INFORMATION LEADERSHIP

AI and Privacy

The Foundation You Can’t Ignore

Sarah Heal
CHIEF EXECUTIVE AND FOUNDER

informationleadership.com | 0800 001 800 | sarah@informationleadership.com

Microsoft
AI Cloud Partner

Microsoft
Solutions Partner
Modern Work

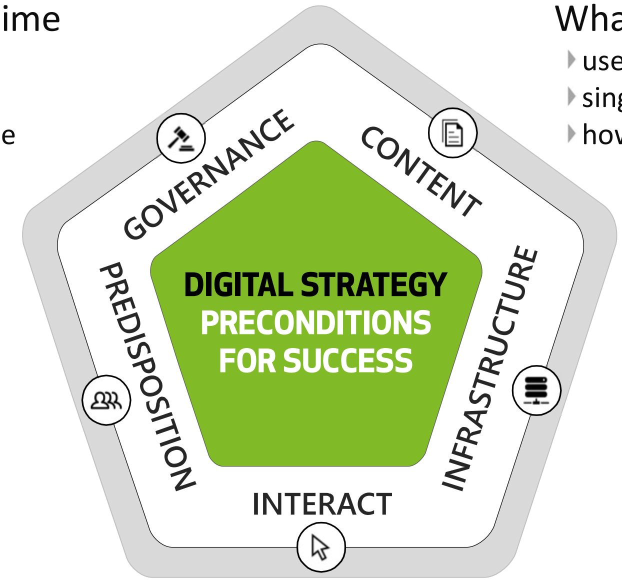
Our beginnings … 2004

How maintained over time

- principles & policies
- rules, processes, measures
- roles, resourcing & influence

What information

- useful & complete
- single source of truth & up to date
- how captured & created



DIGITAL STRATEGY
PRECONDITIONS FOR SUCCESS

GOVERNANCE

CONTENT

INFRASTRUCTURE

INTERACT

PREDISPOSITION

Willingness to use

- how positioned
- buy in & WIIFM
- training & support

Underlying systems

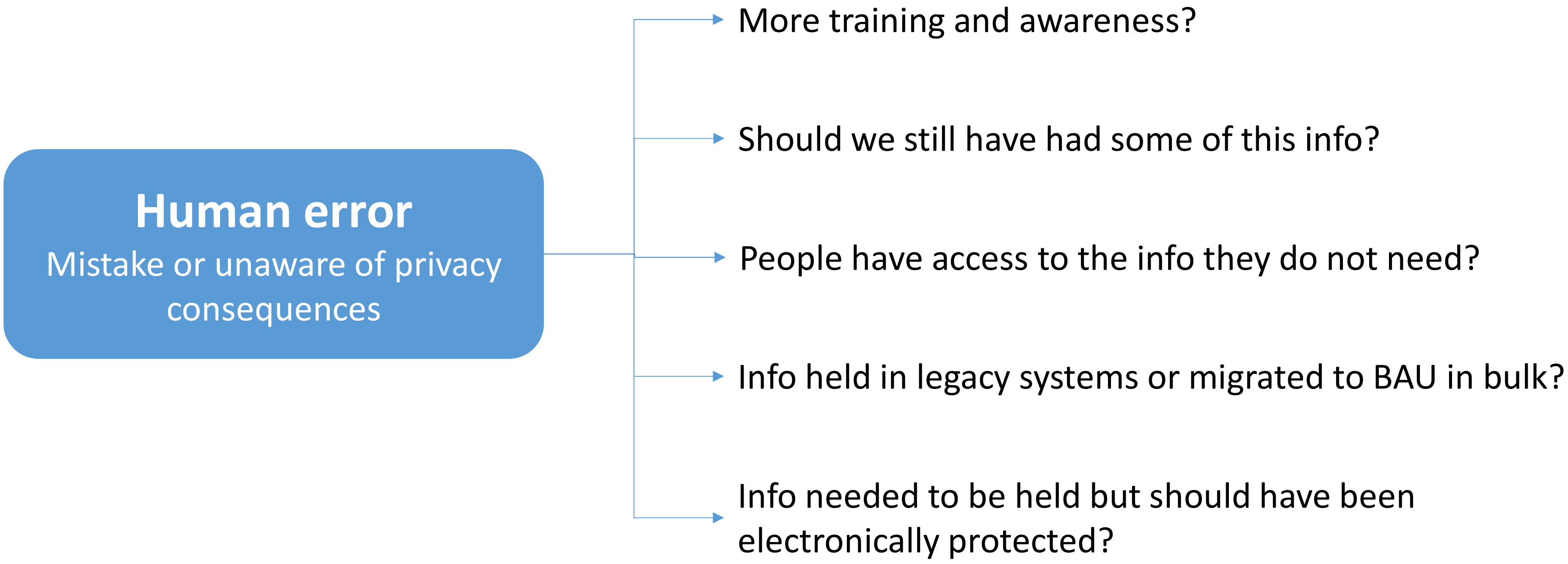
- systems architecture
- security & functionality
- integration & processes

How people interact

- information architecture
- usable interfaces, whatever device
- delivering value & convenience

INFORMATION MANAGEMENT

Moving from blame to root causes



Human error

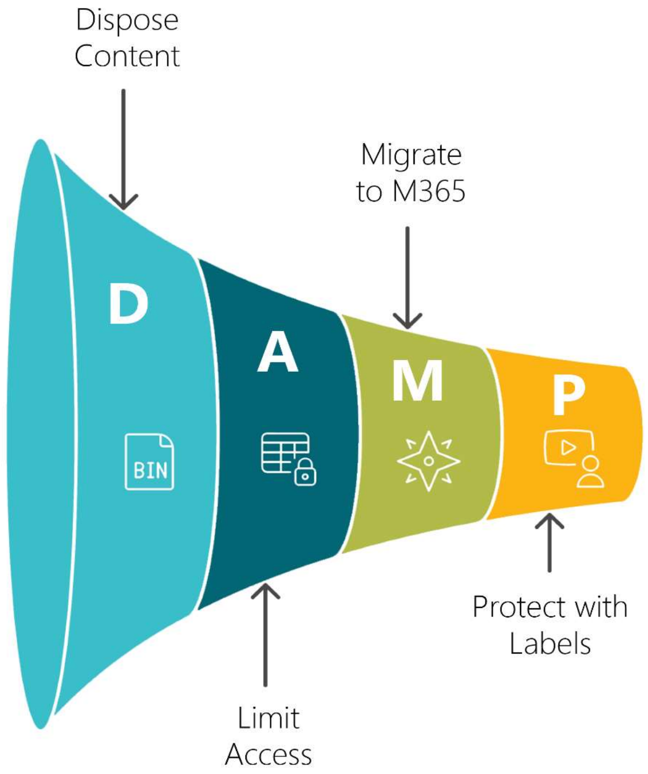
Mistake or unaware of privacy consequences

- More training and awareness?
- Should we still have had some of this info?
- People have access to the info they do not need?
- Info held in legacy systems or migrated to BAU in bulk?
- Info needed to be held but should have been electronically protected?

DAMP

Proactive Information Governance

Your messy information ...



... much better protected

INFORMATION MANAGEMENT RISKS BECOME PRIVACY RISKS

Legacy Content: What could possibly go wrong?

SEARCH EXAMPLE

Jamie collates and write a report on a Health+Safety risk

DAMP

Info used was over 5 years old and no longer needed – it should have been removed as part of systematic deletion

DAMP

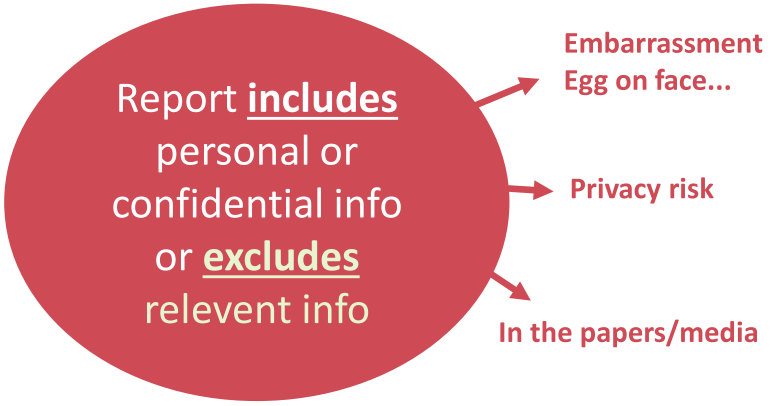
Info overshared by through OneDrive or Teams collaboration

DAMP

Relevant info was unavailable because it was in a OneDrive, legacy system or fileshare

DAMP

Info returned is sensitive and should have had a sensitivity label with DLP added as part of the case



