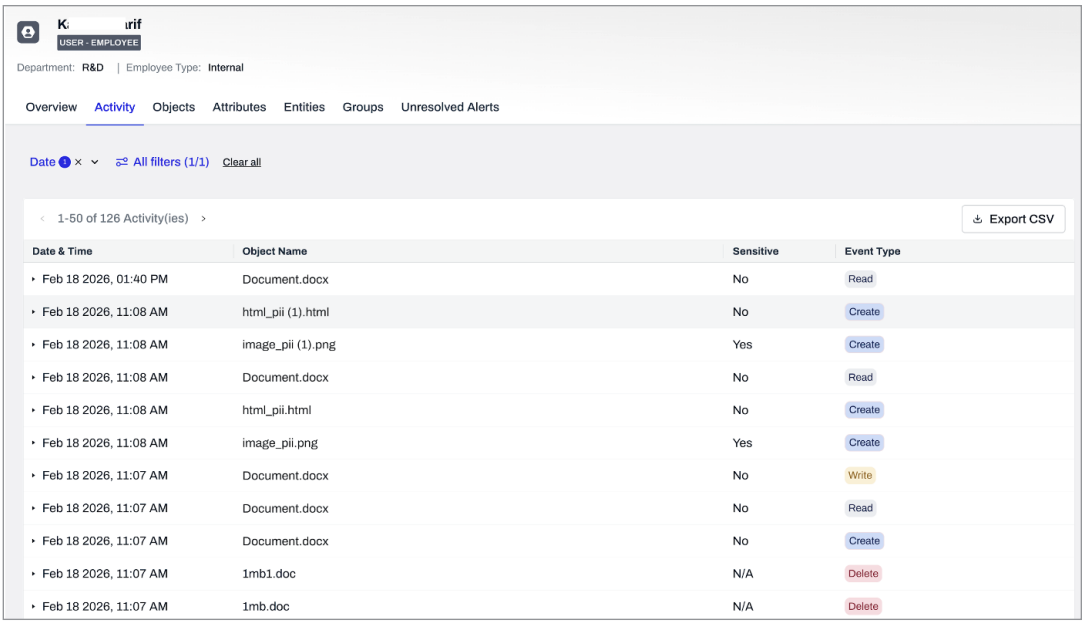


Figure 26. User Behavior Profiling in a Data Source

Privacy Operations

Lightbeam has an entire category of privacy-focused functionality labeled "Privacy Ops." Given the criticality of meeting privacy requirements and protecting sensitive personal data, we spent time looking at the various capabilities the platform offers. First, we explored the Privacy Ops dashboard, which provides a range of privacy-related data at a glance, including data subject requests (DSRs), consent requests and consent management, risk and privacy impact assessments, and records of processing activity (RoPA). See Figure 27.

