



Få styr på persondata, før du giver AI adgang

Data Loss Prevention & Deletions (DLP-D)

- Du kan ikke *lække* data, der ikke er *der*.

SOHO – 8 oktober 2026



DAGENS PROGRAM

09.00 Velkomst og ramme: hvorfor privatliv ændrer sig med AI

09.10 Klassificering, privatliv og sikkerhed - **v. Sophia Laura von Ferber & Mikkel Nielsson**

09.50 DLP – **v. Kristian Boe Helweg Hansen**

10.30 Pause

10.45 Copilot – **v. Kristian Billeskov Bøving**

11.25 Opsamling/Afrunding

11.30 Tak for i dag

Sådan klassificerer Data & More data

Agenda



01

Nålen i høstakken



02

**Data&More's
klassifikation af
persondata**



03

**Sikkerheds-
klassificering**

Sophia von Ferber

Titel: Data Classification Specialist, Data & More
From Germany - we take data privacy very seriously 😊

Academic Background:

- MSc in Management of Innovation and Business Development at CBS

Data & More:

- Data Classification Specialist at Data & More since January 2023
- First as a student assistant, now full time >2 years
- Managing German PoC projects
- Primarily responsible for German data classification

Freetime:

- Skiing, Snowboarding, Kitesurfing, Bouldering

Location:

- Copenhagen, Denmark



EMNE 01 / 03

01

Nålen i høstakken



● 01 Nålen i høstakken

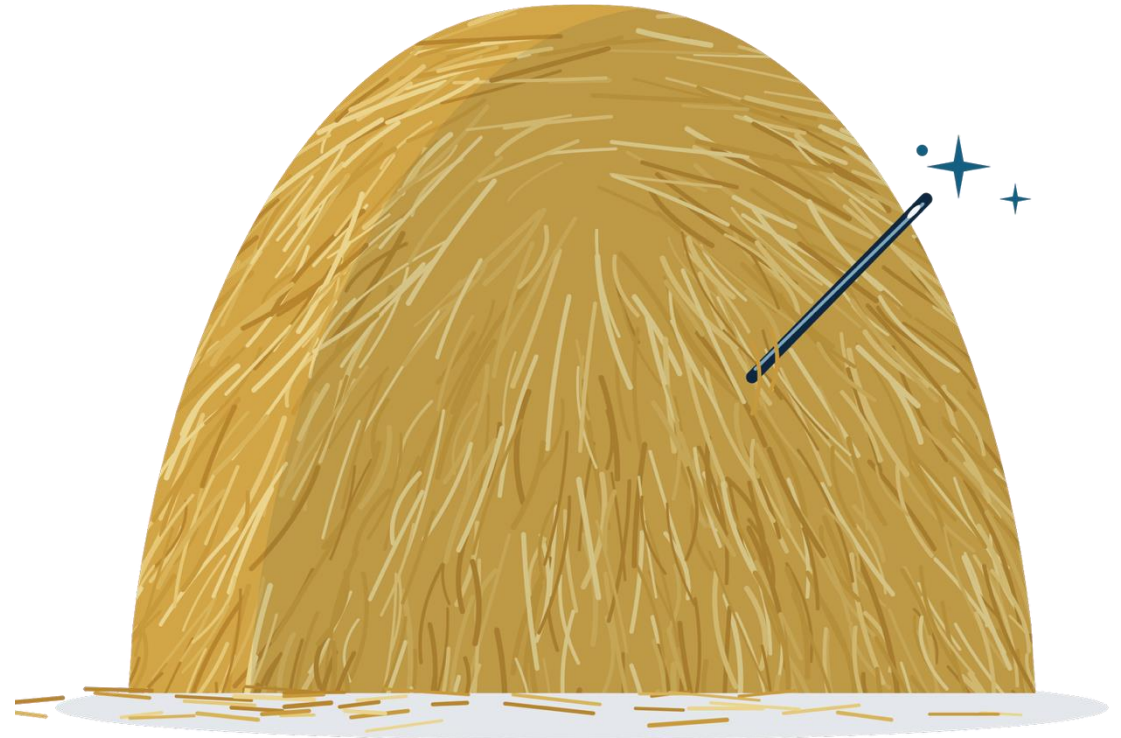
○ 02 Data&More's klassifikation af persondata

○ 03 Sikkerhedsklassificering

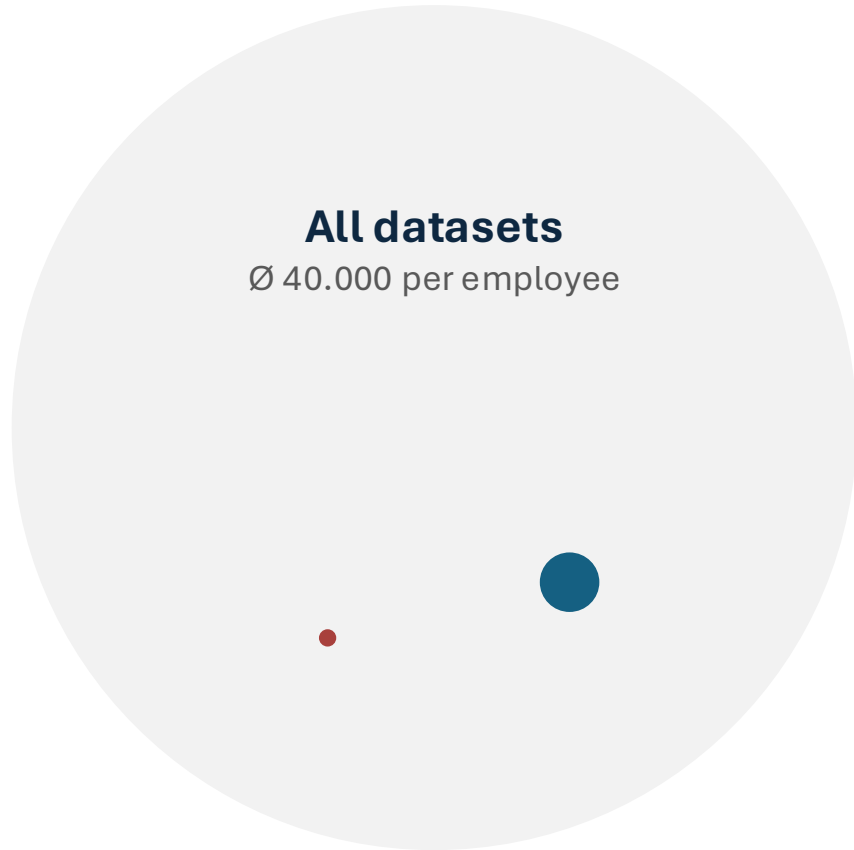
The Needle in a haystack

**Have you ever tried
finding a needle in a
haystack?**

We have. Every single day.



