



SIXTUS

COMPLIANCE

# Dagens Talere

- Bo Pyskow – CEO og Co-founder.
  - IT Dude
  - Product owner & developer
  - [bo@sixtus-compliance.dk](mailto:bo@sixtus-compliance.dk)
  
- Jonas Funch – Data Lingvist
  - Sprog Nørd
  - Secret Sauce
  - [jonas@sixtus-compliance.dk](mailto:jonas@sixtus-compliance.dk)



SIXTUS  
COMPLIANCE

# Hvem er vi?

- Legal Tech Software og Compliance
- 2 Founders – Jurist og en IT Dude
- 8 kollegaer incl. Eksterne udviklere.
- Founded i 2019
- Nationale og internationale kunder i de fleste sektorer, dog ikke den offentlige og farma.
- Kontor i Nordhavn



**SIXTUS**  
COMPLIANCE

# Problemet vi gerne vil løse

Kæmpe mængder af Dark Data

Source: see last slide



SIXTUS  
COMPLIANCE



Mange fordele ved Data Management!

- Oprydning
- Kategorisering og Klassificering
- Case: GDPR
- Ressourcer



SIXTUS  
COMPLIANCE

# Hvordan?

## Sixtus Scanner

- Natural Language Processing
- On Premise Solution
- Alle filtyper.....også PDF!
- 1 time vs. 1 ½ År



SIXTUS  
COMPLIANCE

**Nu bliver det langhåret**

**Over til Jonas!**



**SIXTUS**  
COMPLIANCE

# Hvad er det vi gør, lavpraktisk....

- Lokalisering og verificering af sensitive "cases"
- Hvad er sensitiv data, og hvornår er noget sensitivt?
- Hvad er en 'case', er det kun tekst?
  - Multimodalitet: Lyd, billede, tekst, og i kombination
- Parring af enhed og egenskab



SIXTUS  
COMPLIANCE



# Kan det Machine Learnes?

- Regelbaseret vs. Machine learning
- Problemet med data
- Hybrid: En sprogteknologisk løsning
- Transparens on-demand



SIXTUS  
COMPLIANCE

# Verificeringen af følsomme egenskaber

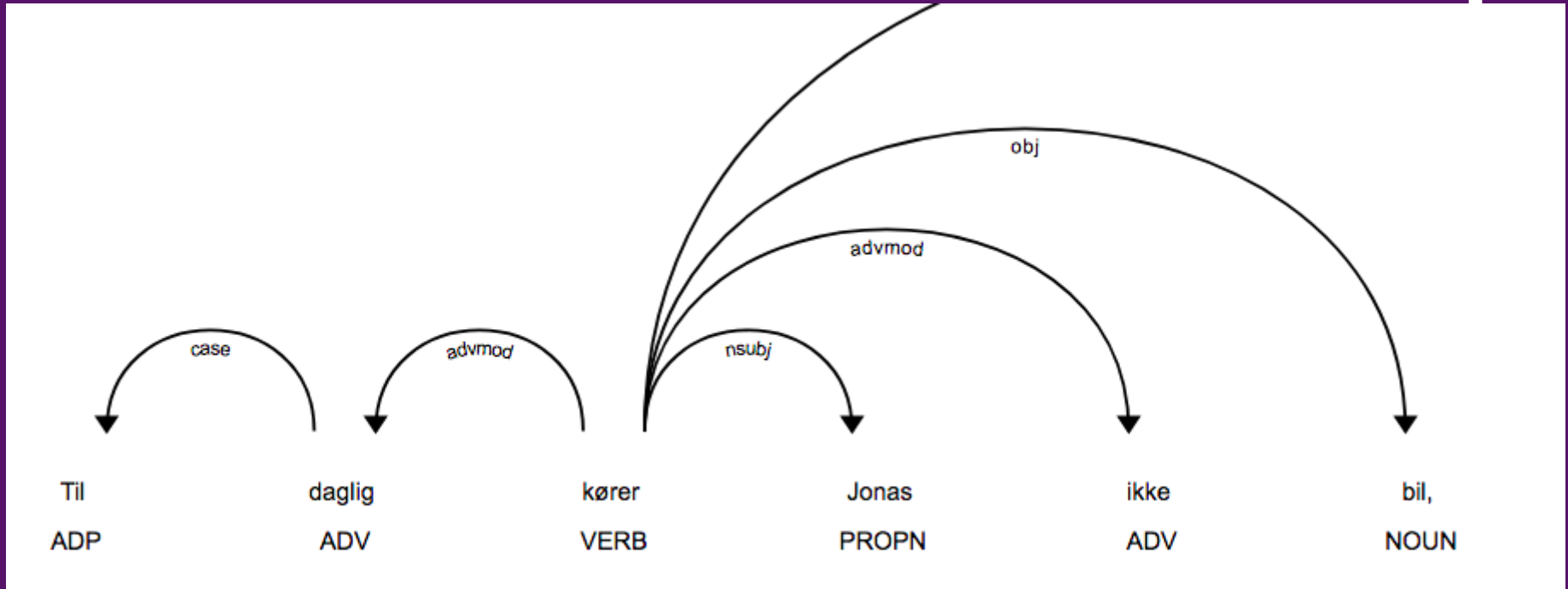
Hvad er kravene til udsagn i relation til personfølsom data?