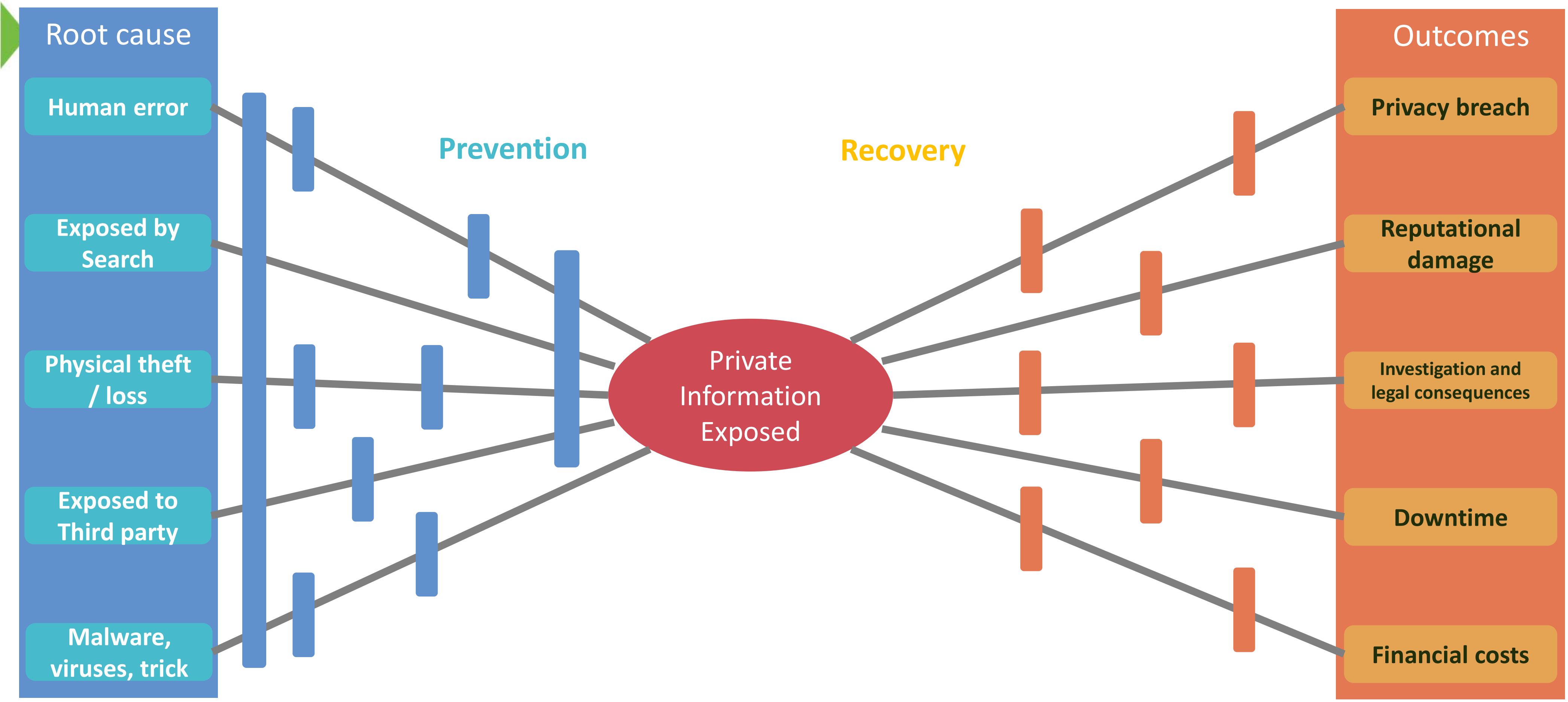
Report **includes** personal or confidential info or **excludes** relevent info

Embarrassment
Egg on face...

Privacy risk

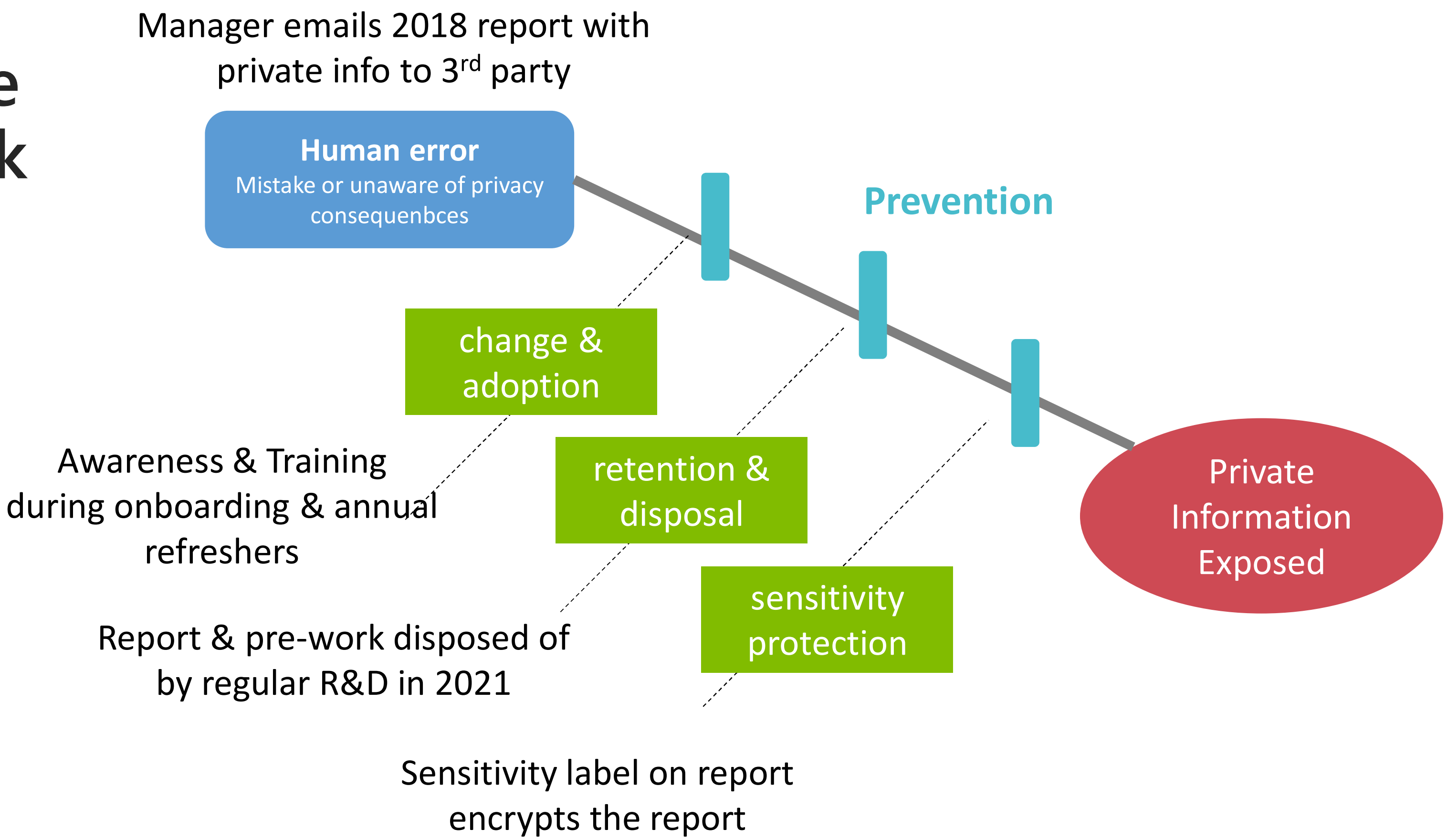
In the papers/media

INFORMATION ACCIDENT BOWTIE



RISK MANAGEMENT

Any one gate would eliminate or lower the risk


Manager emails 2018 report with private info to 3rd party

Human error
Mistake or unaware of privacy consequenbces

Prevention

change & adoption

retention & disposal

sensitivity protection

Private Information Exposed

Awareness & Training during onboarding & annual refreshers

Report & pre-work disposed of by regular R&D in 2021

Sensitivity label on report encrypts the report

Five underlying tactics

1. Heightened Awareness

TACTICS

Awareness is the First Defence

When staff and contractors understand what private information looks like they:

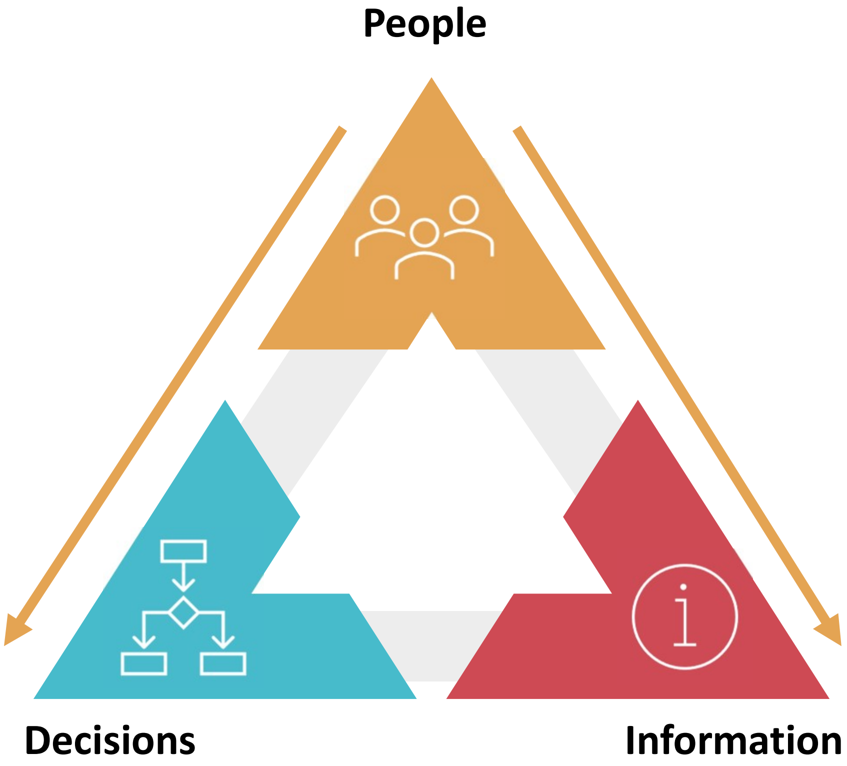
1. Make better decisions about where to store and share it
   (e.g. not dumping it into a Teams chat or OneDrive folder)

2. Treat the information more responsibly
   – tagging, protecting and disposing of it when needed.

If awareness is low, people:
1. Won't recognise private information in the first place
2. Which leads to poor decisions e.g. sharing in unsecured spaces
3. And mishandled information (e.g. exposed to AI or external users)