The Needle in a haystack



○ **40.000**

datasets per employee on average

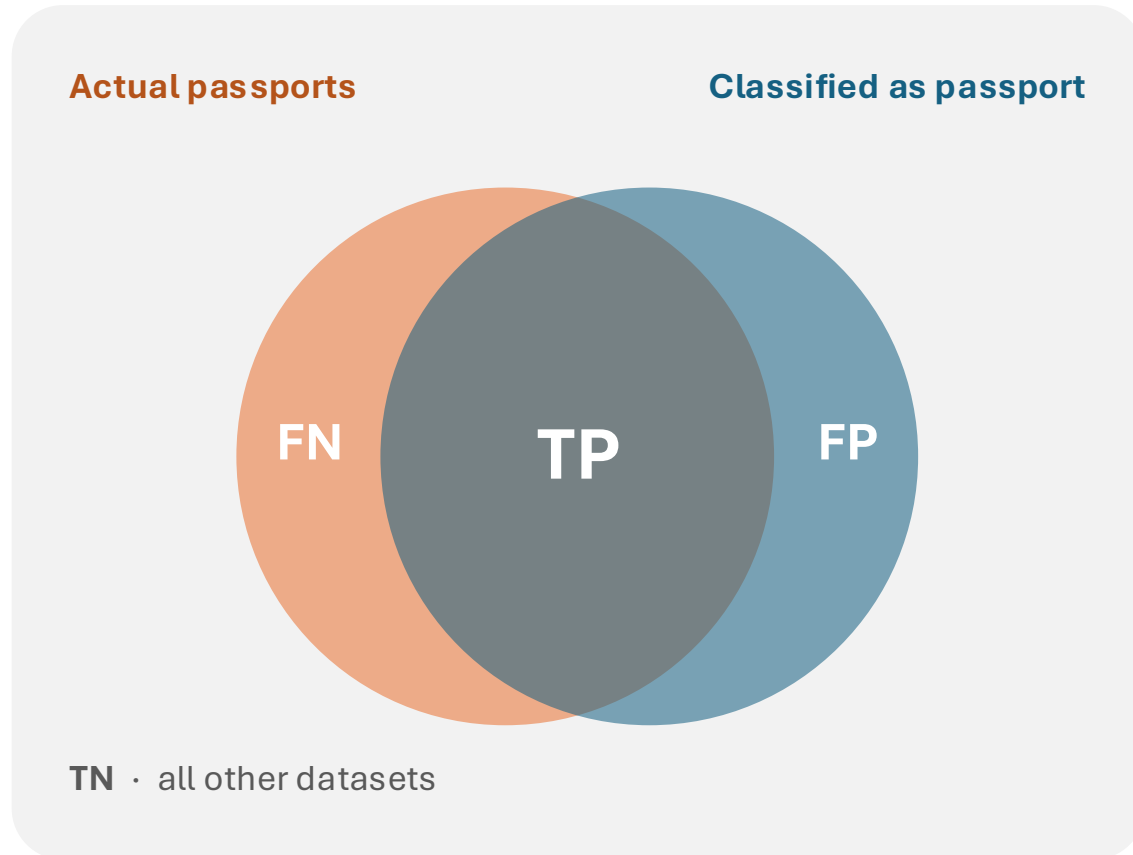
● **0,5%**

privacy data · $\approx$ **200**
datasets per employee

● **0,04%**

security data · $\approx$ **16**
datasets per employee

The Needle in a haystack



TP

True positives

Passports correctly classified as passports

FP

False positives

Flagged as a passport, but isn't one

< 5%

FN

False negatives

Passports we missed

as low as possible

TN

True negatives

Other datasets, correctly left alone

The Needle in a haystack

01

95%

True Positives

02

1 in 200

Privacy data

→ rare-event classification

03

20.000

documents must be
classified to find 100
privacy documents

04

5

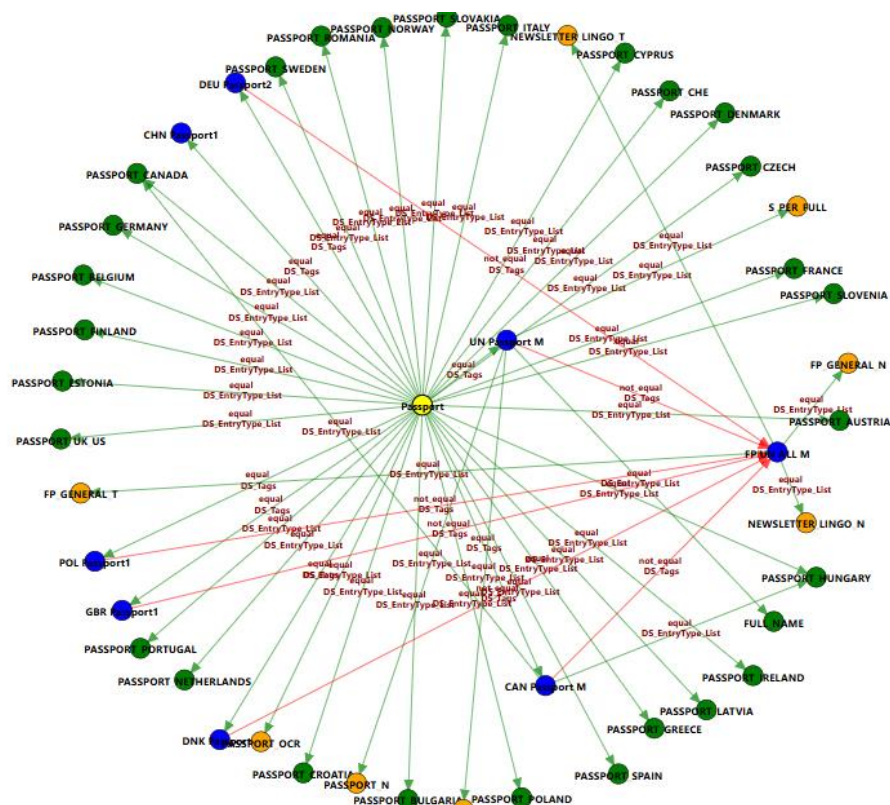
false positives allowed
across those 20.000
documents to stay at 95%

99,975%

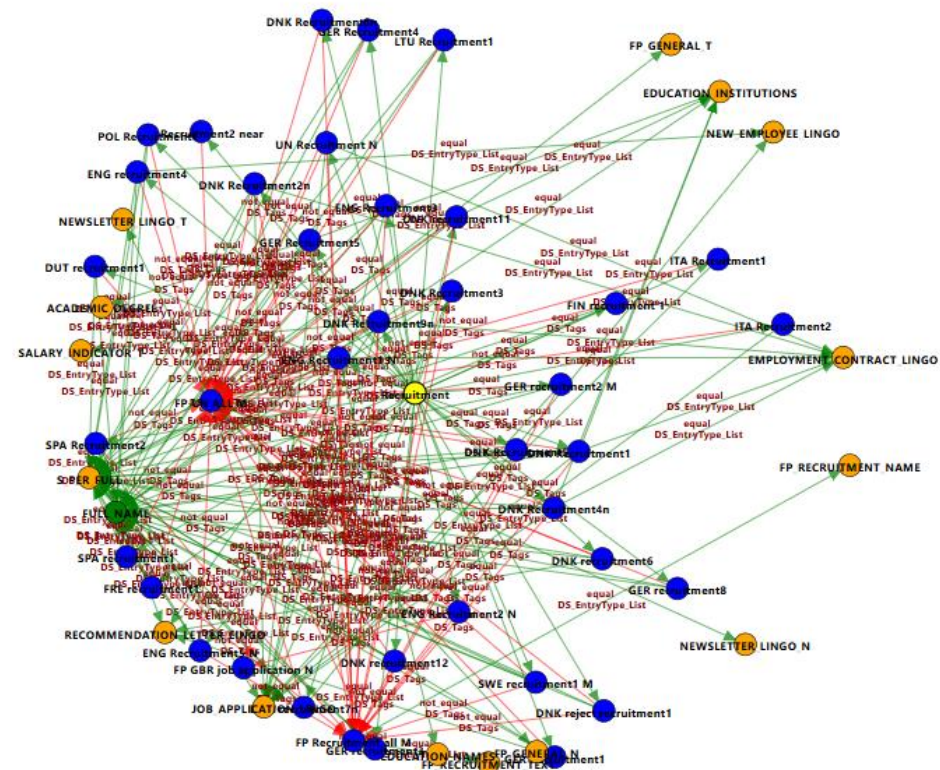
How often each individual classification has to
be correct.