- Syntaktiske krav
- Semantiske krav
- Faktivitet og præsupponering



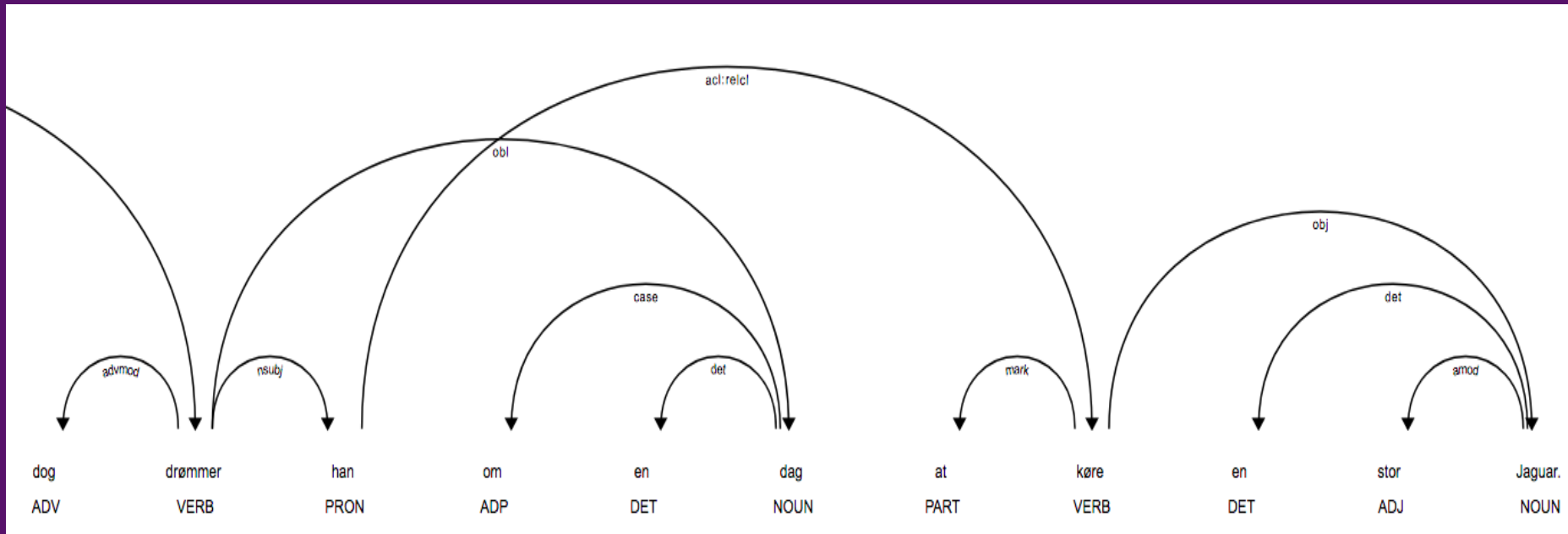
SIXTUS  
COMPLIANCE

# Eksempler



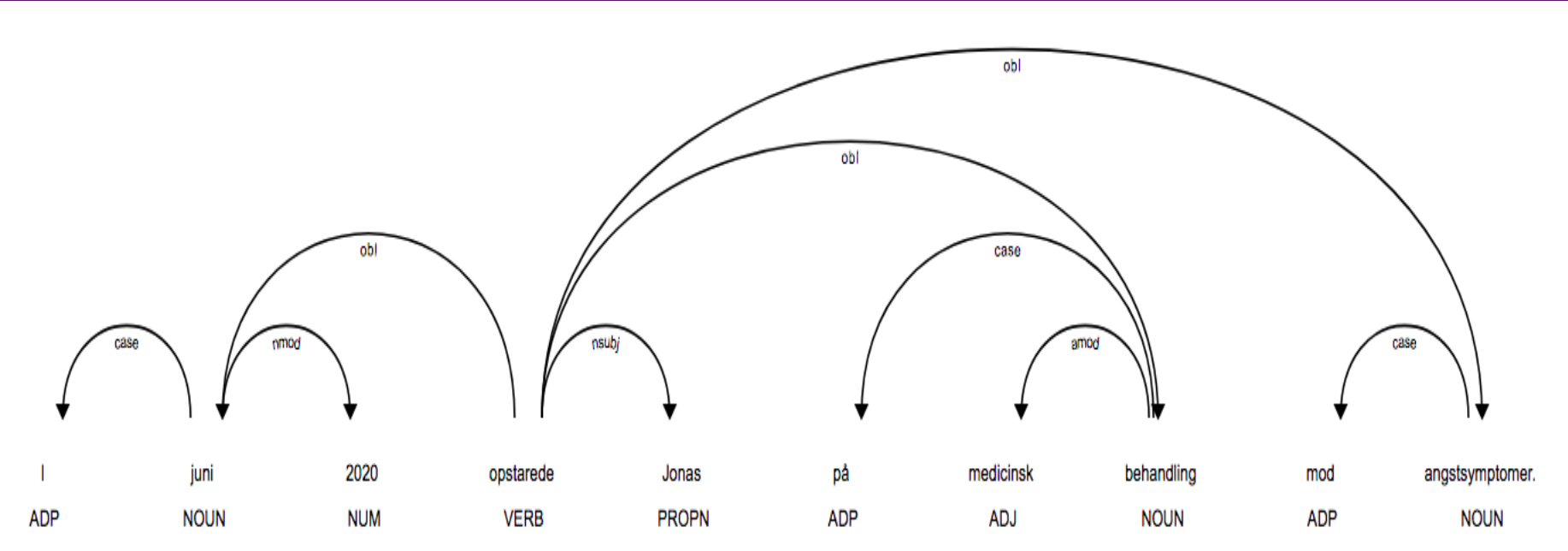
SIXTUS  
COMPLIANCE

# Eksempler



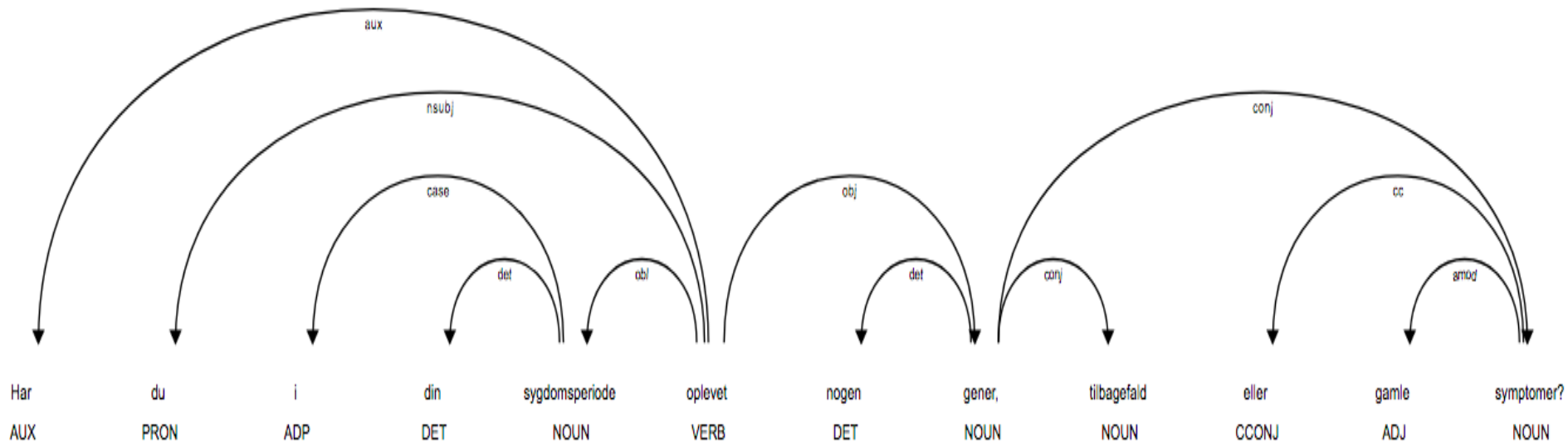
SIXTUS  
COMPLIANCE

# Eksempler



SIXTUS  
COMPLIANCE

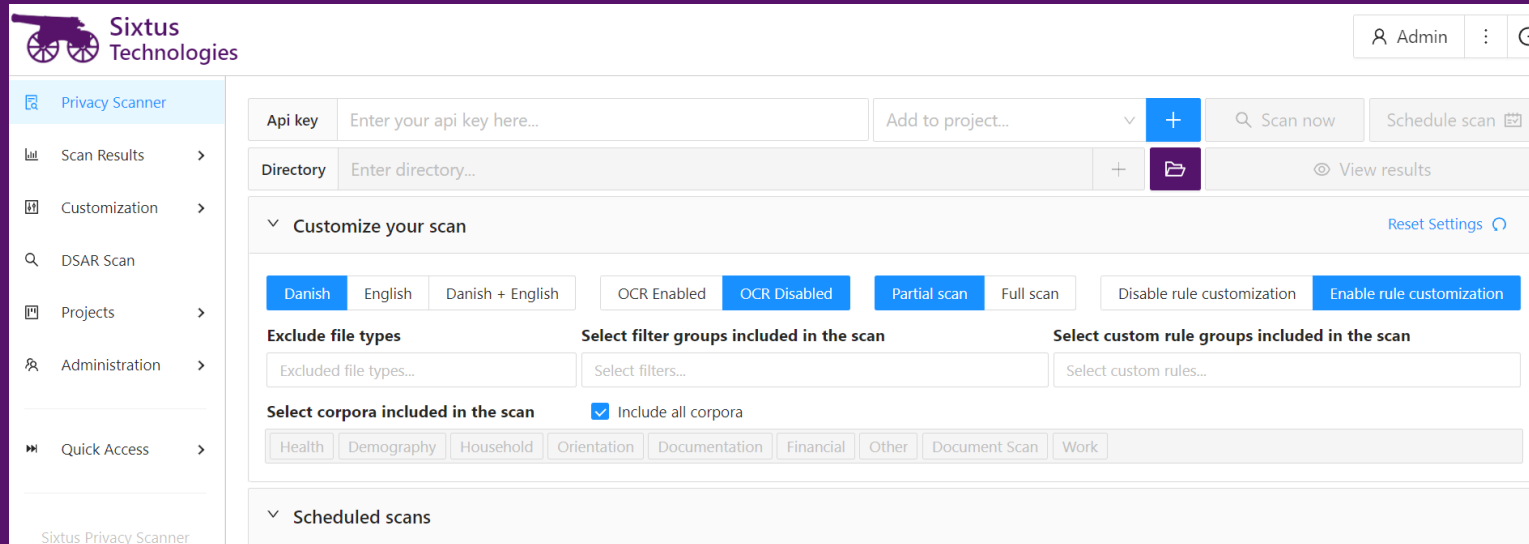
# Eksempler



SIXTUS  
COMPLIANCE

# Sixtus Privacy Scanner Screens

Easy interface to start scanning