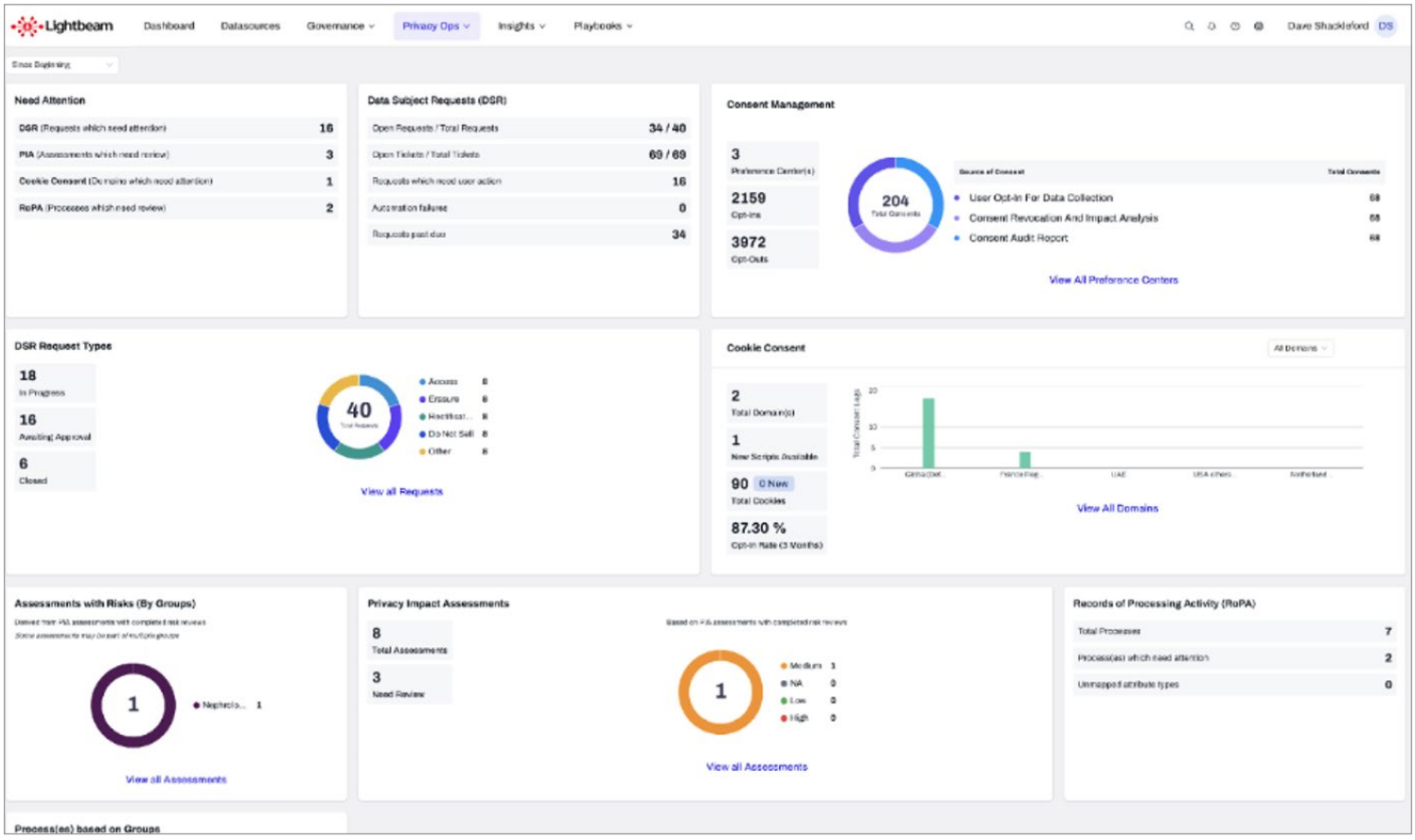


Figure 27. The Lightbeam Privacy Ops Dashboard

Our first function to explore within Privacy Ops was the Data Subject Requests (DSRs), which are a set of rights provided to individuals under GDPR, CCPA, and other privacy regulations. Privacy Ops can handle the entire life cycle and workflow associated with DSRs, including real-time tracking and continuous monitoring and change detection. Unlike privacy tools that require you to collect the DSR data through tickets sent to people in the organization, Lightbeam already had a map of all the data associated with the data subject through its data identity graph and can automate the production of the report to the data subject immediately. The platform has an easy-to-use form builder (with a number of prebuilt templates) to generate and manage DSRs. We generated a new test form that included inputs like first and last name, email, and country for a request that included data access, data rectification, data erasure, and more (see Figure 28).