What this looks like in practice

Passport



Recruitment



Mikkel Eidset Nielsson

Titel: Project and Data compliance Lead Specialist

Baggrund:

- Cand.ling.merc (Tysk som fremmedsprog)
- Har erfaring med kundeimplementeringer i IT-branchen i 9+ år

Data&More:

- Ansat siden 2022
- Kundeimplementeringer + support i Danmark og Tyskland
- En del af Data&More's klassifikations-team
 - Hovedansvarlig for dansk data-klassifikation

Fritid:

- Fodbold, løb, klaver

Privat:

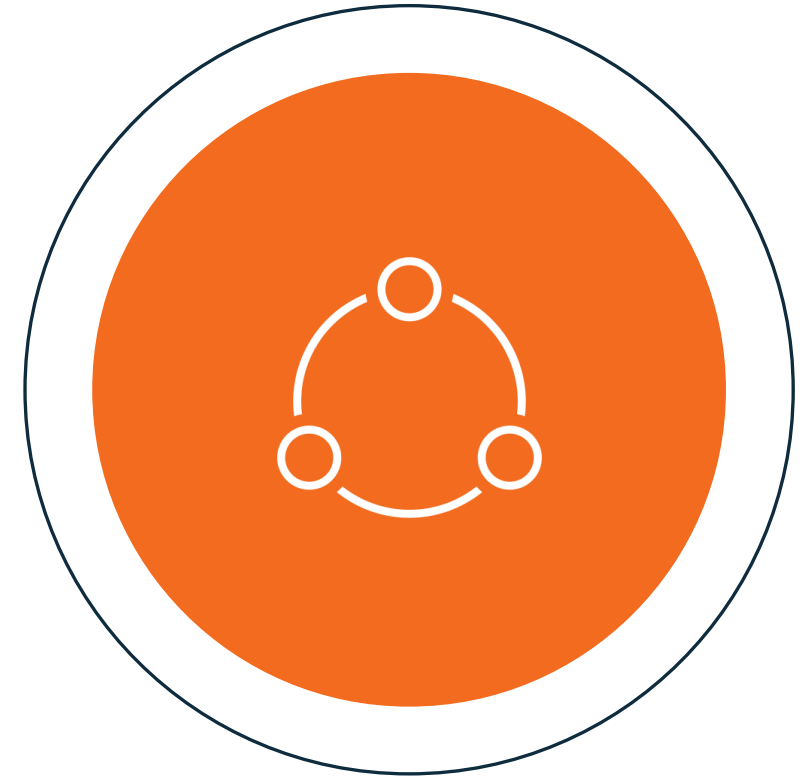
- Bor på Frederiksberg med kone og søn på 4 år



EMNE 02 / 03

02

Data&More's klassifikation af persondata



○ 01 Nålen i høstakken

● 02 Data&More's klassifikation af persondata

○ 03 Sikkerhedsklassificering



MØD BENTE

Bente har rigeligt at lave i forvejen

Når hun rydder op med Data&More, skal hun kun se de data, der kræver en beslutning. Ikke endnu en bunke.

Det, hun overser, kan Copilot stadig finde.

Billedet er AI-genereret

Fejler klassificeringen, virker beskyttelsen mod AI ikke

S SharePoint › Økonomi › Dokumenter › Fakturaer

fakturaoversigt.xlsx			
	A	B	C
1	Fakturanr.	Kunde	Beløb
2	2409871133	Nord A/S	48.750 kr.
3	0103920457	Hansen & Co	12.300 kr.
4	1511784026	Fjordtrans	7.980 kr.
5	3006883391	Lyngby	23.415 kr.

S SharePoint › HR › Dokumenter › Medarbejdere

medarbejdere.xlsx			
	A	B	C
1	Navn	CPR-nummer	Afdeling
2	Anne Jensen	011280-9392	Salg
3	Mikkel Holm	010691-5117	Økonomi
4	Sofie Nielsen	102110-0838	HR
5	Lars Madsen	010661-2005	IT

Det er filen på Bentes HR-SharePoint-site med rigtige CPR-numre, der skal findes og beskyttes mod AI, så medarbejdernes persondata ikke ender i Copilot.

Sådan genkender vi et CPR-nummer

Selve nummeret er ikke nok. **Ordene ved siden af nummeret afgør**, om det er et CPR-nummer.

Hvad kunne dette nummer være?

0107729995

Også som 010772-9995 · 01 07 72 9995 · 01.07.72-9995 ...

Startord

= CPR-nummer

CPR-nummer 0107729995

Andre startord, der peger på CPR:

CPR	cpr.nr
CPR-nr	cpr#
CPR-numre	Cprnr
Cpr:	cprnummer
cpr:nr:	

Stopord

= Ikke CPR-nummer

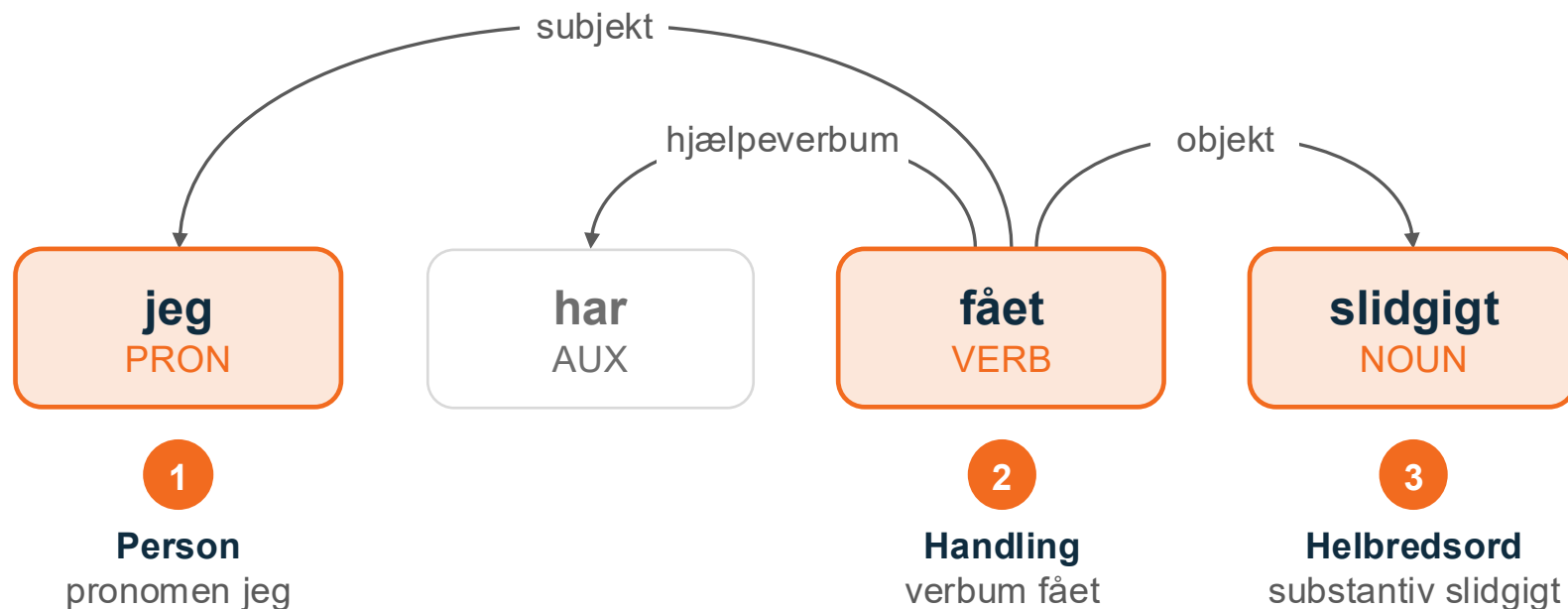
Fakturanummer 0107729995

Andre stopord, der udelukker CPR:

Ordrenummer	Telephone:
EAN-nummer	Forsendelsesnummer
Billetnummer	Produktkode
Telefonnummer	Produktnummer

Grammatikbaseret sætningsanalyse

Eksempel: "Jeg har fået slidgigt"



Sådan virker det

- Vi læser **grammatikken**, ikke kun nøgleord
- Markeres kun, når **alle tre led** er opfyldt
- Virker på flere sprog, såsom **dansk, engelsk og tysk**
- Færre oversete fund og falske positive

Ordbog over sygdomme og symptomer

skoldkopper · sklerose · **slidgigt** · sukkersyge + *mange flere*



Fejler klassificeringen, virker beskyttelsen mod AI ikke

OneDrive › Bente Nielsen › Dokumenter › Studie

Bacheloropgave_sygepleje.docx

3.2 Analyse af casen

I casen møder vi en fiktiv borger på 58 år, som fortæller:

»Jeg har fået slidgigt i knæet.«

Udsagnet bruges til at analysere, hvordan borgere beskriver kroniske smerter over for sygeplejerske.