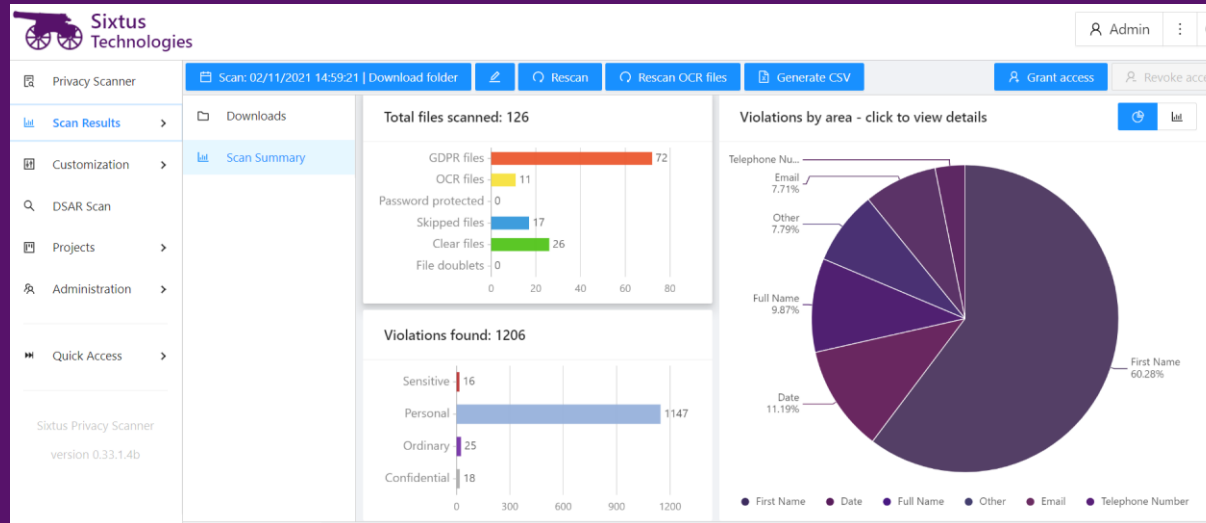
The screenshot displays the Sixtus Privacy Scanner web interface. The top left features the Sixtus Technologies logo and a navigation menu with items like Privacy Scanner, Scan Results, Customization, DSAR Scan, Projects, Administration, and Quick Access. The top right shows user information (Admin) and navigation icons. The main content area is titled 'Customize your scan' and includes fields for 'Api key' and 'Directory', along with buttons for 'Add to project...', 'Scan now', and 'Schedule scan'. Below these are options for language (Danish, English, Danish + English), OCR status (Enabled/Disabled), and scan type (Partial/Full). There are also sections for selecting filter groups and custom rule groups, and a 'Select corpora included in the scan' section with a checked 'Include all corpora' option and various category buttons like Health, Demography, Household, etc. A 'Scheduled scans' section is partially visible at the bottom.