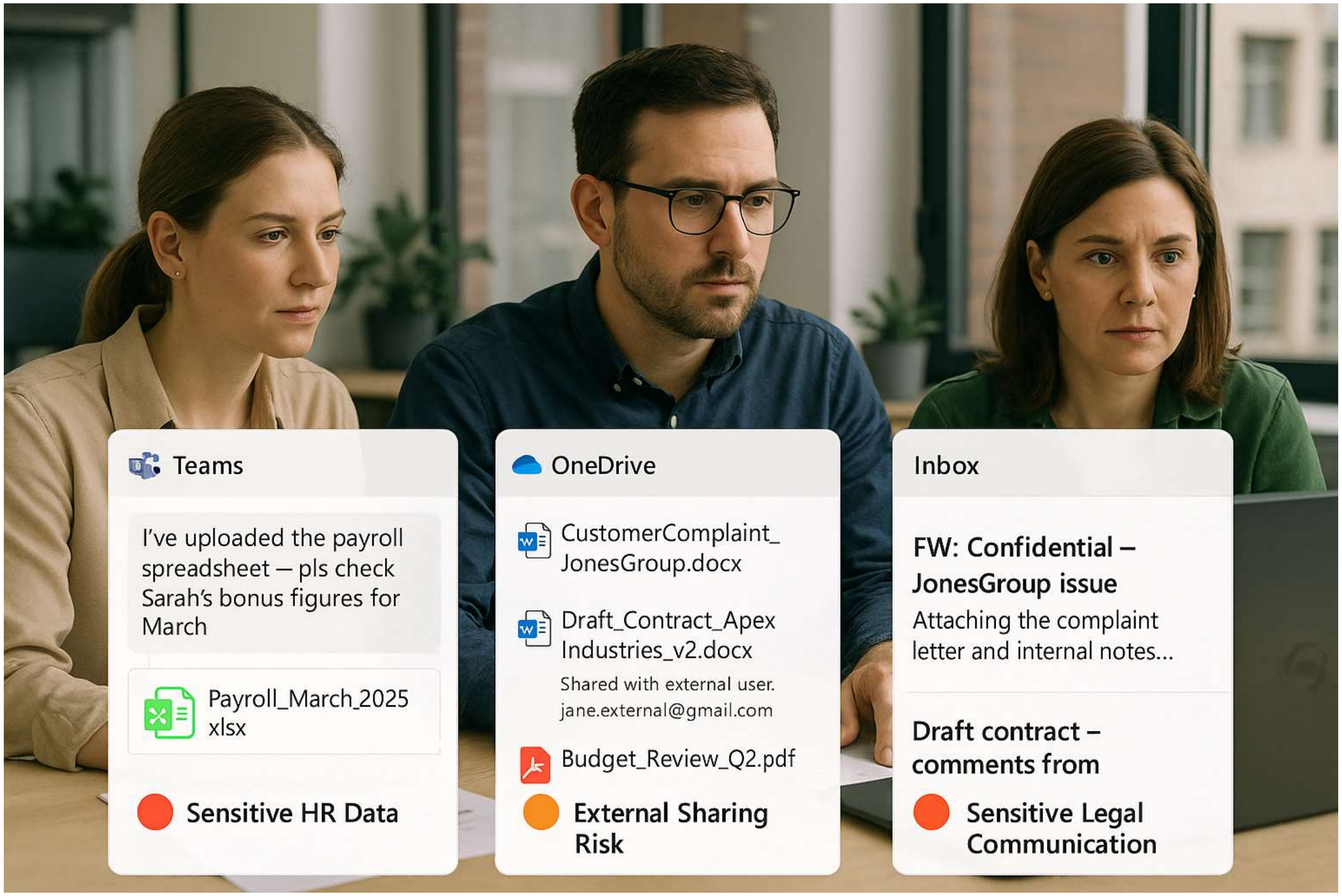
The root of most privacy mishaps isn't maliciousness.
It's simply people not realising the risk.

Awareness is the upstream fix that prevents downstream problems.



TACTICS

What 'Private' Info Looks Like in BAU



Teams

I've uploaded the payroll spreadsheet – pls check Sarah's bonus figures for March

Payroll_March_2025 xlsx

● Sensitive HR Data

OneDrive

CustomerComplaint_JonesGroup.docx

Draft_Contract_Apex Industries_v2.docx

Shared with external user. jane.external@gmail.com

Budget_Review_Q2.pdf

● External Sharing Risk

Inbox

FW: Confidential – JonesGroup issue

Attaching the complaint letter and internal notes...

Draft contract – comments from

● Sensitive Legal Communication

TACTICS

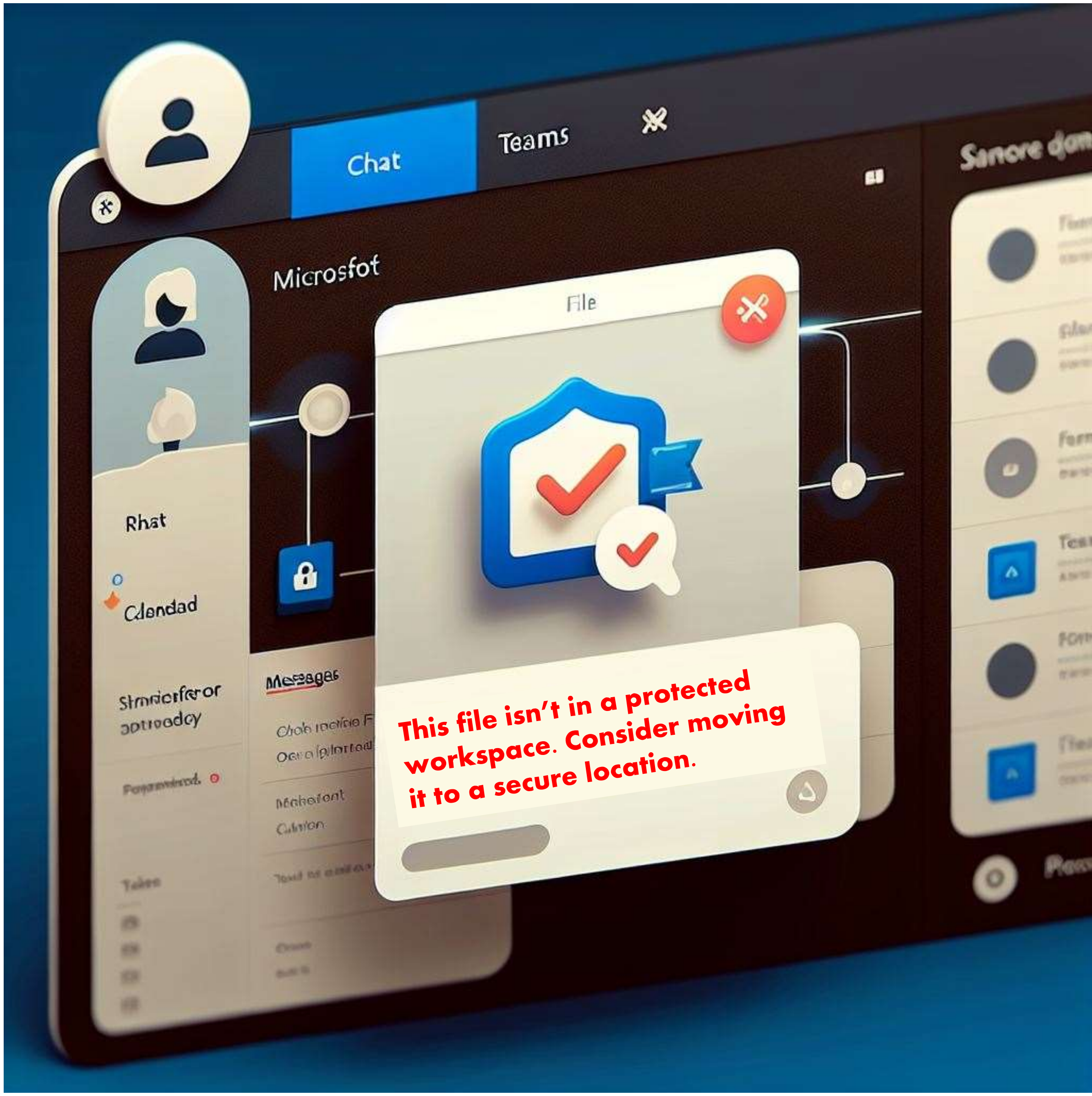
Ask Your Team...

1. Can they identify private information they create or use in their role?
2. Do they understand what's sensitive or regulated?
3. Have they received clear, localised, examples

TACTICS

Boosting Awareness Across the Organisation

- Localised, role-based training
- ‘Just-in-time’ prompts (tooltips, nudges)
- Quick reference guides and AI “what not to share” reminders



This file isn’t in a protected workspace. Consider moving it to a secure location.