OneDrive › Bente Nielsen › Dokumenter › Sundhed

Journaludskrift_Karen_Holm.pdf

Journalnotat · Ortopædkirurgisk ambulatorium

Patient: Karen Holm

Dato: 14.03.2026

Patienten oplyser: »Jeg har fået slidgigt i knæet.«

Røntgen bekræfter artrose i højre knæ. Henvises til fysioterapi.



Det er lægejournalen på Bentes OneDrive med helbredsoplysninger om en navngiven patient, der skal findes og beskyttes mod AI, så følsomme persondata ikke ender i Copilot.

Frasortering af falske positive

Forskellige metoder kan bruges som løsning



Nyhedsbrev

Ariana Grande trækker sig fra musical

Fans har udtrykt bekymring for popstjernens helbred og spekuleret i, om hun lider af anoreksi.

[Klik her for at læse nyhedsbrevet](#)



Løsning: fast sætning

Nyhedsbreve med "Klik her for at læse nyhedsbrevet" frasorteres



GDPR-vejledning.pptx

Hvad er helbredsoplysninger?

Helbredsoplysninger er følsomme personoplysninger efter artikel 9.

Et eksempel er udsagnet "Hun er sygemeldt med stress".



Løsning: filtype

PPTX-filer frasorteres for helbredsoplysninger

DATA&MORE'S KLASSIFIKATION AF PERSONDATA | FORBEDRINGSPROCESSEN

Kampen for at ramme rigtigt

Dårlige scanninger, numre der ligner hinanden, regler der skifter fra land til land og tekster, der kun ligner persondata. Her er et lille uddrag af det, vi skal ramme rigtigt hver dag.

Medarbejderbonus eller Coop-bonus?

Personnummer i Polen

Kørekort i dårlig billedkvalitet

Sygemelding fotograferet skævt med mobilen

Fiktive personer i bøger

Lønsedler i Tyskland

Nyhedsbreve i forskellige formater

Opsigelse af lejeaftale eller medarbejder?

ID-kort scannet på hovedet

Landespecifik lovgivning

Overenskomst eller ansættelsesoplysning?

Rekrutteringsdata i Kroatien

Hunden Hans' helbred

Mødelokale eller rejsebooking?

Vejledninger i forskellige formater

Seksualitet i skønlitteratur

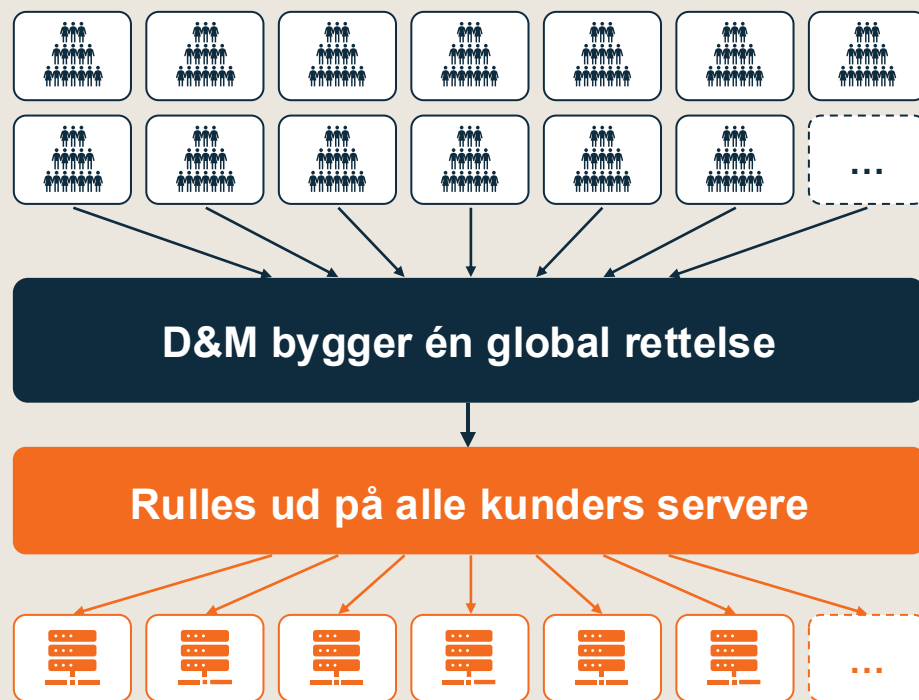
Guiden til det gode CV

Personlige attester i hvert EU-land



Jeres feedback, alles gevinst

Global rettelse, bygget på alle kunder



KUNDEFEEBACK

Kundefeedback fra mere end 175.000 brugere er en afgørende drivkraft:

Vi har mange gange fået at vide fra jer, at vi er forkert på den. Vi ville ikke være hvor vi er i dag uden alt det input vi får fra jer.

EMNE 03 / 03

03

Sikkerhedsklassificering



○ 01 Nålen i høstakken

○ 02 Data&More's klassifikation af persondata

● 03 Sikkerhedsklassificering

THE SHIFT

AI is changing the rules of the game.

Until now: Protect the network!

- The data sits behind a defined perimeter.
- Data inside the network is implicitly trusted.
- Risks move at human speed — phishing, misconfigurations, lost laptops.
- The SOC team can detect most events before damage occurs.

From now on: Protect the data!

- Company data is now the attack surface.
- Agents act autonomously across systems — they combine APIs, tools and multi-step workflows.
- AI inherits the access of human actors, but at machine speed.
- The SOC team loses oversight — millions of events per minute.

The conclusion: Data is no longer at risk because attackers are looking for it — but because it sits where an AI can reach it at any time.



PERSONAL DATA

Found in emails, Fileshares, OneDrive & SharePoint.

Name	Description of the document class
Payment card	Payment-card data; specifically searched for credit cards via number logic.
Identity documents	Various identity documents used to identify a person.
Driver's license	Driver's-license data incl. unique numbers and photo of the person.
Ethnic origin	Indications of ethnic affiliation or country of origin.
Grant application	Personal data in applications for financial funding.
Health card	Health-insurance cards, mostly recognized via OCR scanning.
Health information	Sick notes, diagnoses and medical details.
Union membership	Indications of membership in a union.
National ID number	National identification numbers (also via OCR from image/PDF).
National ID card	National ID cards of identifiable persons from various countries.
Passport	Passports with photo of the person and unique country codes.
Political orientation	Party membership or political orientation.
Recruitment	Applications, speculative applications, CVs and rejections.
Religious belief	Religious affiliation and membership in recognized churches.
Salary & financial data	Payslips, bonuses and fees.
Sexual orientation	Details about a person's sexual orientation.
Tax information	Tax data, especially annual statements (often as PDF).