SIXTUS  
COMPLIANCE

# Sixtus Privacy Scanner Screens

Understandable overview

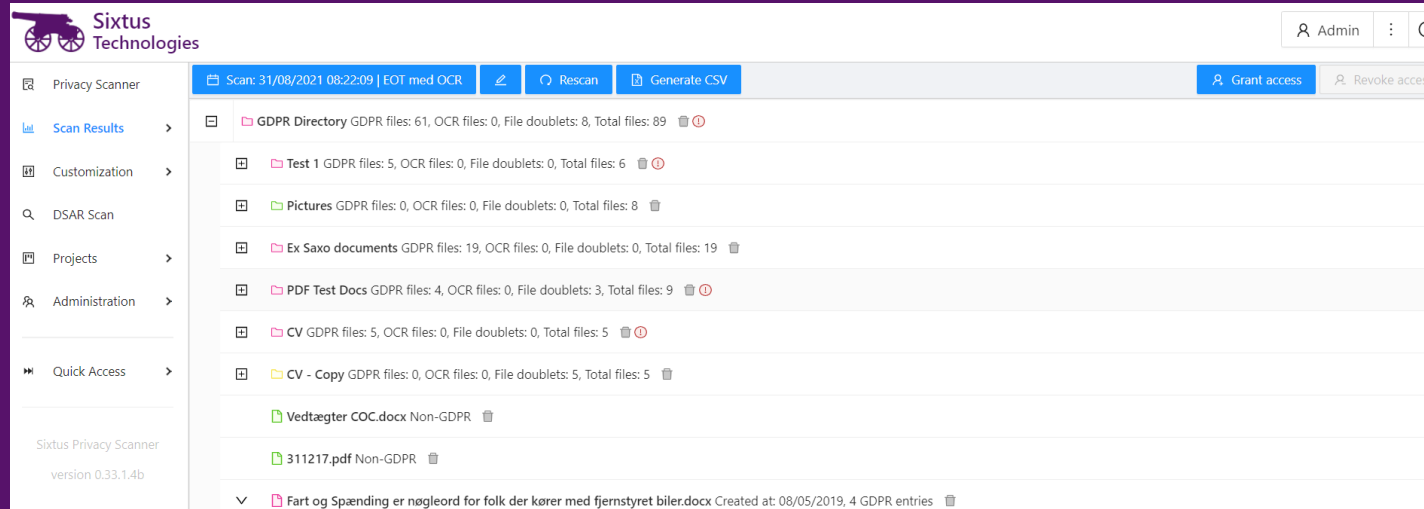


SIXTUS  
COMPLIANCE



# Sixtus Privacy Scanner Screens

Overview of folders,files and findings



The screenshot displays the Sixtus Privacy Scanner web interface. The top navigation bar includes the Sixtus Technologies logo, a user profile for 'Admin', and a search icon. Below the navigation bar, a status bar shows the scan details: 'Scan: 31/08/2021 08:22:09 | EOT med OCR'. Action buttons for 'Rescan' and 'Generate CSV' are visible. The main content area is a list of scanned folders and files, each with a summary of findings:

- GDPR Directory**: GDPR files: 61, OCR files: 0, File doublets: 8, Total files: 89
- Test 1**: GDPR files: 5, OCR files: 0, File doublets: 0, Total files: 6