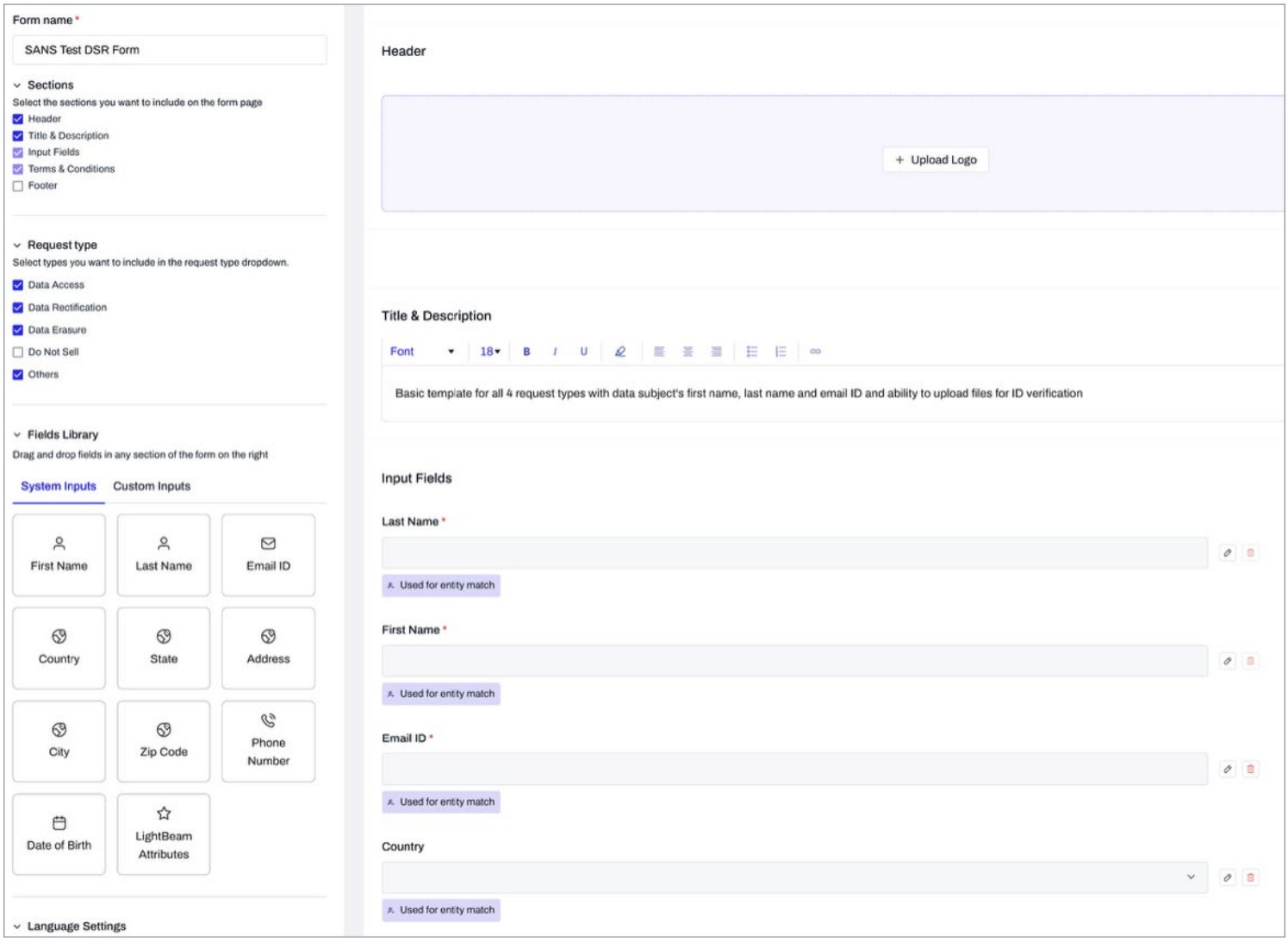


Figure 28. A Sample DSR Form

We then filled out the test form (see Figure 29).

Once the request was logged, we approved it in the DSR module, as shown in Figure 30.