CRITICAL SECURITY INFORMATION

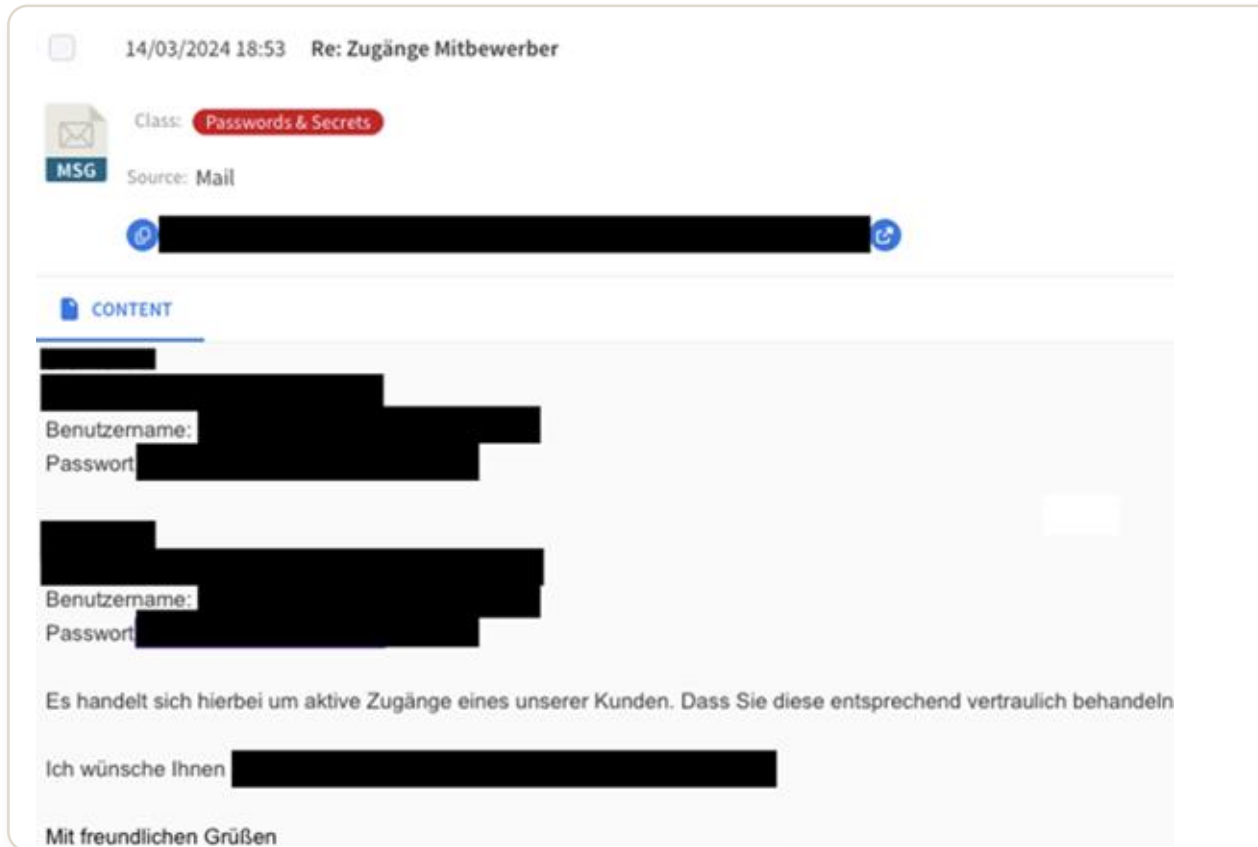
Found in emails, network drives, OneDrive & SharePoint.

Class	Description
Passwords & secrets	Credentials and keys for system and machine-to-machine access.
Source code	Source code that enables IP theft and vulnerability analysis.
Log files	Log files from systems and servers.
Infrastructure configuration	Configuration data incl. automation (e.g. Ansible).
Vulnerability assessment	CVE assessments and penetration-test results.
Security incidents	Reports that describe and assess security incidents.
Digital certificates	Certificates for authentication & encryption (pki, pem, cert).
Physical Security	Information on the location of surveillance cameras.
Security requirements	Security requirements for external service providers.
Network access control	Firewall rules, VPN settings and network segmentation that define who can reach which systems.



CRITICAL SECURITY INFORMATION: EXAMPLE

Passwords & secrets



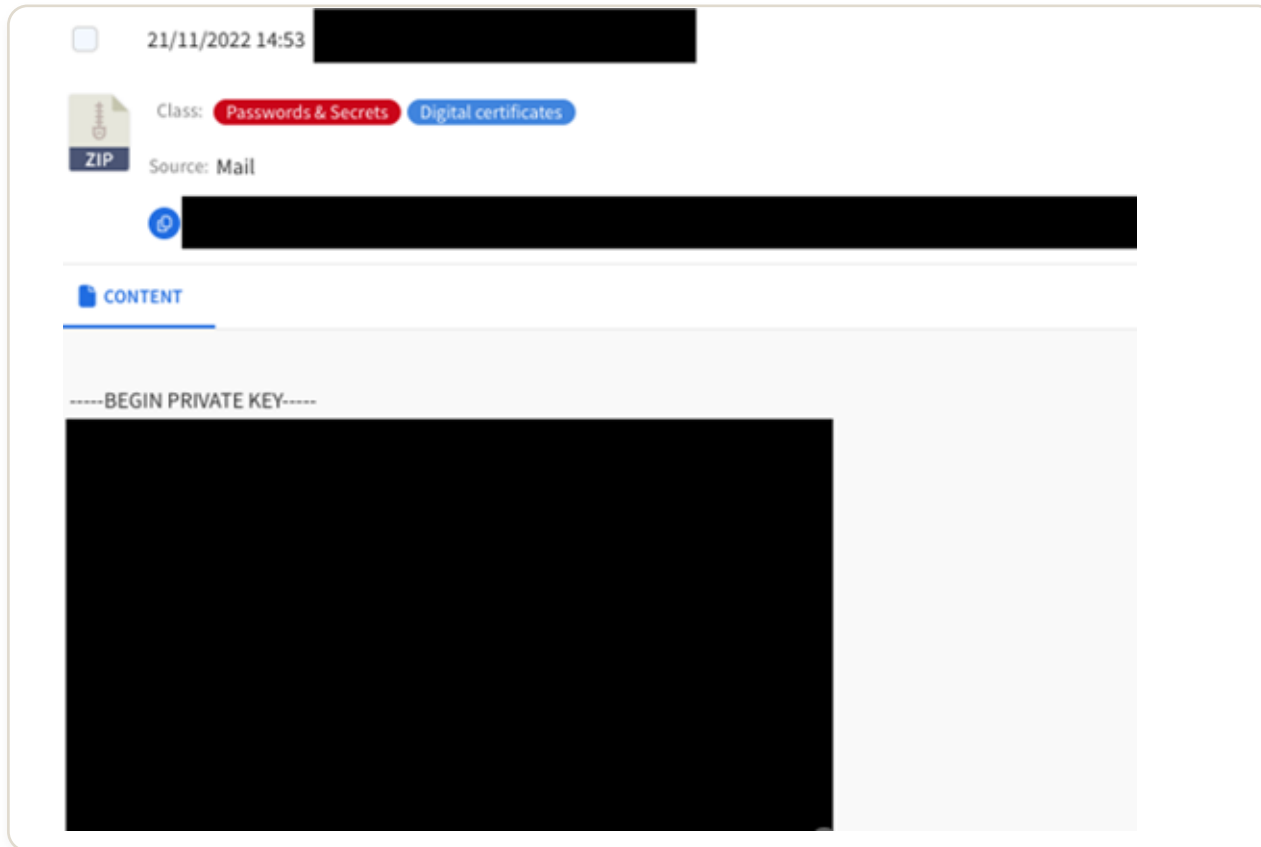
WHY IT MATTERS

- An obvious risk in the wrong hands – or in an AI
- Once in Copilot's index, users can't delete them
- **We also find “Secret Values” – the master key to all of M365**



