

Maskinlæring og kunstig intelligens

v/Geir Harbak - Løsningseier

Geomatikk IKT

Webinar Norsk Arkivråd 10.11.2021

GEOMATIKK IKT

- Eget selskap siden 2009
- Våre kunder er offentlig virksomhet og privat næringsliv
- Digitalisering av papirarkiver for siden 1998
 - ca 10 000 000 sider per år
- Over 80 ansatte fordelt på to lokasjoner i Trondheim
- Solid utviklingsmiljø med over 30 systemutviklere



EKSTERNE FORRETNINGSOMRÅDER



Arkiv og digitalisering

- Offentlige aktører
- Private aktører



Vegforvaltning

- Kommuner og fylkeskommuner
- Entreprenører
- Energiverk
- Statens vegvesen



Taksering for eiendomsskatt

- Kommuner

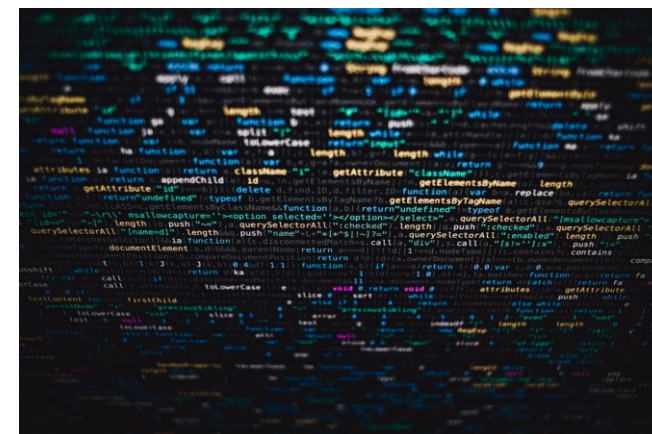
SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



Løsninger innen arkiv

braArkiv – arkivløsning/innsynsløsning for alle typer historiske data
docTorg – selvbetjeningsløsning for innbygger (www.doctorg.no)
docHub – løsning for datafangst/skanning



SAMLE

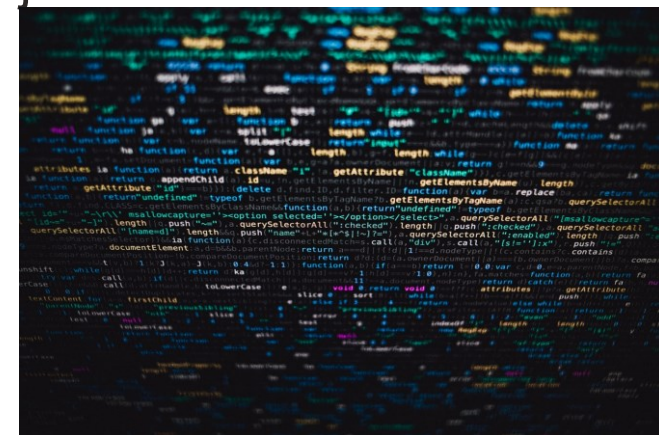
- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



docAI - Maskinlæring

Maskinlæring automatiserer ryddeprosessen

- Tilfører metadata, som for eksempel dokumenttyper/kategorier
- Kategoriserer og rekategoriserer enkeltdokumenter i journalposten iht KS ArkivLett-standard
- OCR-tolking på enkeltdokumenter/sider i journalposten
- Identifiserer dokumenter med sensitiv informasjon
- Identifiserer stempler og innhold i disse
- Gir bla muligheter for automatisk sladding



SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



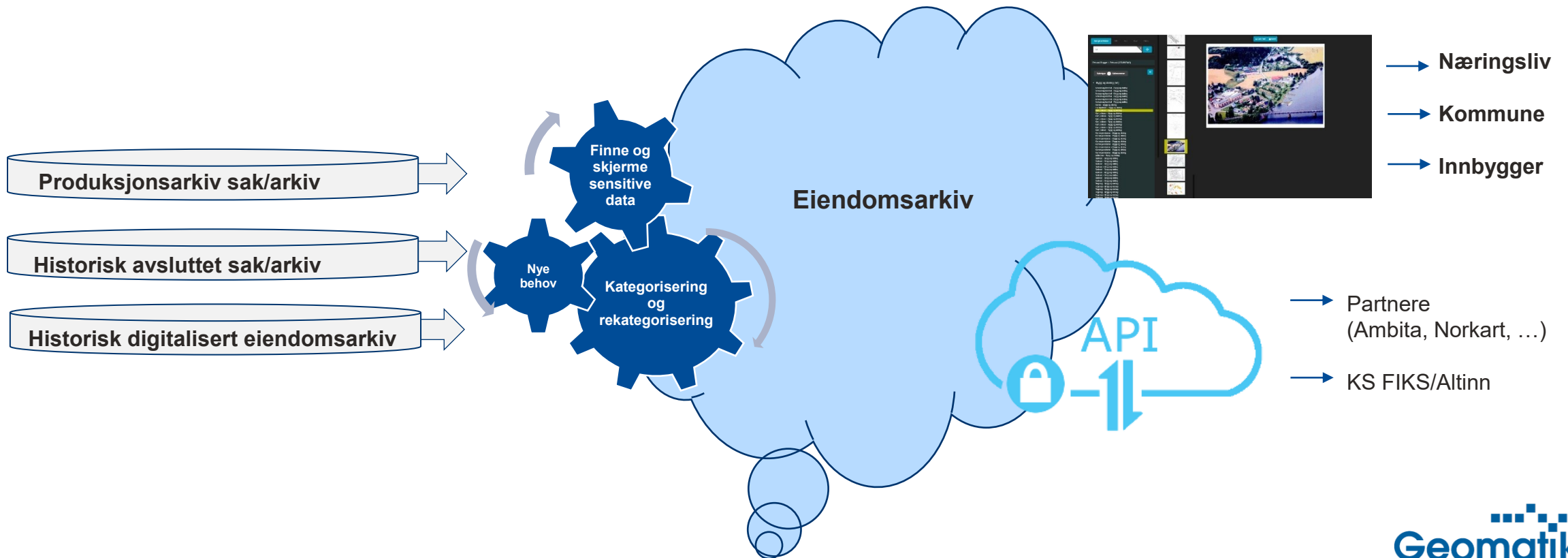
Skann meg!

braArkiv Synk

Kopiere data fra historiske- og produksjonsbaser

- **Kopierer** saker og dokumenter fra historiske baser
- **Kopierer** saker og dokumenter løpende fra produksjonsbasen til braArkiv
- Sammen med skannet papirarkiv gir dette en komplett kilde for etterspurte data i braArkiv
- Benytter GI-standard og etter hvert FIKS IO/FIKS Arkiv
- Rydder og kategoriserer dokumentene med docAI
- Vil automatisk sjekke for personnummer og sensitivitet

Ett komplett braArkiv



SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



GEVINSTER

All data samlet på et sted i sky legger til rette for:

- ✓ Ett sted å søke
- ✓ Ett sted å integrere
- ✓ En komplett og lik samling av data for innsyn uavhengig kilde
- ✓ Frigjøring av ressurser
- ✓ Innsparing av tid
- ✓ Automatisk matrikkel oppdatering
- ✓ Komplette datagrunnlag for publisering i innbyggerløsninger
- ✓ Sikrer beredskap



SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



Sensitivitetsanalyse med docAI

- Gjenkjenner sensitivinformasjon ut fra liste med ord og fraser
- Finner personnummer også håndskrift
- Viktig før publisering

SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



Sensitivitetsanalyse med docAI

- Kjøres mot historiske data før publisering
- Kjøres automatisk ved kopiering fra produksjonsarkiv

SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



Skann meg!