- Pictures**: GDPR files: 0, OCR files: 0, File doublets: 0, Total files: 8
- Ex Saxo documents**: GDPR files: 19, OCR files: 0, File doublets: 0, Total files: 19
- PDF Test Docs**: GDPR files: 4, OCR files: 0, File doublets: 3, Total files: 9
- CV**: GDPR files: 5, OCR files: 0, File doublets: 0, Total files: 5
- CV - Copy**: GDPR files: 0, OCR files: 0, File doublets: 5, Total files: 5
- Vedtægter COC.docx**: Non-GDPR
- 311217.pdf**: Non-GDPR
- Fart og Spænding er nøgleord for folk der kører med fjernstyret biler.docx**: Created at: 08/05/2019, 4 GDPR entries

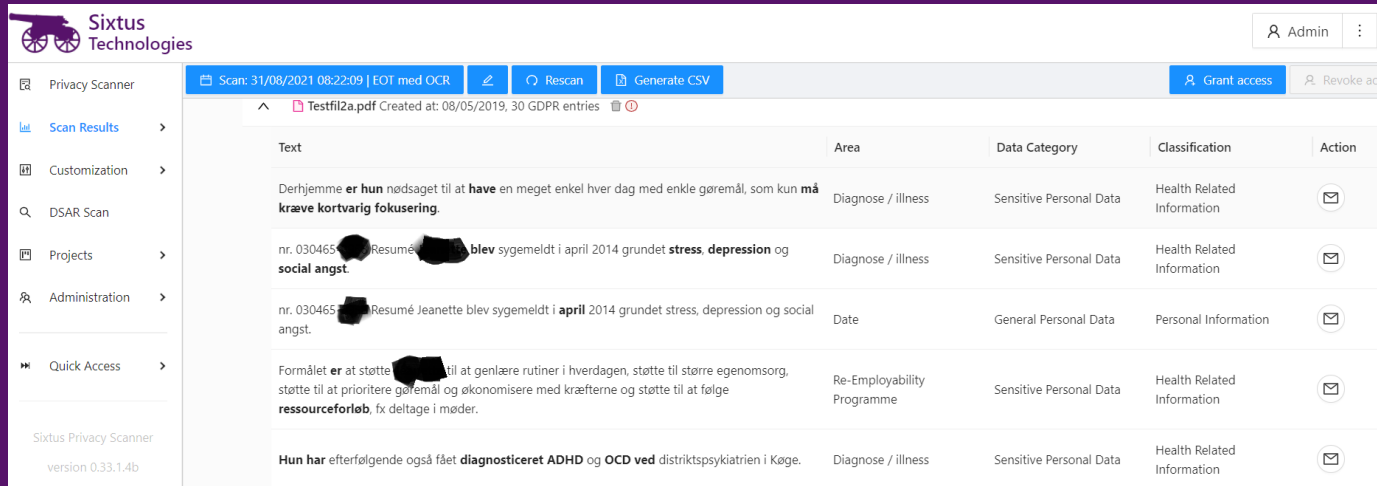
The left sidebar contains navigation options: Privacy Scanner, Scan Results, Customization, DSAR Scan, Projects, Administration, and Quick Access. The bottom left corner of the interface shows the version: 'Sixtus Privacy Scanner version 0.33.1.4b'.