CRITICAL SECURITY INFORMATION: EXAMPLE

Private Keys

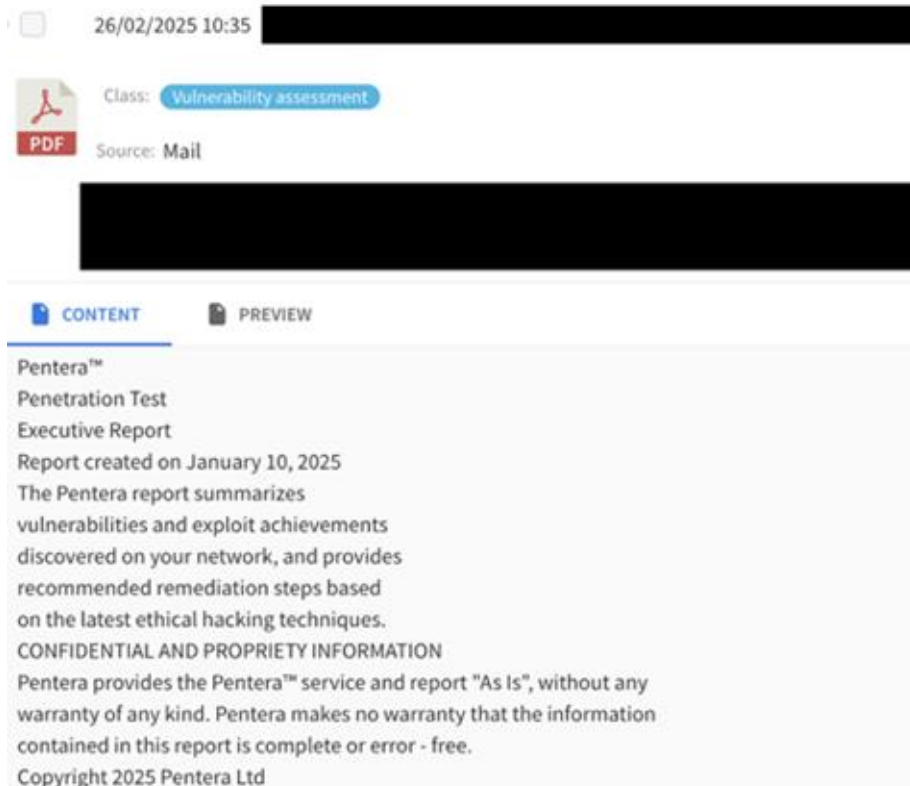


WHY IT MATTERS

- A free pass for identity theft
- Anyone holding it can impersonate the owner
- **Encrypted messages can no longer be trusted**

CRITICAL SECURITY INFORMATION: EXAMPLE

Penetration Tests



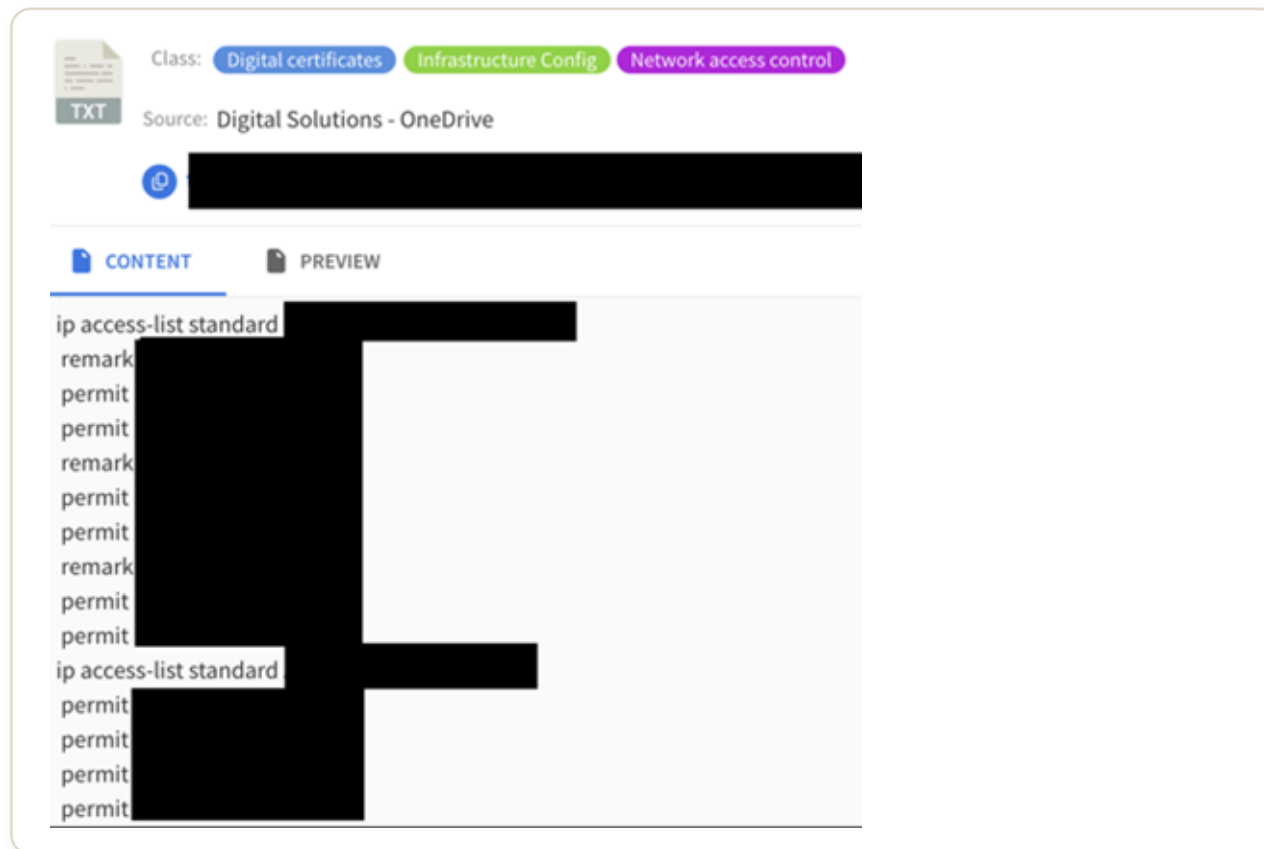
WHY IT MATTERS

- Every system has gaps
- A pentest report lists them all
- **For an attacker: a ticket to the next break-in**



CRITICAL SECURITY INFORMATION: EXAMPLE

Network Access Control



The screenshot shows a document viewer interface. At the top, there are three colored tags: "Digital certificates" (blue), "Infrastructure Config" (green), and "Network access control" (purple). Below these, it says "Class: Digital certificates" and "Source: Digital Solutions - OneDrive". A large black redaction box covers the top part of the document content. Below the redaction, there are two tabs: "CONTENT" (selected) and "PREVIEW". The "CONTENT" tab shows a list of network configuration commands, with several lines redacted by black boxes. The visible text includes:

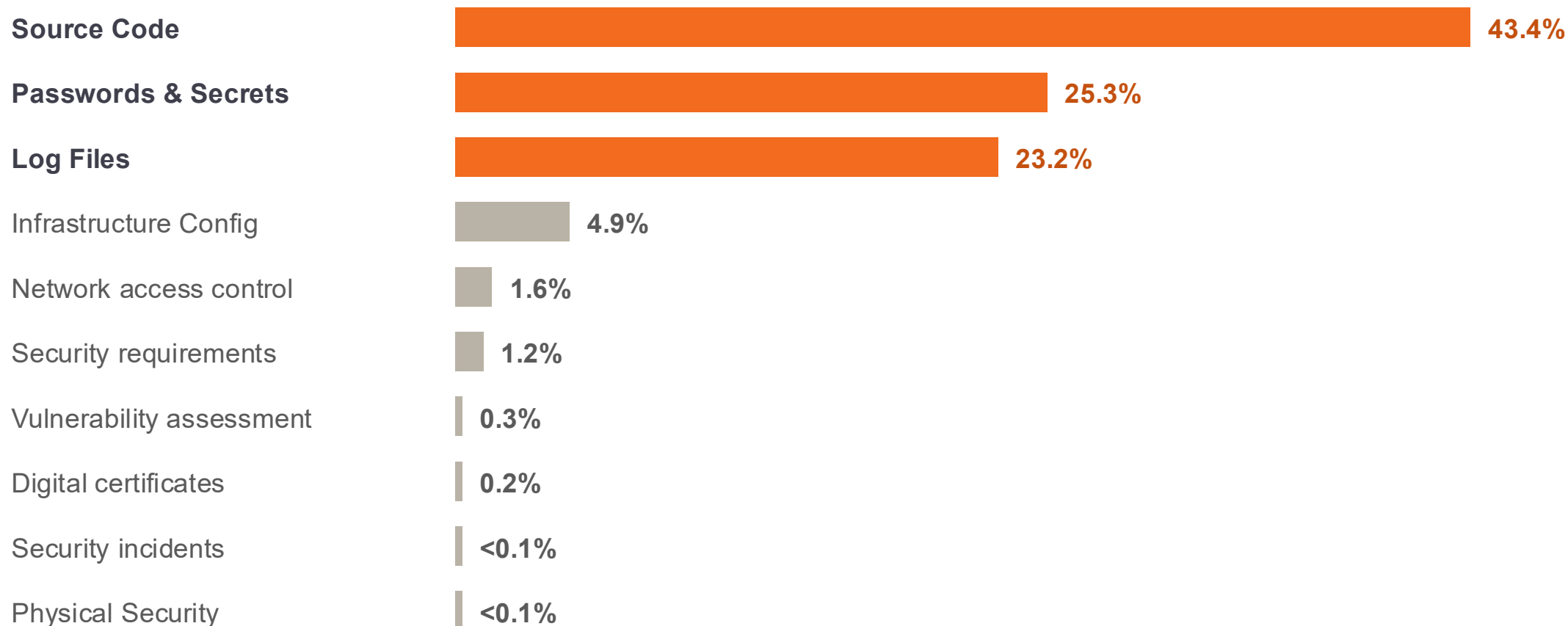
```
ip access-list standard
remark
permit
permit
remark
permit
permit
remark
permit
permit
ip access-list standard
permit
permit
permit
permit
```

WHY IT MATTERS

- Firewalls and VPNs are the heart of security
- Their configuration maps the whole network
- **For an attacker: a ticket to the entire network**

SHARE OF 1,644,149 SECURITY INSTANCES ACROSS ~ 160 SCANS

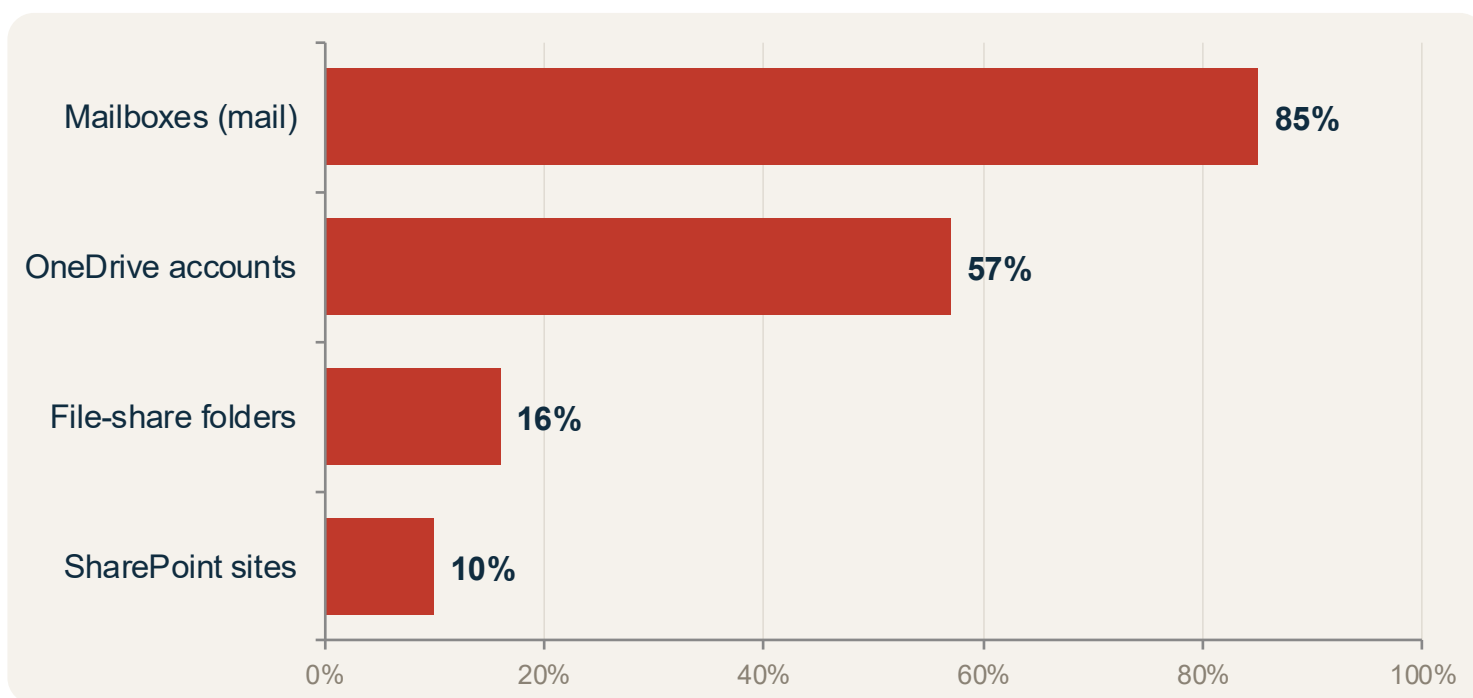
Share of security instances found, by category



AI-READINESS-CHECK

Copilot knows the passwords¹⁾

Average share of accounts, sites and folders that contain non-compliant data.



PER EMPLOYEE

16

**emails with security
information in plain text, on
average**

¹⁾On average, 16 emails with security information in plain text are found per employee.