Sensitivitetsanalyse med docAi

DokumentId	Sidenr	Arkivdel	GID	Dokumentkategori	Gradert	SensitivMetadata
20045633	3	Byggesak	4/1/0/0	05 Søknad	0	[{"searchterm":"erklæring*","found":["erklæring","erklæring","erklæring","Erklæring"]},{"searchterm":"innlagt*","found":["innlagt","innlagt","innlagt"]}]
20045638	1	Byggesak	4/1/0/0	10 Avtale, erklæring	1	[{"searchterm":"dom","found":["dom"]}]
20045639	2	Byggesak	4/1/0/0	11 Ansvar og kontroll	1	[{"searchterm":"vitnemål*","found":["vitnemål"]}]
20045639	3	Byggesak	4/1/0/0	11 Ansvar og kontroll	1	[{"searchterm":"erklæring*","found":["erklæring","Erklæring"]}]
20045643	3	Byggesak	4/1/0/0	15 Interne dokument	1	[{"searchterm":"erklæring*","found":["Erklæring","erklæring"]},{"searchterm":"fullmakt*","found":["fullmakt"]}]
20045645	1	Byggesak	4/1/0/0	02 Vedtak	0	[{"searchterm":"fullmakt*","found":["fullmakt"]}]
20045648	1	Byggesak	4/1/0/0	05 Søknad	0	[{"searchterm":"attest","found":["attest"]}]
20045654	1	Byggesak	4/1/0/0	11 Ansvar og kontroll	1	[{"searchterm":"attest","found":["attest"]}]
20045654	2	Byggesak	4/1/0/0	11 Ansvar og kontroll	1	[{"searchterm":"erklæring*","found":["erklæring"]}]
20045679	2	Byggesak	4/1/0/0	05 Søknad	0	[{"searchterm":"erklæring*","found":["erklæring"]},{"searchterm":"innlagt*","found":["innlagt","innlagt"]}]
20045679	4	Byggesak	4/1/0/0	05 Søknad	0	[{"searchterm":"dommen","found":["dommen"]}]
20045689	2	Byggesak	4/1/0/0	15 Interne dokument	1	[{"searchterm":"innlagt*","found":["innlagt","innlagt"]}]
20045689	3	Byggesak	4/1/0/0	15 Interne dokument	1	[{"searchterm":"fengsel*","found":["Fengsel"]},{"searchterm":"sykehus","found":["Sykehus"]}]
20045692	1	Byggesak	4/1/0/0	02 Vedtak	0	[{"searchterm":"fullmakt*","found":["fullmakt"]}]
20045710	1	Byggesak	4/1/0/0	05 Søknad	0	[{"searchterm":"erklæring*","found":["Erklæring"]}]
20045710	2	Byggesak	4/1/0/0	05 Søknad	0	[{"searchterm":"erklæring*","found":["erklæring"]}]
20045710	3	Byggesak	4/1/0/0	05 Søknad	0	[{"searchterm":"ufør*","found":["Ufør"]}]
20045720	2	Byggesak	4/1/0/0	15 Interne dokument	1	[{"searchterm":"innlagt*","found":["innlagt","innlagt"]}]
20045720	3	Byggesak	4/1/0/0	15 Interne dokument	1	[{"searchterm":"fengsel*","found":["Fengsel"]},{"searchterm":"sykehus","found":["Sykehus"]}]
20045739	2	Byggesak	4/1/0/0	02 Vedtak	0	[{"searchterm":"dom","found":["dom"]}]
20045742	1	Byggesak	4/1/0/0	05 Søknad	0	[{"searchterm":"sykehus","found":["sykehus","sykehus"]}]
20045744	2	Byggesak	4/1/0/0	07 Kart	0	[{"searchterm":"sykehus","found":["sykehus"]}]
20045744	3	Byggesak	4/1/0/0	07 Kart	0	[{"searchterm":"sykehus","found":["sykehus"]}]
20045744	5	Byggesak	4/1/0/0	07 Kart	0	[{"searchterm":"sykehus","found":["sykehus"]}]
20045744	6	Byggesak	4/1/0/0	07 Kart	0	[{"searchterm":"sykehus","found":["sykehus"]}]