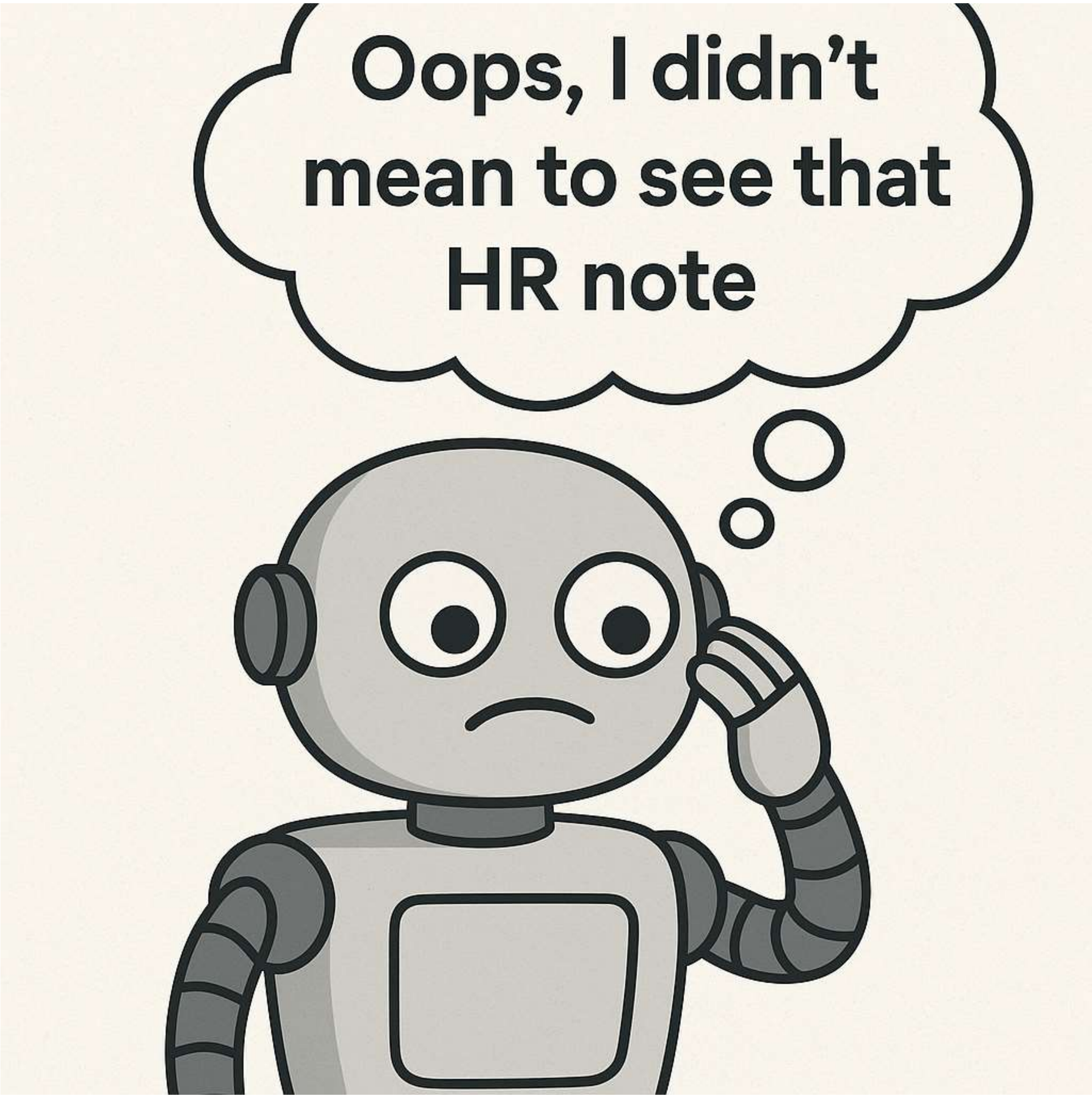
TACTICS

Why Awareness Matters for AI

So why does all this awareness matter when it comes to AI? Because AI tools like Copilot are only as safe as the data and context we give them.

If staff do not recognise what private information looks like, they will not think twice about including it in prompts or pulling it from the wrong place.

That means AI can end up surfacing or summarising content that was never meant to be seen more widely.



2. Bias to structured workspaces

INFORMATION LEADERSHIP

let's make work better

EMAIL, ONEDRIVE, FILESHARES, AD-HOC TEAMS

Biggest risk is the unstructured content

Because its like herding cats

Better to provide ways they can create, collaborate, file and find that are ***better*** than email, OneDrive…

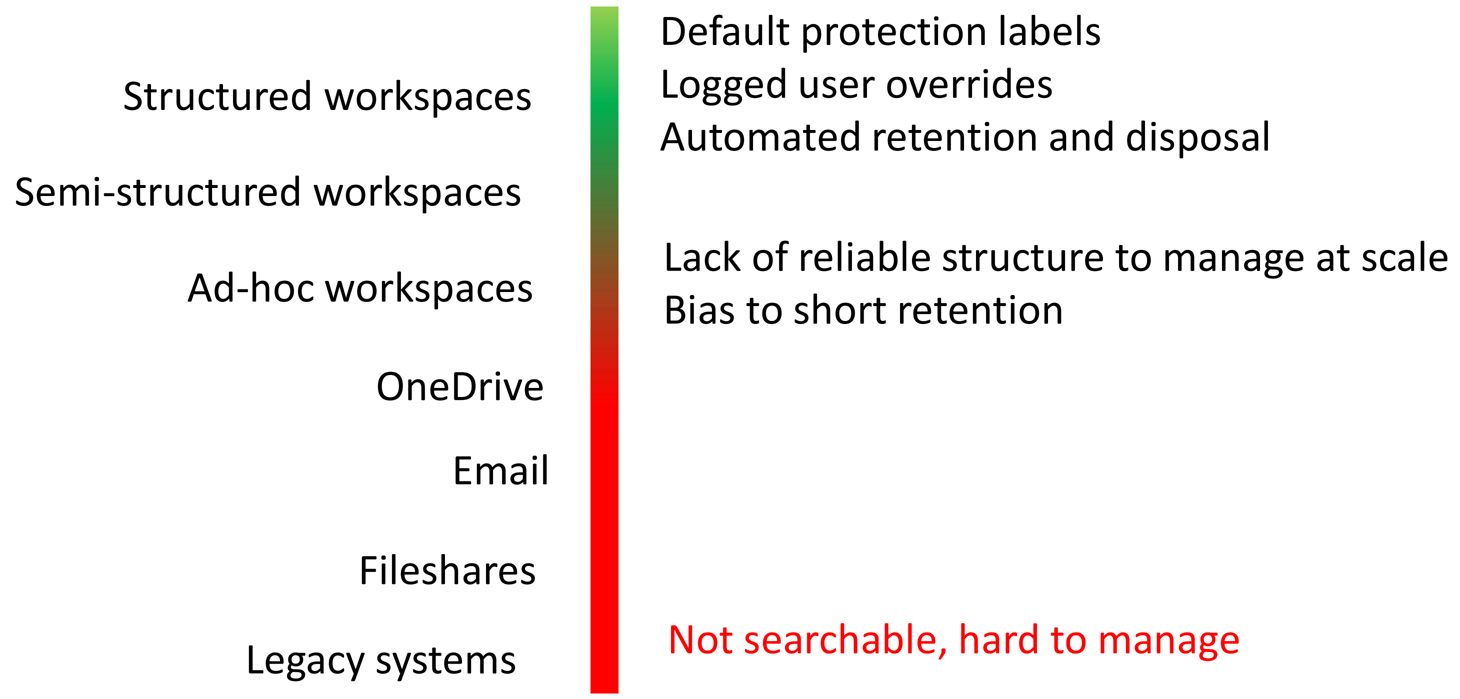
Yes, “change management” and “policy/procedures” can help but the real enduring gamechanger is when they see and then use what for them ***makes work better***

FOR KEEPING PRIVATE INFO PRIVATE

Info stores: the best, good, bad + ugly

Where people do work

Where search and structured search get info from



Structured workspaces

Semi-structured workspaces

Ad-hoc workspaces

OneDrive

Email

Fileshares

Legacy systems

Default protection labels
Logged user overrides
Automated retention and disposal

Lack of reliable structure to manage at scale
Bias to short retention

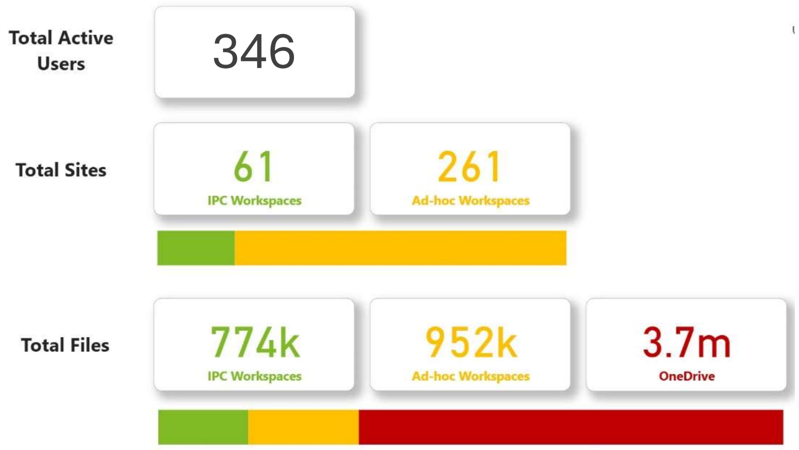
Not searchable, hard to manage

DASHBOARD ANALYTICS OF ONEDRIVE SILOS

So measure it for your org, ponder then act

High OneDrive use means it's easier and better for users to do their work there …

