FIELD NOTES FROM USC DATA SCANS

What we find in the files nobody looks at

Five places to look for sensitive data before a migration, Copilot rollout or cleanup

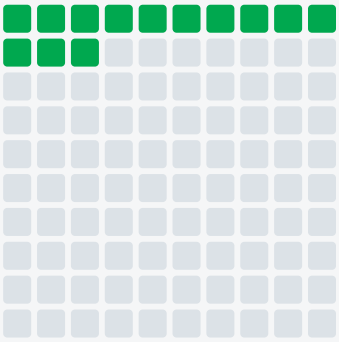
We have spent ten years inside the shared drives, mailboxes and SharePoint sites of regulated businesses. The pattern repeats: the business knows its systems well and has little idea what is inside its files. This is what that looks like in practice, and where to start looking in your own.

ONE WEALTH ADVISORY MIGRATION

13.1%

44,532 of 338,674 files contained personal information.

USC Data scanned Strategic Wealth's files before its move to SharePoint, and remediated those files as part of the project.



Each square = 1% of files scanned

At a second advisory firm, Tribel Advisory, USC Data scanned and remediated nearly one million files before a CRM migration.

Five places it hides, and where to start

Place	What we typically find	First action
1 Mailboxes of people who left	IDs, payroll and contracts sent as attachments over years	Confirm who owns the mailbox and its retention status
2 Scanned documents on shared drives	Driver's licenses, passports and signed forms saved as images, which keyword search misses	Make sure your scan reads images, not just text files
3 Exports from business systems	Spreadsheets pulled from payroll, CRM or accounting "for one report"	Find them, and confirm the system of record still holds the data
4 Old sharing links	Files shared with a link years ago that are still open	Review links on files holding personal data first
5 Laptops	Downloads and desktop folders that never went back to a governed location	Include laptops in the scan, not just servers

What a finding looks like

A useful scan does not stop at a count. Each finding should tell you where the data is, what it is, who can reach it and what to do next.

Location	Data found	Who can reach it	Recommended action
Finance › Payroll › 2019 payroll.xlsx	Social Security numbers, bank account numbers	Everyone except external users	Restrict to the payroll group; review against retention
HR › Onboarding › Scans	Driver's licenses, passports (images)	"Anyone" link, no expiry	Remove the link; move to a restricted library
Mailbox of a former employee	Client dates of birth and account numbers	Two delegates still assigned	Confirm retention; remove delegates; archive

Illustrative example. Not client data.

29% of IT leaders say AI has already surfaced sensitive data

In a March 2026 survey of 851 IT leaders across the US, Canada and Europe, 29% said Copilot or another AI tool had surfaced sensitive internal data that should not have been accessible. Another 8% were not sure. Source: ShareGate, State of Microsoft 365 2026.

Run your own first pass this month

1. **Pick the system you trust least.** For most firms it is the oldest file share or SharePoint site.
2. **Scan a defined scope first:** one site, one share or a set of mailboxes. 20,000 files is a workable first scope, not a sample of your whole estate.
3. **Close broad access on files that hold personal data:** company-wide grants, old links, files with no owner.
4. **Review aged files against your approved retention schedule and any legal holds** before anyone authorizes disposal.
5. **Record what you did and when.** That record is what an insurer, auditor or court will ask for.

See your own numbers, free

USC Data will scan up to 20,000 of your files inside your own environment and walk you through the findings in 30 minutes. Nothing leaves your network.

Get the free 20,000-file scan

Sources: USC Data case studies, uscdata.com/case-studies (Strategic Wealth; Tribel Advisory); ShareGate, State of Microsoft 365 2026.