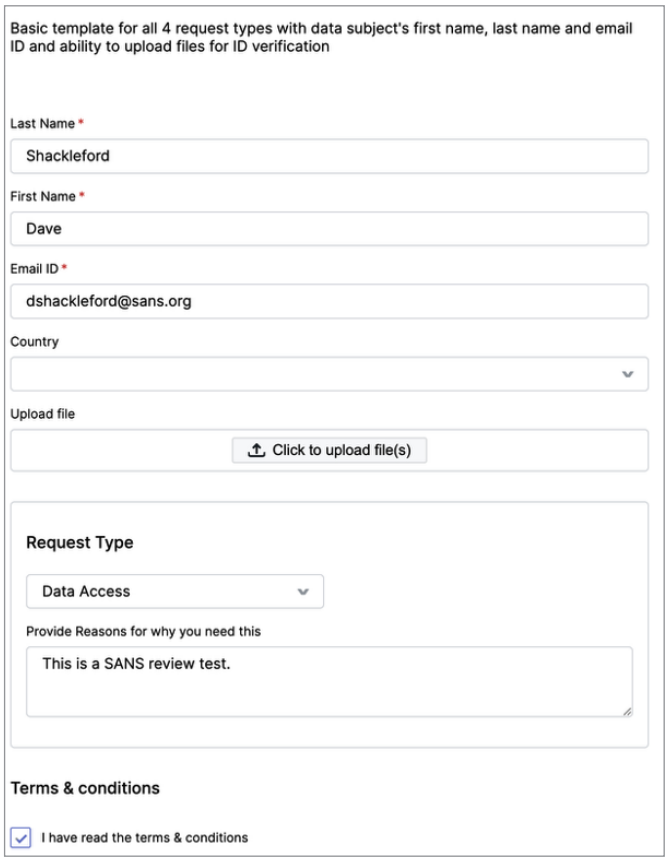


Figure 29. Testing a Sample DSR Form

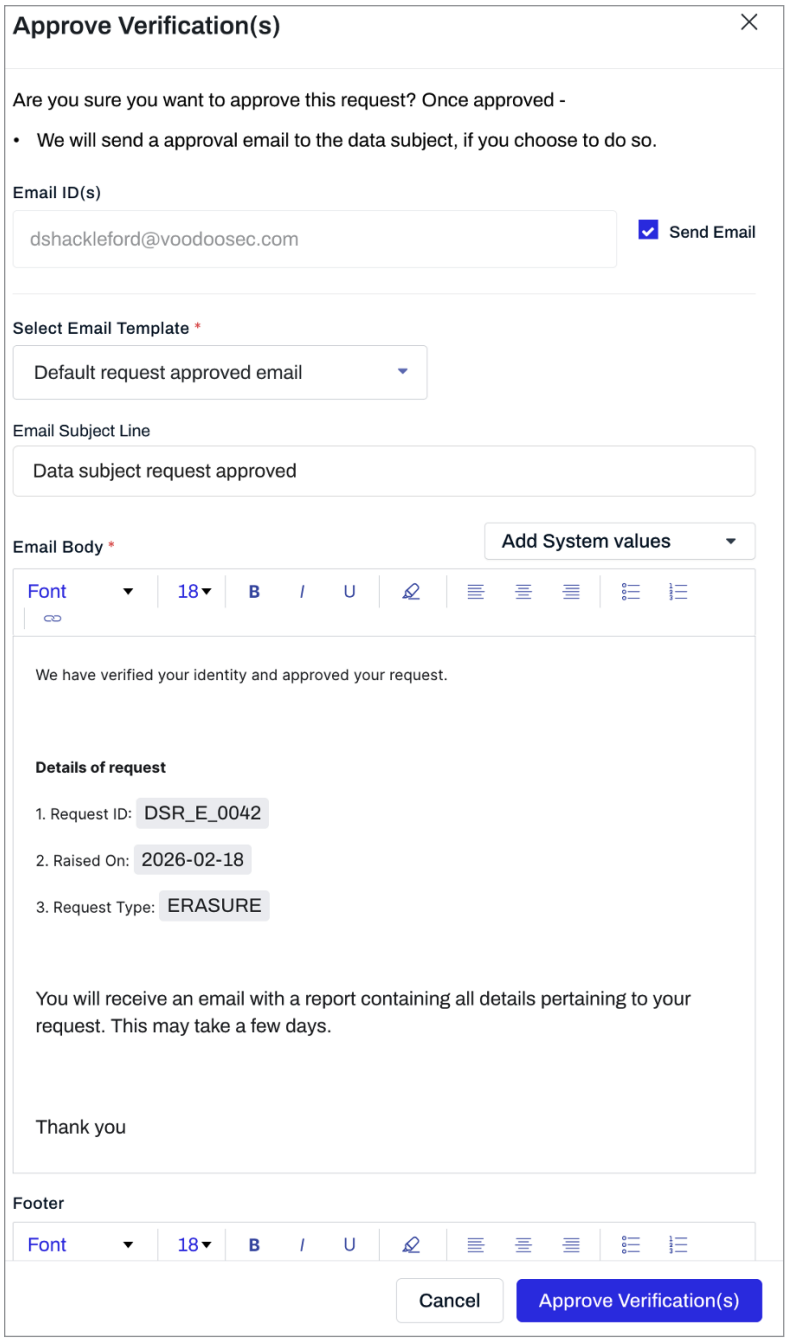


Figure 30. Approving a DSR Request

Once approved, an email confirmation was sent, and then we could proceed to close the request if/as warranted. We explored the DSR capabilities and workflows to ensure that private data associated with identities was validated, too. Figure 31 shows manual validation during a review of a subject's sensitive data.

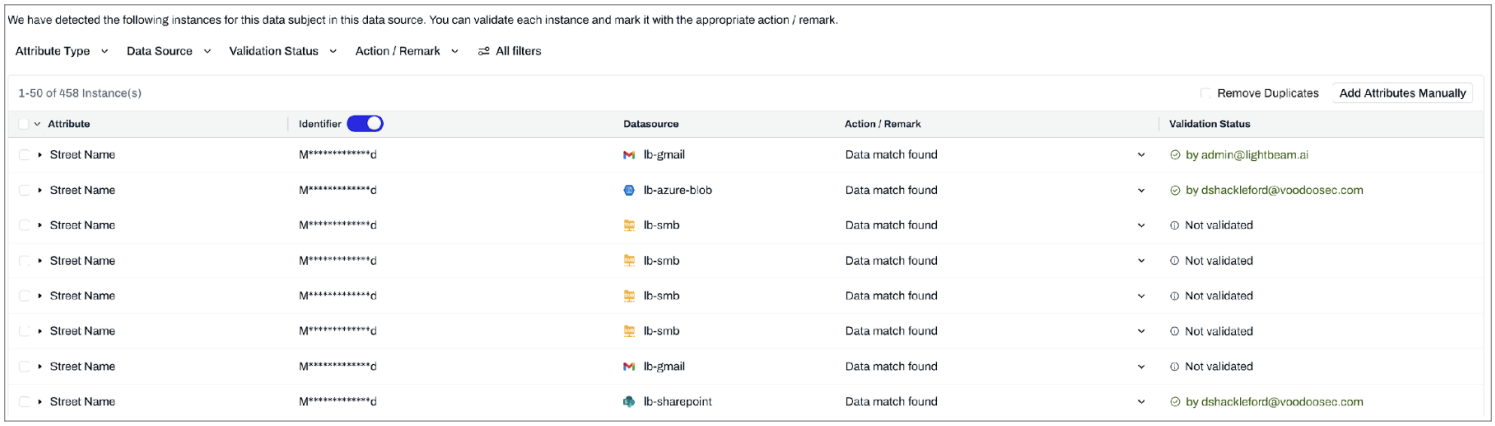


Figure 31. Validation of Sensitive Data Across Multiple Data Sources

For several DSRs, we generated reports of the private information discovered across data sources. The reports are generated very quickly and can be downloaded in PDF format or automatically sent to the data subject requesting the DSR. We also generated a new "Data Erasure" DSR for a person's sensitive data detected across multiple data sources. Lightbeam made it simple to generate and assign a ticket to the specific data source owner, as shown in Figure 32.