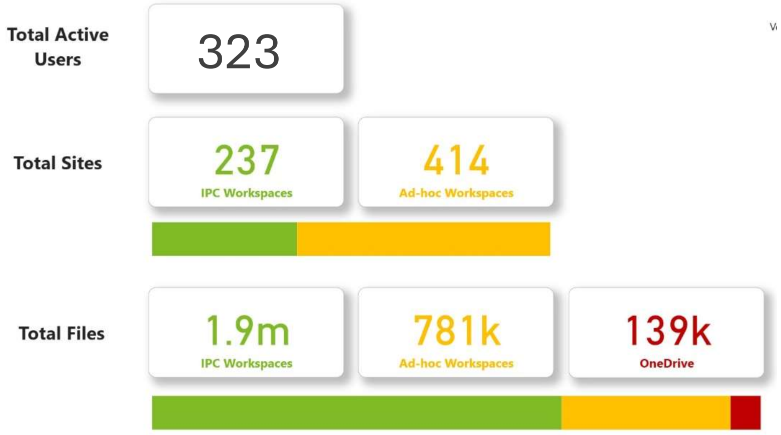
ORG: A ORG: B



Total Active Users: 346

Total Sites: 61 IPC Workspaces; 261 Ad-hoc Workspaces

Total Files: 774k IPC Workspaces; 952k Ad-hoc Workspaces; 3.7m OneDrive



Total Active Users: 323

Total Sites: 237 IPC Workspaces; 414 Ad-hoc Workspaces

Total Files: 1.9m IPC Workspaces; 781k Ad-hoc Workspaces; 139k OneDrive

High OneDrive usage

So higher risk that files with private info will be shared with others

Low and controlled OneDrive usage

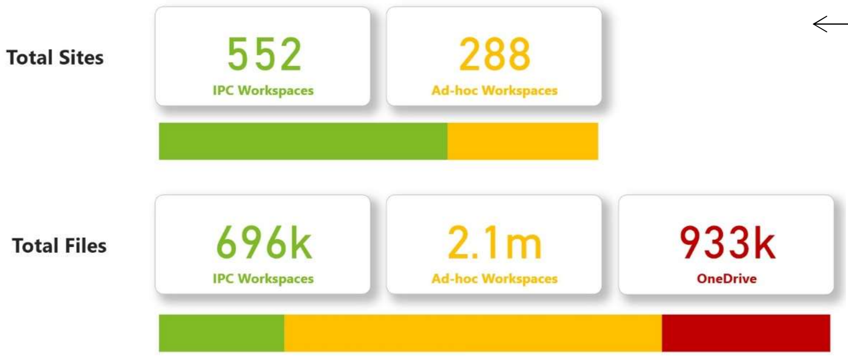
So much lower risk that files with private info will be shared with others

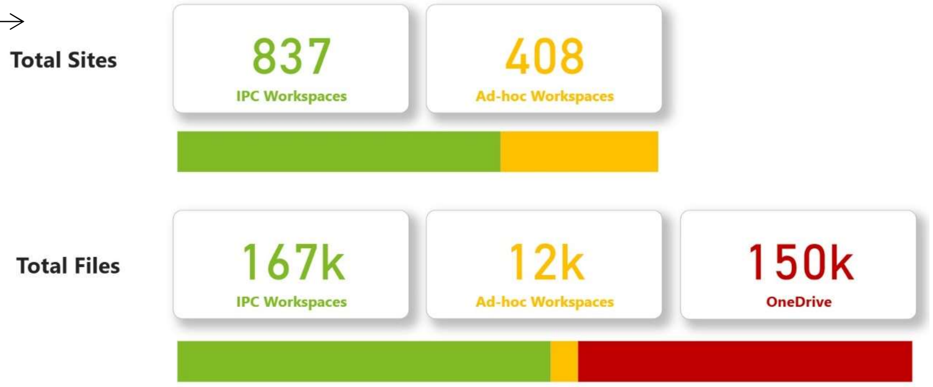
DASHBOARD ANALYTICS OF ONEDRIVE SILOS

Ad-hoc Teams a key risk

… and overuse is a sign that your digital workplace is not meeting user needs

ORG: A ORG: B



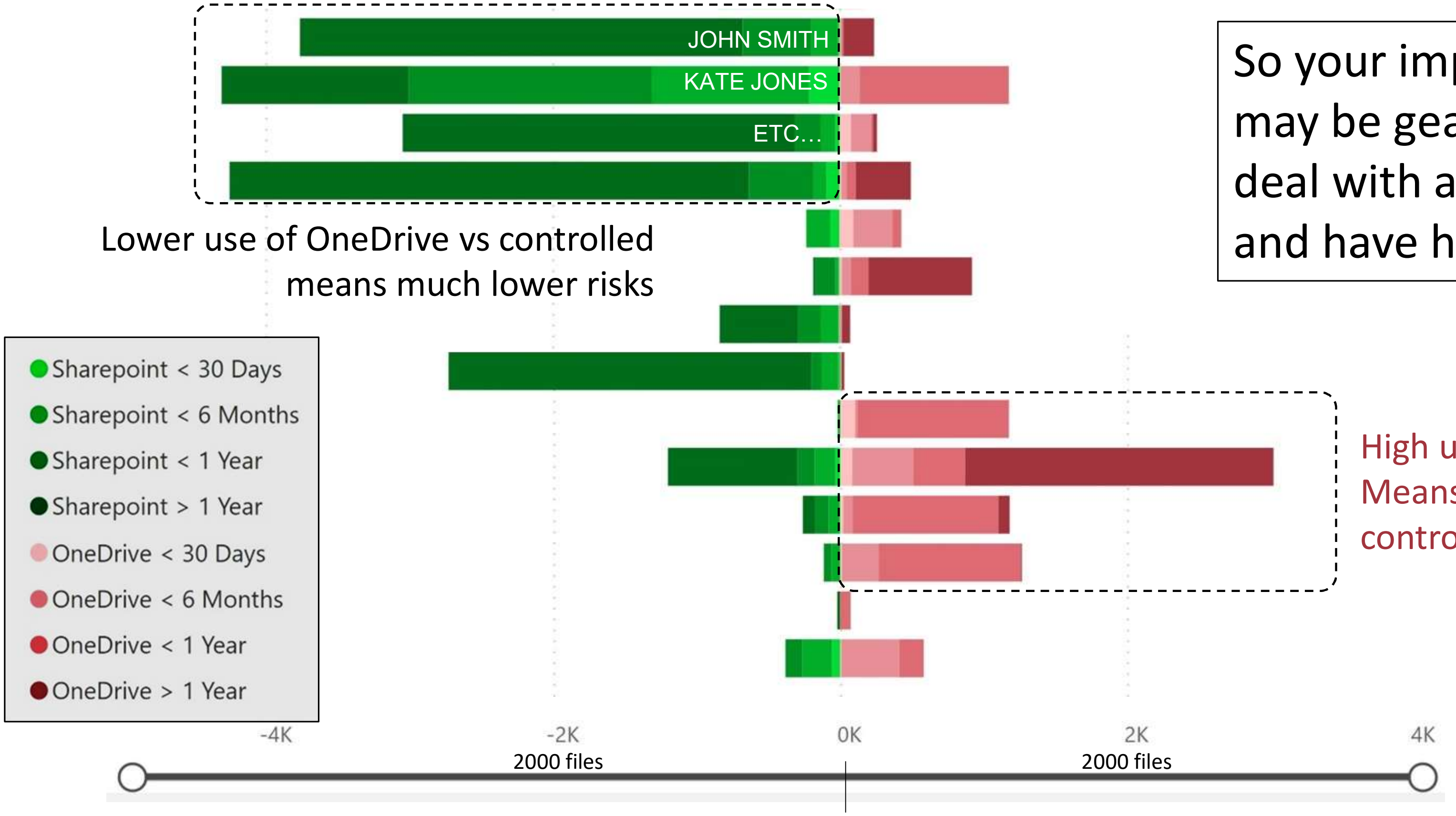


Most files going into ad-hoc Teams workspaces so higher risk of privacy info accidents

While a lot of ad-hoc workspaces most content going into structured workspaces

DASHBOARD ANALYTICS OF M365 USAGE

Controlled workspaces vs OneDrive use by person



Lower use of OneDrive vs controlled
means much lower risks

High use of OneDrive vs controlled
Means much harder to ID and
control files with private info

So your improvement efforts may be geared to people who deal with a lot of private info and have high OneDrive use

3. Bias to making work better

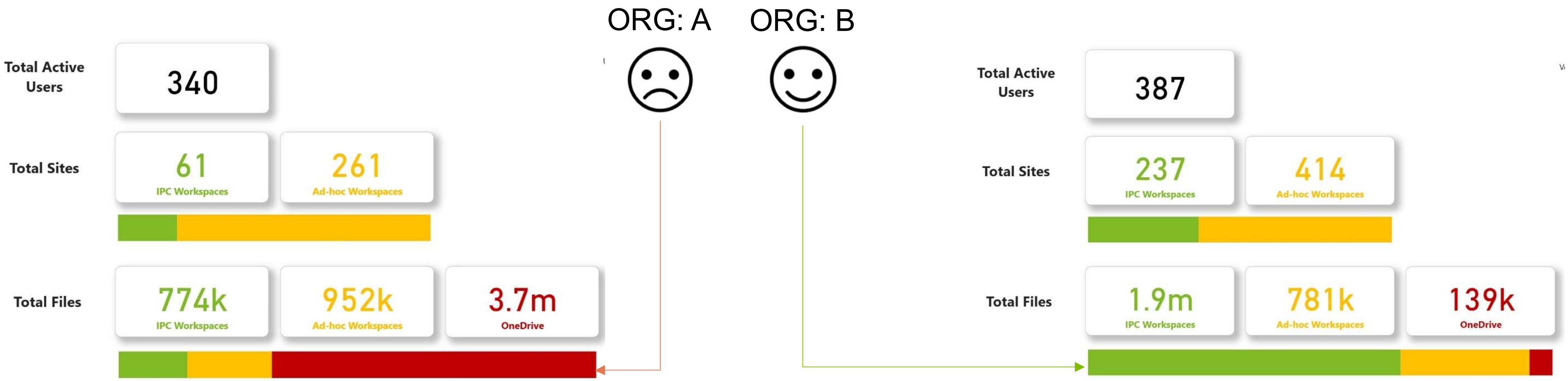
Making the right thing to do, the easiest thing to do

INFORMATION LEADERSHIP

let's make work better

DASHBOARD ANALYTICS EXAMPLE

Why the difference?