SIXTUS  
COMPLIANCE

# Sixtus Privacy Scanner Screens

## Overview of folders, files and findings



The screenshot displays the Sixtus Privacy Scanner interface. The top navigation bar includes the Sixtus Technologies logo, a user profile for 'Admin', and a search icon. Below the navigation bar, the current scan is identified as 'Scan: 31/08/2021 08:22:09 | EOT med OCR'. Action buttons for 'Rescan' and 'Generate CSV' are visible. The main content area shows a table of findings from a PDF file named 'Testfil2a.pdf', which was created on 08/05/2019 and contains 30 GDPR entries. The table has five columns: Text, Area, Data Category, Classification, and Action. The findings are as follows:



Text	Area	Data Category	Classification	Action
Derhjemme <b>er hun</b> nødsaget til at <b>have</b> en meget enkel hver dag med enkle gøremål, som kun <b>må kræve kortvarig fokusering</b> .	Diagnose / illness	Sensitive Personal Data	Health Related Information	✉
nr. 030465 [redacted] Resumé [redacted] blev sygemeldt i april 2014 grundet <b>stress, depression</b> og <b>social angst</b> .	Diagnose / illness	Sensitive Personal Data	Health Related Information	✉
nr. 030465 [redacted] Resumé Jeanette blev sygemeldt i <b>april</b> 2014 grundet stress, depression og social angst.	Date	General Personal Data	Personal Information	✉
Formålet <b>er</b> at støtte [redacted] til at genlære rutiner i hverdagen, støtte til større egenomsorg, støtte til at prioritere gøremål og økonomisere med kræfterne og støtte til at følge <b>ressourceforløb</b> , fx deltage i møder.	Re-Employability Programme	Sensitive Personal Data	Health Related Information	✉
<b>Hun</b> har efterfølgende også fået <b>diagnosticeret ADHD</b> og <b>OCD</b> ved distriktspsykiatrien i Køge.	Diagnose / illness	Sensitive Personal Data	Health Related Information	✉










SIXTUS  
COMPLIANCE

# Sixtus Privacy Scanner Screens

PDF Scan!

^  Bo Pas og Sygesikring.pdf Created at: 08/05/2019, 6 GDPR entries 

Text	Area	Data Category	Classification	Action
Bo Pas og Sygesikring.pdf	Document: Passport	Confidential Information	Identification Information	
Tlf. 46 35 11 30.	Telephone Number	General Personal Data	Contact Detail	
DNK 206212278.	Passport Number	General Personal Data	Identification Information	
Christian <b>Bo Pyskow</b> .	Full Name	General Personal Data	Personal Information	
3/220570- 	CPR number	National Identification Data	DK Civil Registration Number	
<b>Bo</b> Pas og Sygesikring	First Name	General Personal Data	Personal Information	



SIXTUS  
COMPLIANCE

# Sources

Slide 4: <https://medium.com/untrite/what-is-the-dark-data-and-why-should-organisations-start-looking-into-it-61cdba7aab8f>

# Thank you for your time

E [bo@sixtus-compliance.dk](mailto:bo@sixtus-compliance.dk)

T +45 30 65 45 65



**SIXTUS**

COMPLIANCE