Kristian Boe Helweg Hansen

Titel: Customer Success & Compliance Manager

Baggrund:

- Produktions Ingeniør + HD IT & Økonomistyring
- Har erfaring med kundeprojekter og været i IT-branchen i +25 år
- Certificeret DPO

Data&More:

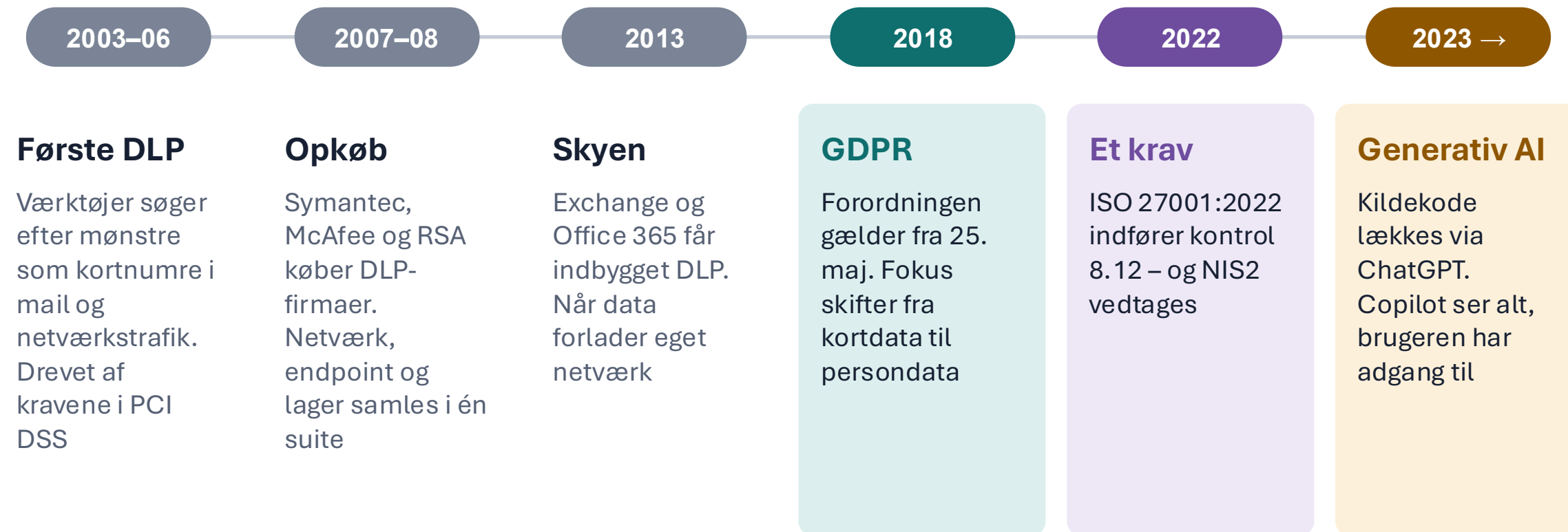
- Ansat siden 2020
- Kundeimplementeringer i Danmark og Sverige
- Hovedansvarlig for Customer Success + D&M Compliance Manager

Fritid:

- MC ture i skoven og udlandet
- Bor i Græsted og har 2 voksne børn + 2 Børnebørn :-*



Fra kortnumre til Copilot



Pointen: DLP er gået fra at kigge på trafikken til at **kende data** – også når AI læser dem.

DLP: én kontrol, tre regelsæt

ISO/IEC 27001:2022 · ANNEX A · KONTROL 8.12

Data leakage prevention

Ny kontrol i 2022: forebyg og opdag uautoriseret videregivelse og udtræk af følsom information.

1

Kend data

Find og klassificér følsom information

2

Overvåg

Mail, fildeling, chat, AI og endpoints

3

Stop og log

Blokér eller advar, og dokumentér hver hændelse

GDPR

Beskyt persondata

Art. 5(1)(f) fortrolighed og integritet

Art. 25 databeskyttelse gennem design

Art. 32 passende tekniske foranstaltninger

Art. 33–34 brud anmeldes inden for 72 timer

NIS2

Beskyt forretningen

Art. 20 ledelsen er ansvarlig

Art. 21 risikostyring og sikkerhedstiltag

Art. 23 hændelser: varsel inden for 24 t

DK NIS2-loven gælder fra 1. juli 2025

Pointen: 8.12 er **hvordan** – GDPR og NIS2 er **hvorfor**. DLP-loggen er jeres **bevis**.

Eksempler på data-læk der kan stoppes

Source / Incident	CPR in file	Passport (OCR)	Customer data	Health data	Salary information	IBAN / payment card
Exchange	Block	Block	Block	Block	Alert	Alert
OneDrive	Block	Block	Block	Block	Alert	Alert
SharePoint	Block	Block	Block	Block	Alert	Alert
Teams	Block	Block	Block	Block	Alert	Alert
CoPilot	Block	Block	Block	Block	Alert	Alert

DLP ligger i den eksisterende platform

ADMINKONSOL

Discovery

Klassificering

DLP (ny sektion)

Rolletilladelse: DLP-admin

DLP: egen tilladelse, så ikke alle administratorer kan se den.

I DMCS

”DLP ligger i den samme DMCS platform, som I allerede bruger til oprydning og klassificering. I behøver ikke endnu en konsol eller agent.”

Vis

- Åbn DLP fra hovedmenuen
- Fremhæv DLP-rolletilladelsen

Én politik dækker alle fire kanaler

POLITIK · PERSONDATA MÅ IKKE SLIPPE UD

Kilder: Exchange, OneDrive, Teams, Copilot

Personnummer / CPR: **Blokér**

Helbredsdata: **Blokér**

IBAN / betalingskort: **Advarsel**

Lønoplysninger: **Advarsel**

Gem: alle mailregler oprettes automatisk.

Fejlsikring: mailregler lever i 10 minutter og fornyes hvert andet minut, så de udløber, hvis DLP stopper.

Projekt

”Vi anbefaler at man starter i ‘Alert’ tilstand og skifter derefter udvalgte klasser til ‘Block’”

Fremhæv

- Ekstern mail bliver altid kontrolleret
- Scannede filer køres gennem OCR
- Hvis DLP stopper, kører mailen videre

Fire hændelser, ét revisionsspor

DLP-REVISION · SENESTE 15 MINUTTER

Mail · Bruger · CPR: **Frigivet af Lena**

OneDrive · Bruger · kunder: **Link tilbagekaldt**

Teams · Bruger · helbred: **Advarsel**

Copilot · Bruger · løn: **Advarsel**

Eksport: dokumentation til revisor.

Revision

"Det er jeres dokumentation for ISO 27001:2022 kontrol 8.12, Data leakage prevention, og for GDPR art. 32 og NIS2."

Vis

Kanal, bruger, klasse, handling og hvem der frigav den.

Kristian Billeskov Bøving, PhD

Titel: Head of AI, Data & More

Akademisk Baggrund:

- Ph.D. i Computermedier fra Københavns Universitet

Erhvervserfaring:

- Head of AI hos Data & More
- Tidligere Director hos Gartner
- Speciale i digital transformation & komplekse systemer

Ekspertise & AI:

- Automatisering af databeskyttelse, ML & NLP
- EU-regulering (AI Act & GDPR)
- Ekstern ekspert for EU-Kommissionen

Lokation:

- Vejle, Danmark





GENERATIV AI & COPILOT TIL ORGANISATIONER

Giv Copilot adgang til jeres data – ikke til jeres hemmeligheder.

Sådan beskytter Data & More persondata og sikkerhedsdata, når generativ AI får adgang til organisationens data.

 Data, som Copilot må bruge

 Persondata & sikkerhedsdata – fundet og mærket



UDGANGSPUNKTET

Frigør dine data – sikkert.

Nødvendigheden af dataadgang

- For at generativ AI kan skabe reel værdi i virksomheden, skal teknologien have adgang til organisationens data.
- Copilot gør dette ved at vektorisere og søge i jeres data, så de kan bruges til at gøre hjælpen præcis.

Udfordringen med dataeksponering

- Hvordan forhindrer man, at de forkerte data bliver eksponeret og potentielt misbrugt?
- Adgang til data er afgørende for, at AI kan fungere, men det medfører betydelige risici for dataeksponering.



TRE VÆRKTØJER

Vi hjælper med at reducere Copilot-risiko med tre værktøjer

KLASSIFICERING

CoPilot Protection

Klassificerer data, som via et label undtager dem fra Copilot.

SLETNING & RAPPORTERING

CoPilot deletion

Introducerer chats og uploadede filer til chat i vores rapporter og sletningsprocedurer.

ADVARSLER & REGULERING

Copilot DLP

Giver mulighed for at advare brugere om forkert brug af Copilot.