ORG: A

Total Active Users	340		
Total Sites	61 IPC Workspaces	261 Ad-hoc Workspaces	
Total Files	774k IPC Workspaces	952k Ad-hoc Workspaces	3.7m OneDrive

ORG: B

Total Active Users	387		
Total Sites	237 IPC Workspaces	414 Ad-hoc Workspaces	
Total Files	1.9m IPC Workspaces	781k Ad-hoc Workspaces	139k OneDrive

Why? Users perceive its quicker/easier to use OneDrive

- It might be (ouch!)
- Training
- Workspaces & folders don't meet their needs
- Email is default for some, that makes more people use it

People opt out when tools don’t work

If it’s too hard, people will find another way

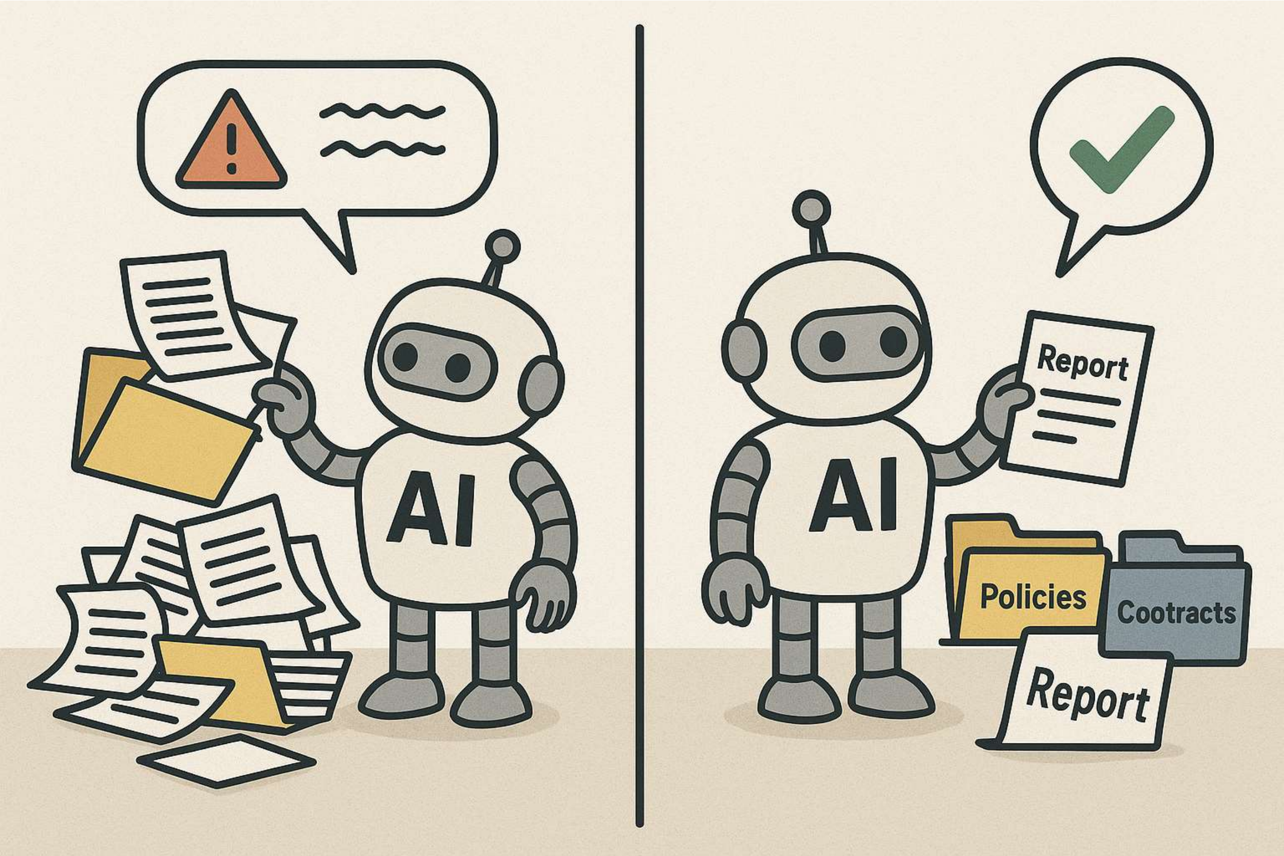
When tools feel clunky or slow, people will find ways around them. That is human nature. We have seen this again and again, especially with structured environments that are too complex, too slow, or just not matched to the task.

The more effort it takes to work the right way, the more likely people are to store sensitive information in the wrong place, or use tools like email or OneDrive just to keep things moving.

This is why good design and usability matter just as much as governance. If the structured workspace is easier than the workaround, people will use it.

Smart privacy still needs smart design

AI works best when content is structured, labelled and relevant



4. Bias to disposal

Most organisations hoard information

Most organisations hoard information.

This creates risk.

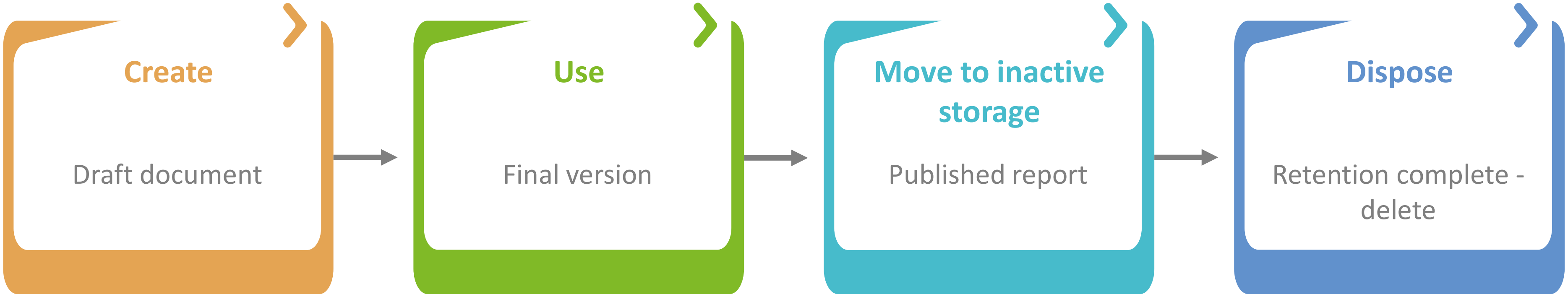
Let’s be honest. Most organisations are information hoarders. Not because people are careless, but because there is no clear line between what should be kept and what can be safely deleted.

We end up with multiple versions of documents, old reports, sensitive customer details, and draft content sitting around for years in inboxes, personal folders, and abandoned sites.

The problem is not just clutter. All of that content remains searchable, shareable, and potentially visible to AI tools.

Dispose of what’s no longer needed

Deliberate disposal reduces exposure to AI, audits and breaches



Create

Draft document

Use

Final version

Move to inactive storage

Published report

Dispose

Retention complete - delete

It is not about deleting everything. It is about applying retention schedules with confidence, knowing that the disposal is justified and documented.

A useful way to frame this is as a content lifecycle. Content is created, used, sometimes moved to inactive storage, and then disposed of according to approved rules.

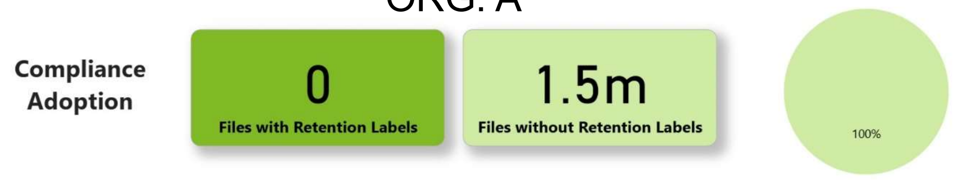
Doing this reduces what AI has access to by default, lowers your risk in audits and privacy investigations, and keeps information environments cleaner and easier to manage.

Disposal should be deliberate, not aggressive. When it is based on approved disposal authorities, it becomes a key part of your organisation's privacy, compliance, and operational strategy

CONTENT DISPOSAL

Bias to controlled disposal

ORG: A



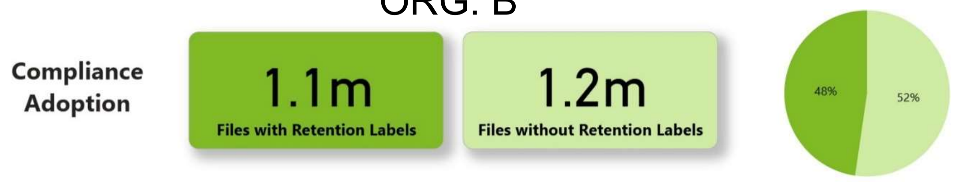
Compliance Adoption

0
Files with Retention Labels

1.5m
Files without Retention Labels

No controlled disposal
More chance of old content
being surfaced as they have

ORG: B

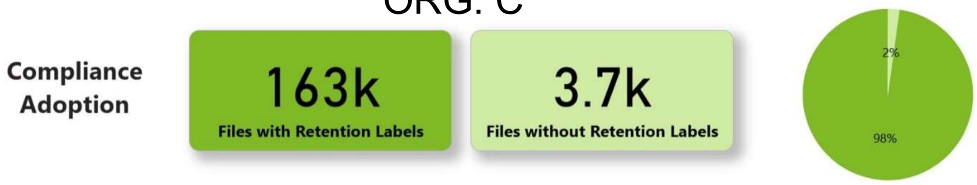


Compliance Adoption

1.1m
Files with Retention Labels

1.2m
Files without Retention Labels

ORG: C



Compliance Adoption

163k
Files with Retention Labels

3.7k
Files without Retention Labels

Safer with deliberate
disposal for most files

When you can’t dispose…

Redact or summarise

Restrict access

Apply labels

When you can’t dispose of it, contain it!