Overall, we found the entire DSR module to be easy to use, granular in reporting and workflow control, and flexible in configuring templates, automation sequences, and DSR forms.

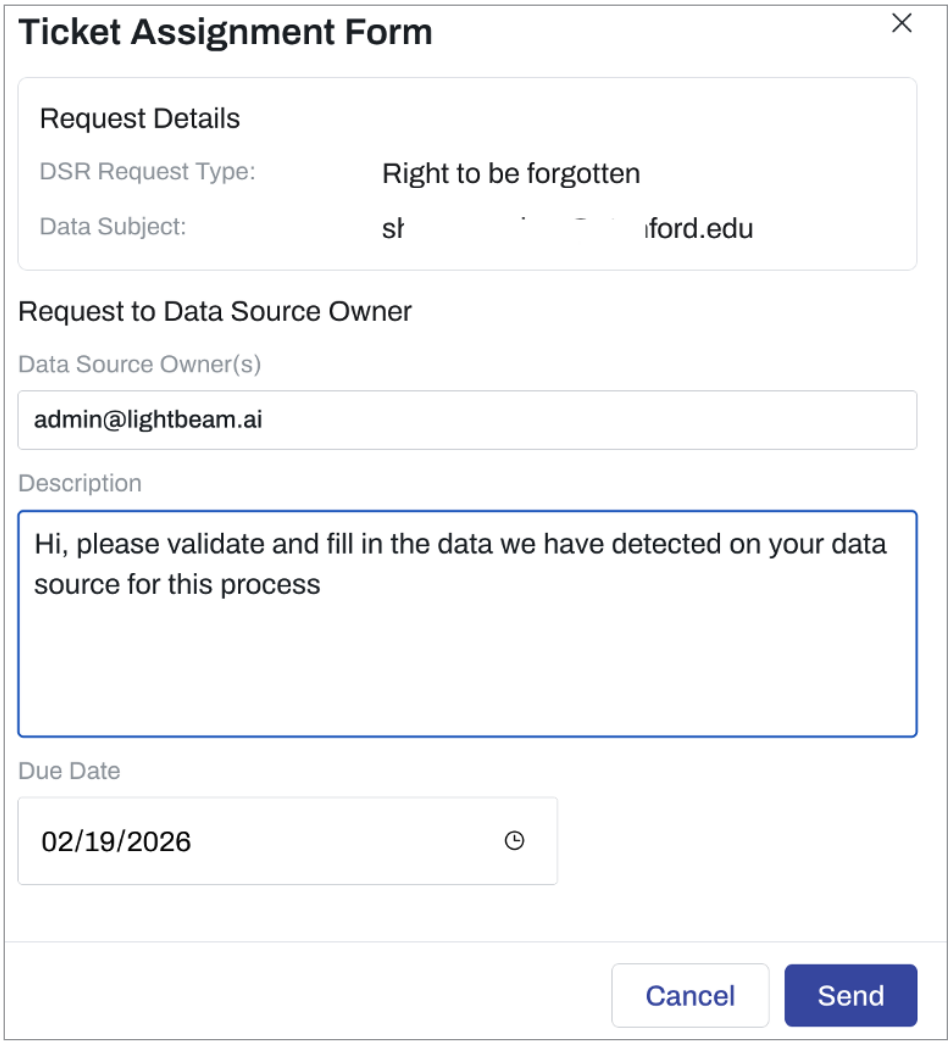


Figure 32. Data Source Owner Notification Workflow

The next major module we explored in Privacy Ops was the Record of Process Activity (RoPA). RoPA refers to documentation that organizations must maintain regarding their data processing activities under GDPR and other privacy regulations. Lightbeam provides a number of RoPA default templates that organizations can use or clone into new, customized models. We created a new template (somewhat generic) and then created a new processing activity request. The platform has an extensive level of details you can add in, including data types, users, classifications, data sources and storage locations, and more. An example of an in-process RoPA is shown in Figure 33.