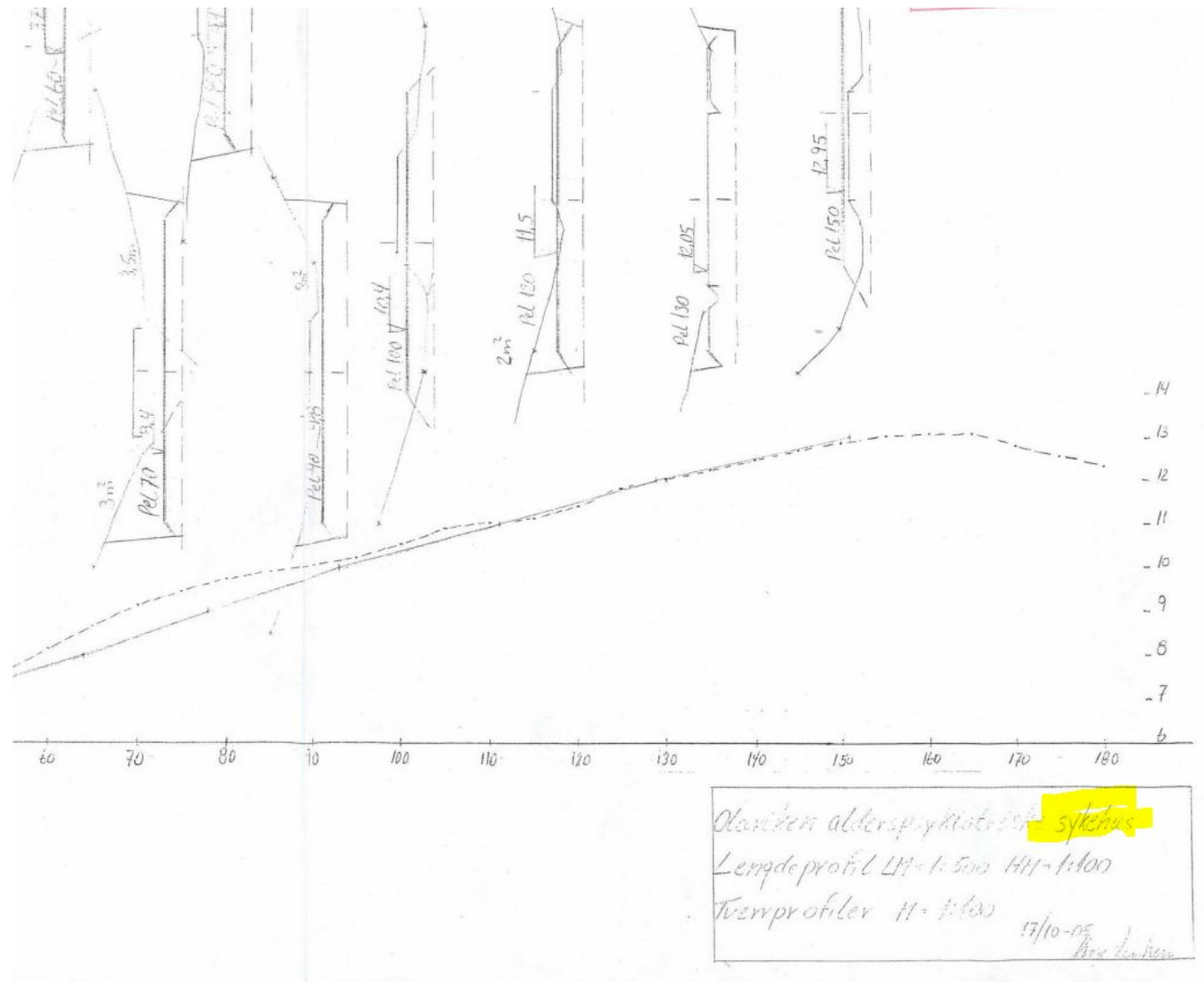
SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



Skann meg!