Sometimes, disposal is not possible. You might need to keep the content for legal, audit, or historical reasons. In those cases, the goal shifts from deleting to containing.

There are three main fallback options that help reduce risk when you cannot dispose.

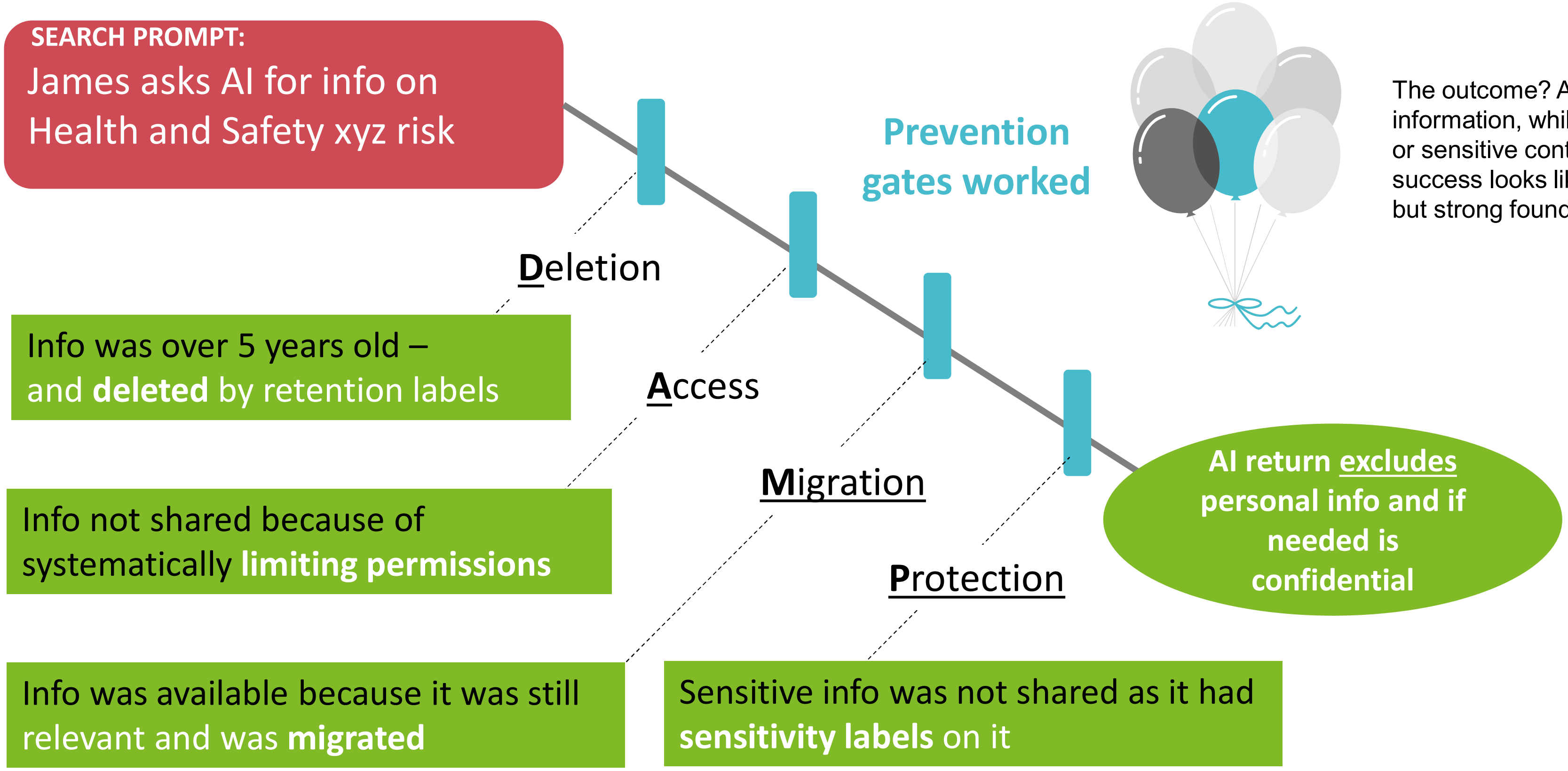
First, you can redact or summarise. This means removing sensitive parts or creating high-level summaries that do not expose private details.

Second, you can restrict access. Make sure only the right people and the right systems can see the content. That includes removing it from general AI access where possible.

And third, you can label the content clearly using metadata. Apply a sensitivity label or a 'Not for AI use' tag so that systems and users treat it with extra care.

None of these are as strong as proper disposal, but they are far better than doing nothing. They reduce exposure and give you a level of control when deletion is off the table.

Done right, risks reduced by better Information Management



SEARCH PROMPT:
James asks AI for info on Health and Safety xyz risk

Prevention gates worked

Deletion

Info was over 5 years old – and **deleted** by retention labels

Access

Info not shared because of systematically **limiting permissions**

Migration

Info was available because it was still relevant and was **migrated**

Protection

Sensitive info was not shared as it had **sensitivity labels** on it

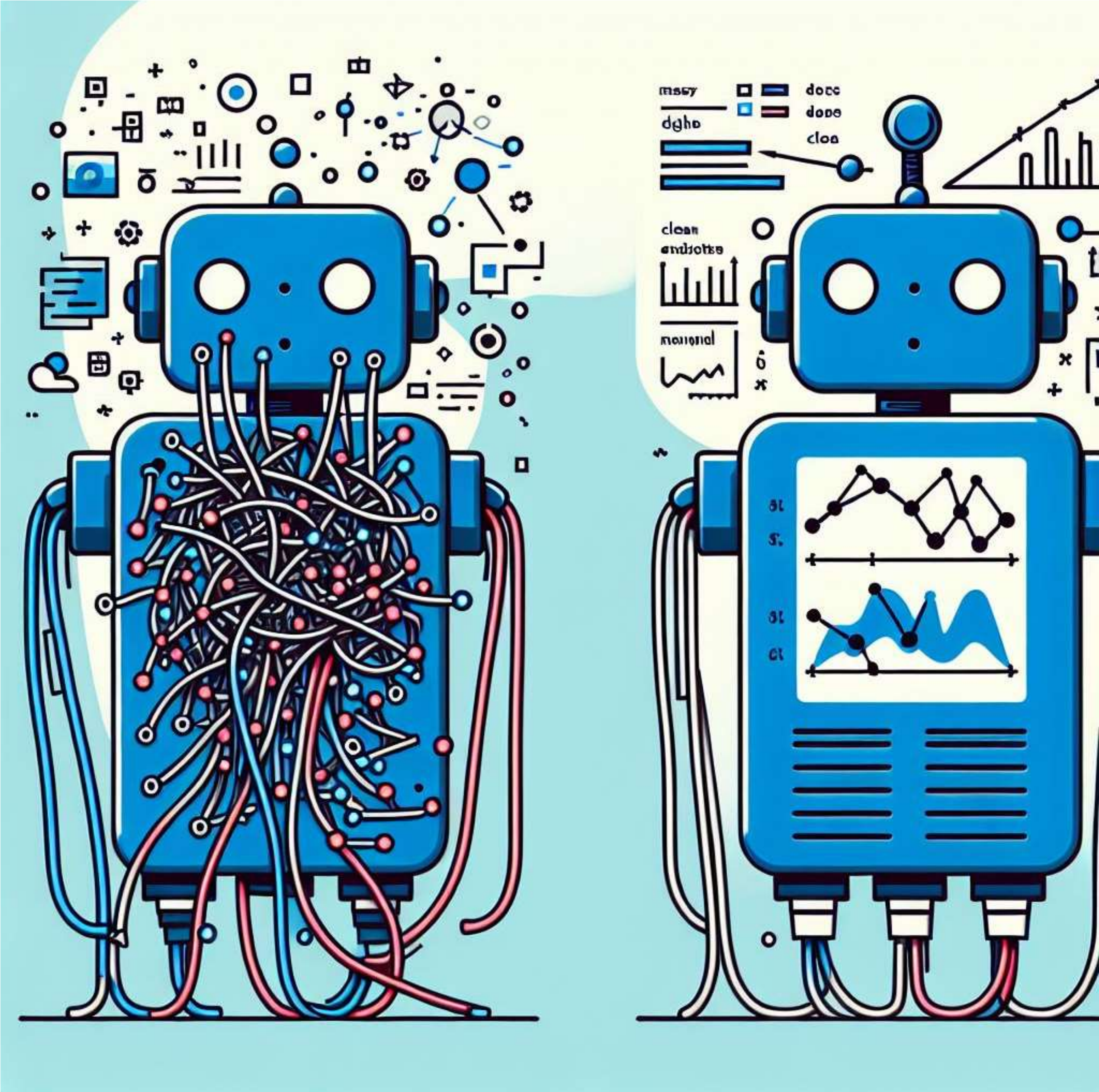
AI return excludes personal info and if needed is confidential

The outcome? AI returned useful information, while excluding personal or sensitive content. This is what success looks like - not just clever AI, but strong foundations underneath it.

