Cull Early, Save Big: How Harvesterâ??s Filtering Transforms eDiscovery

Description

Reducing data volumes before review is always an option. With rising ESI volumes and tighter budgets, legal teams need smarter ways to collect, filter, and deliver data. Pinpoint Labsâ?? Harvester is built for just that.

As a forensically sound, high-performance collection engine, Harvester lets users cull data intelligently during initial collection or apply filters later in the workflow. Whether filtering upstream, downstream, or both, Harvesterâ??s flexibility and power have helped clients dramatically reduce costs and timelines.

Trusted On-Prem Collection for Maximum Privacy and Control

Harvester runs as a fully on-premises solution, enabling organizations to collect data directly to secure local drives or network shares without routing through the cloud. This approach reassures legal teams, IT departments, and custodians that sensitive data remains within their control throughout the process, meeting privacy standards and regulatory requirements with confidence.

Especially in high-security environments or sensitive investigations, Harvesterâ??s ability to perform powerful, targeted collections locally is a major advantage. Thereâ??s no need to upload large data sets to external portals or third-party platforms just to filter them laterâ??Harvester puts the tools directly in your hands.

For workflows that require remote data delivery, Harvester also includes an optional uploader engine. When configured by the administrator, this secure uploader can transfer collected data to trusted destinations such as AWS S3, Microsoft Azure Blob Storage, or SFTP servers. Importantly, this feature is never automatic, collections are only uploaded when explicitly instructed, ensuring that even the most security-conscious environments remain in full control of their data. Harvester jobs can be preconfigured and scripted in advance, allowing automated targeting, filtering, and output to designated collection

locations.

Early Filtering That Delivers Results

Harvester supports filtering by date range, keyword, file type, deduplication, and more, before any data hits review. With PST and OST regeneration, filtered messages and attachments are saved to new containers, allowing for leaner review sets. Combined with deNISTing and hash-based deduplication across backups or overlapping custodian archives, the reductions can be dramatic.

Case Study 1: 6TB Reduced to 96GB

A national service provider was tasked with covertly collecting 6 terabytes of data from Exchange, Outlook, and Lotus Notes. With 64 Boolean and proximity keyword filters, plus deNIST and dedupe, Harvester narrowed the data to 96 gigabytes, just 0.016% of the original size. The collection was executed remotely using the Harvester Server and reported daily to the legal team.

Read the Evidox Case Study

Case Study 2: 10TB Across 34 Custodians

A litigation support group faced 10 terabytes of data across 34 custodians. Using Harvesterâ??s remote job deployment, deNISTing, and extension-based filtering, they performed preconfigured collectionsâ??even amid unpredictable networks. The result: faster collections, significantly lower data volumes, and avoided travel and disruption costs.

Read the Case Study

Case Study 3: Compliance Triaged in 400TB

A forensic consultant used Harvester to assess and triage 300 to 400TB of data during a security incident. By filtering in place, they avoided extracting everything and identified only 8.7TB as affected. Harvesterâ??s data assessment mode and indexing tools gave the team the ability to run multiple search passes across the full dataset, saving over $200K in alternative tools and labor costs.

Read the Ellwood Evidence Case Study

Built for Multiple Stages and Strategies

Some teams prefer broad collection with downstream filtering once more context is available. Harvester fully supports this approach. Its defensible front-end collection engine allows teams to preserve data to secure local drives or vaults without applying filters, then bring those datasets to in-house systems for later review and targeted filtering.

This enables multiple review passes as search criteria evolve, without needing to recollect or risk missing important content. Harvester supports both minimal and broad collections with full audit trails, providing flexibility without compromising defensibility.

Filtering upfront can:

- Cut review sets from terabytes to gigabytes
- Lower hosting and processing costs
- Improve relevance and reduce privileged data risk

But Harvester also protects teams who need to cast a wider net first, then refine downstream. Whether using keyword filters and PST regeneration up front, or full-set triage and dtSearch passes later, Harvester adapts to your legal strategy.

From rapid in-place assessments to precision-targeted filtering, Harvester continues to redefine what is possible in defensible data culling. Whether your workflow favors early precision or broad preservation with iterative review, Harvester gives you the tools to stay fast, focused, and defensible.

And as a trusted, on-premise solution, Harvester ensures your data stays private, secure, and under your control every step of the way.

Date
04/29/2026
Date Created
01/07/2026