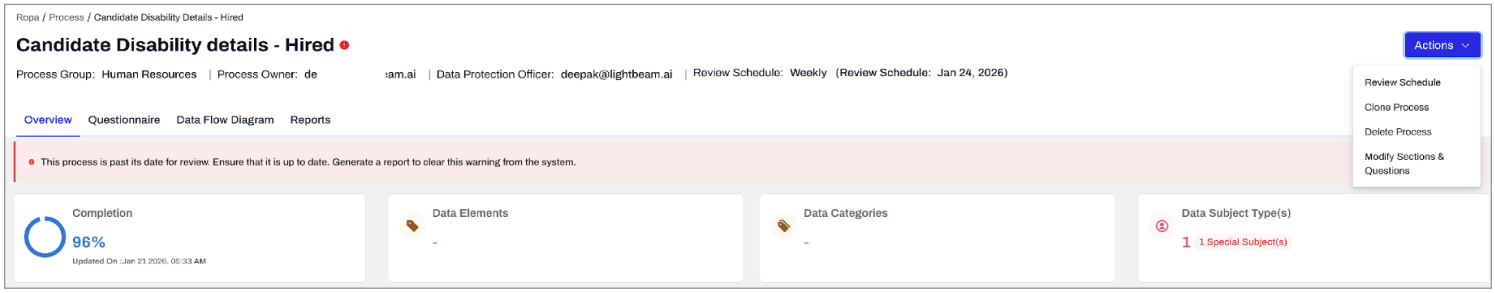


Figure 33. RoPA in Process

The final feature of the Privacy Ops module we explored was Privacy Impact Assessment (PIA), which allows organizations to assess privacy risks associated with data processing activities. PIA helps teams create risk-based questionnaires, assign risk levels to answers, collaborate with stakeholders, and generate assessment reports. We created a new PIA using existing Lightbeam templates for IT software development related to various privacy laws like CCPA and GDPR, and found the process to be straightforward. You create a questionnaire, invite collaborators, then submit for review by privacy officials once questionnaires have been completed. Lightbeam prepopulates information, such as data sources, types of sensitive information present in the data sources, and other information from the DSPM scanning, to streamline the PIA process. A sample dashboard for a completed PIA is shown in Figure 34.