5. Help AI help itself

AI Only Knows What You Feed It

Limiting AI inputs reduces privacy exposure



Limit the Content It Interacts With

OneDrives, Emails

Ad-hoc Teams/libraries

dirty

Almost always too risky
Private info and/or out of date or incomplete

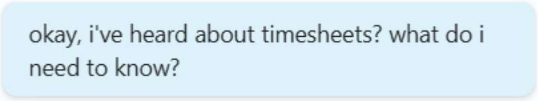
Structured Teams/libraries

clean

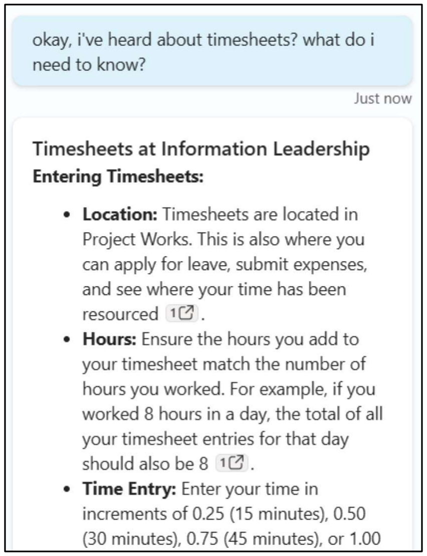
Safe if deliberately designed
Will auto use retention and protection labels to filter by
Labels applied based on confidentiality of files, folders, library
Filter by metadata, patterns within docs (eg NHIxxxx or custID)

Structured/curated
Eg policies, "working here",
governance meeting minutes

Ask AI



okay, i've heard about timesheets? what do i need to know?



okay, i've heard about timesheets? what do i need to know?

Just now

Timesheets at Information Leadership

Entering Timesheets:

- **Location:** Timesheets are located in Project Works. This is also where you can apply for leave, submit expenses, and see where your time has been resourced 1.
- **Hours:** Ensure the hours you add to your timesheet match the number of hours you worked. For example, if you worked 8 hours in a day, the total of all your timesheet entries for that day should also be 8 1.
- **Time Entry:** Enter your time in increments of 0.25 (15 minutes), 0.50 (30 minutes), 0.75 (45 minutes), or 1.00

AI will replace search,
be security + protection trimmed

Markers That Help AI Stay Private

Use Signals, Labels and Metadata

Train your AI on what **not** to touch

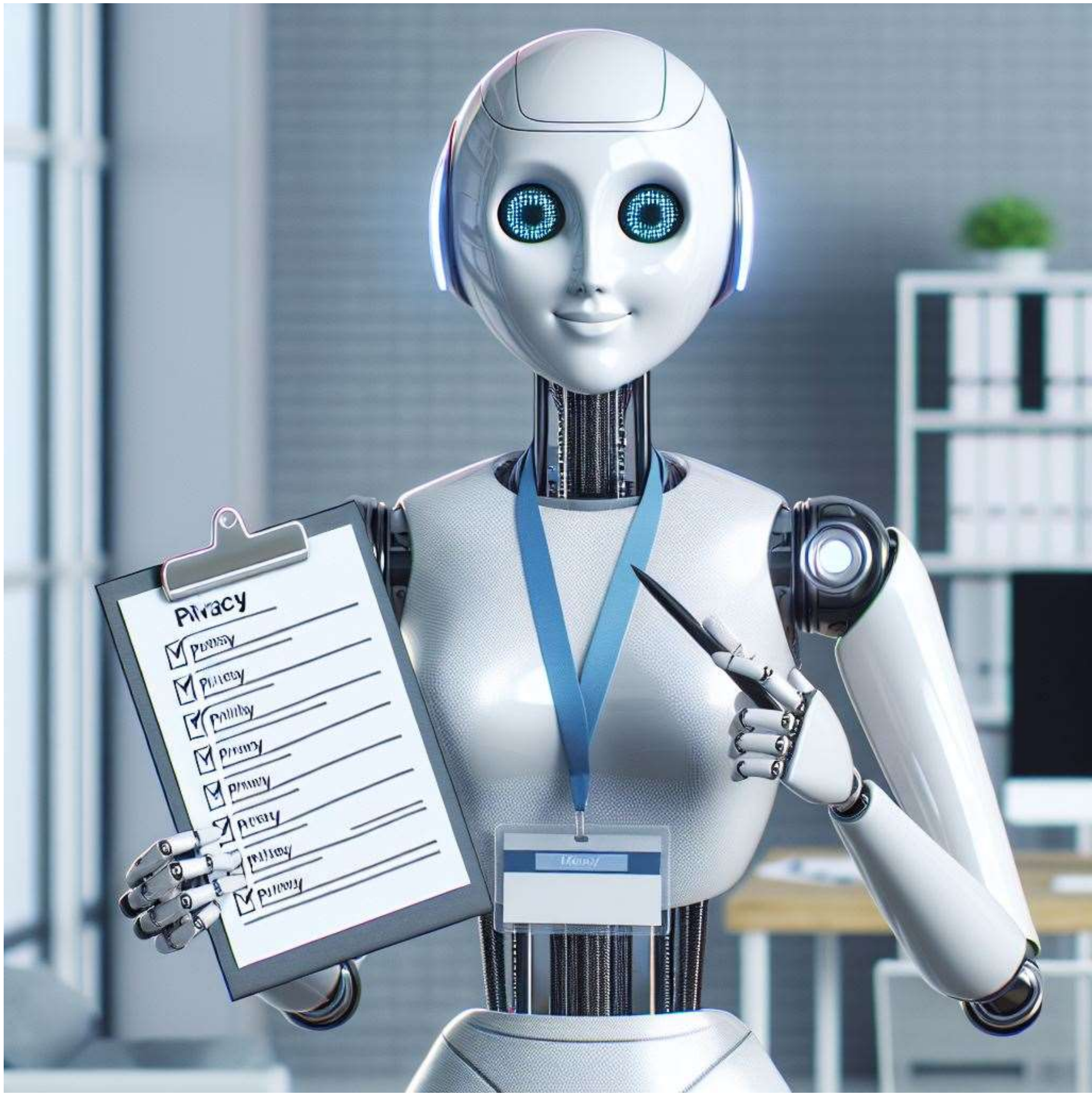
R: Restricted

X: Not for AI

P: Private

Help AI Help You

Smart defaults, restricted access, and markers let AI stay compliant without constant human babysitting



Let’s go recruiting…

High stakes content we all have to manage

INFORMATION
LEADERSHIP

let’s make
work better

RECRUITING CONTENT FOR UNSUCCESSFUL CANDIDATES

	Privacy?	Access By job	Info Protection	Retain for	Content stores
Advert/Job Spec				5 years	
Applicant Letters and CVs	Yes	Recruiters (int/ext) Key HR staff Key business staff		If declined delete after 1 year If successful add to personnel file	HR recruitment Team/workspace case file folder
Interviews	Yes				Email on for logistics
Reference checks Can include drug tests	Yes	Key HR staff Key business staff	High profile roles		
Process/decisions		Key HR staff Key business staff	High profile roles		Comms/docs emailed to doc library or copied to library
"Declined" comms	Yes	Recruiters (int/ext) Key HR staff Key business staff		12 mths	Email default deletion at 12 mths
Successful comms		unrestricted		If successful add to personnel file	

Risk challenges, being human

Email

The default for ext communication, easiest yet hardest to manage risks

Private info gets bundled in with everything else
Hard to separate and “destroy as soon as not needed”

M365 allows from folders and channels to have email addresses
- Get the best of both

OneDrive and “shared”

Worse than email and probably used alongside email!)
OneDrive best to have very limited space and defaults for disposal
“Shared” content is an info accident just waiting to happen as it can be external as well as internal

“Just in case”

We tend to hoard info

M365 example

What a structured approach looks like

Managed library – metadata, structure

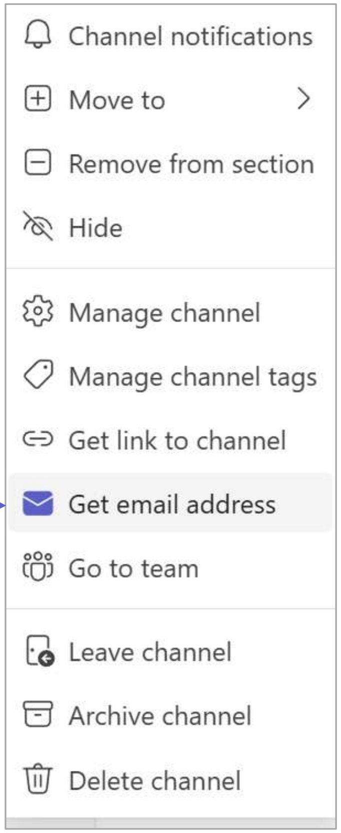
Folder per vacancy (case)

Folder per vacancy…allows

- Access permissions
- Protection settings
- Retention on close

Status of recruitment (open, closed)

Email address for the folder



Having a method of moving content to personnel file for successful then deleting the folder

Manage your digital workspaces

Expect people to use structured workspaces and monitor
Ongoing programme to monitor and reduce OneDrive use and email for high stakes content

Separate treatment and design for “High stakes” content vs BAU”

Most privacy info at high risk of being shared is in case files
e.g.,

- Recruitment (by role)
- Performance reviews (by person by period)
- Disciplinary (by person by incident)

Summary

INFORMATION LEADERSHIP

let's make work better

SUMMARY

Privacy First: The Foundation for Safe and Useful AI

To protect privacy, start with the information:

- **Recognise** private information in daily work
- **Use** structured, shared, well-managed spaces
- **Guide** people by making the right way easy
- **Remove** what is no longer needed, safely
- **Control** what AI can see and use

Stronger privacy is not a separate project – it’s the outcome of better information habits. Get that right, and AI works with you, not against you.

Find out more

Linked In https://www.linkedin.com/in/sarahheal/

I post regularly on info mgmt
+ making work better

Website https://www.informationleadership.com/

What we do and case studies

Public + in-house workshops

On all aspects of making digital workplaces better + safer
Contact me for more details. Next workshops in June…
sarah@informationleadership.com

INFORMATION
LEADERSHIP

let's make
work better

let’s make work better

INFORMATION
LEADERSHIP