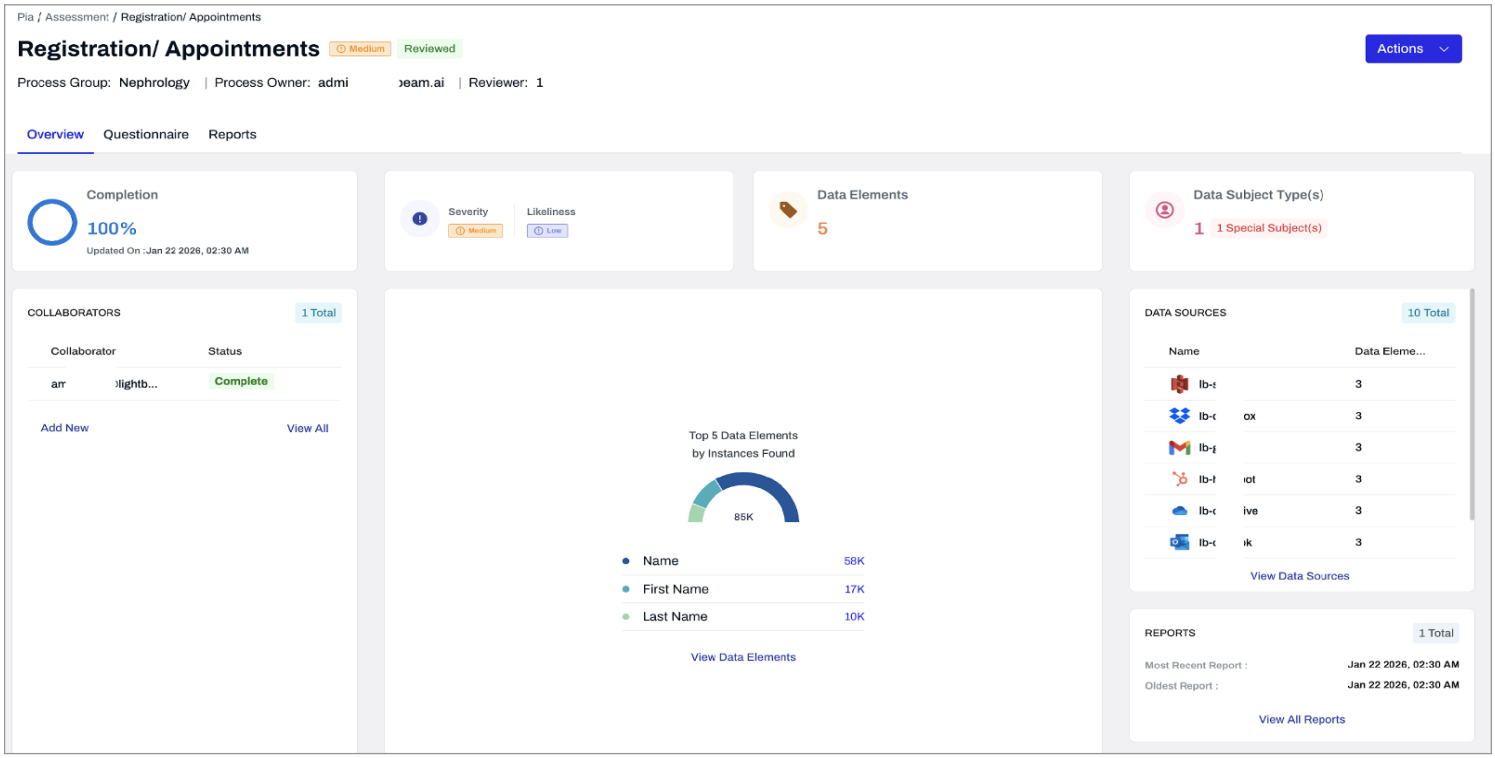


Figure 34. A Completed PIA

Overall, we found the Privacy Ops components of Lightbeam to be easy to use, with numerous capabilities that would readily help privacy, cybersecurity, risk management, and legal teams develop and maintain privacy-specific controls and reporting across their varied data environments.

Conclusion

Data protection across hybrid environments has become critically important as organizations operate simultaneously across on-premises data centers, multiple public clouds, SaaS platforms, and increasingly distributed endpoints and AI-driven services. Sensitive data now moves fluidly between storage tiers, APIs, collaboration tools, and third-party integrations, making traditional perimeter-based controls insufficient. The challenge lies in inconsistent visibility, fragmented policy enforcement, identity sprawl, and the difficulty of classifying and tracking data across diverse platforms that were never designed to operate as a unified security domain. Regulatory pressures, data sovereignty requirements, and the growth of non-human identities and automation further complicate governance. As a result, many organizations struggle not with understanding the importance of protecting data, but with achieving consistent, scalable controls that can follow data wherever it resides, moves, or is processed.

The Lightbeam platform provides a broad variety of controls and deep telemetry around both data stores and the identities affiliated with accessing those data stores, as well as the data subject identity associated with the content (e.g., employee, customers, partners). During our review, we found that we could easily navigate the console, create and manage policies and alerting, and establish a complete end-to-end workflow around DSPM, access governance, and privacy in many environments. There are a range of prebuilt templates and tools to help teams with classification, labeling, grouping, and categorizing data, and identity affiliation and observability are woven in everywhere. Many different stakeholders would benefit from a platform like Lightbeam, including cybersecurity operations, risk management, legal, compliance, and privacy, both from tactical controls and enforcement opportunities as well as comprehensive reporting.