Sensitivitetsanalyse

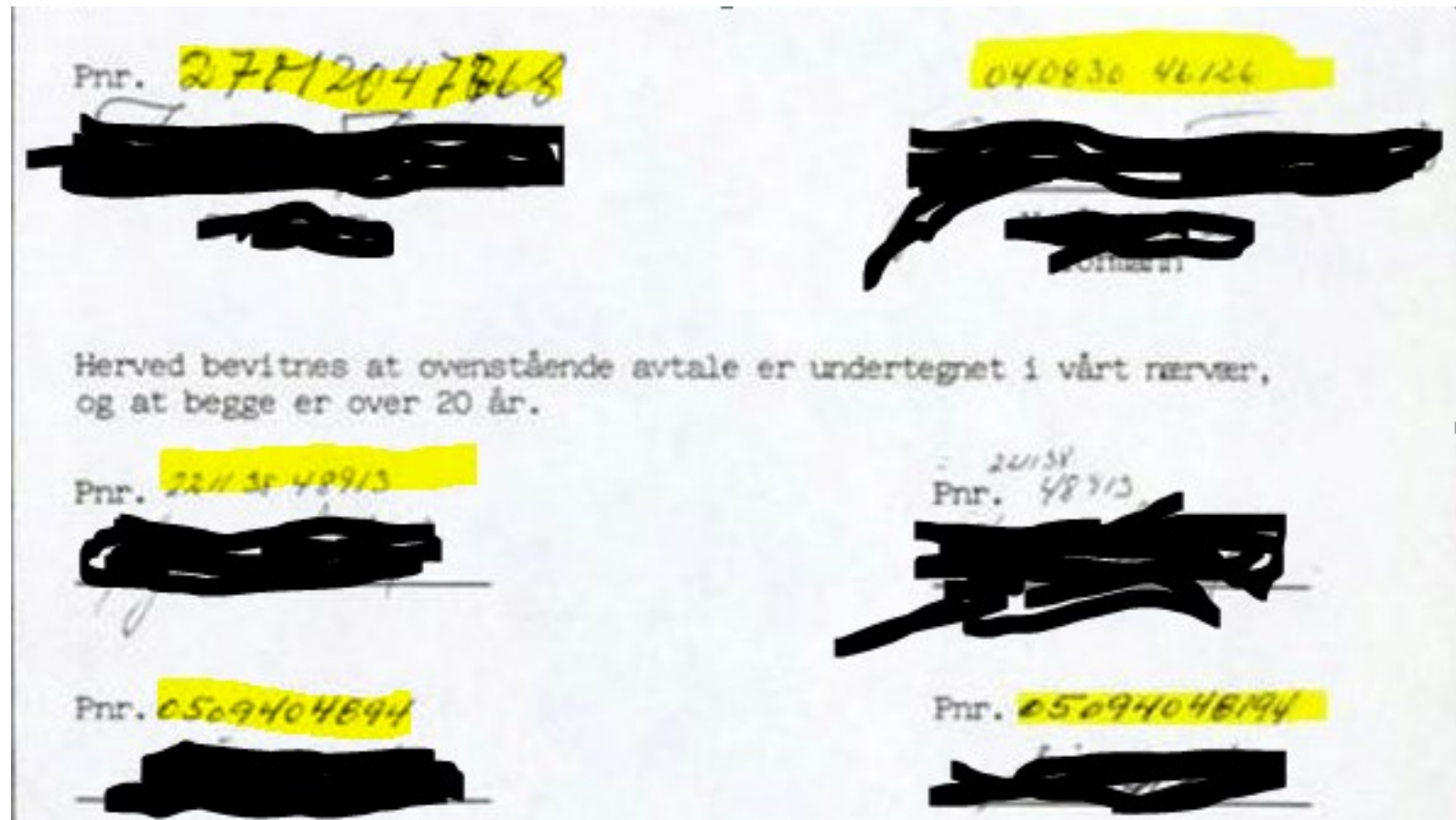


SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



Sensitivitetsanalyse



SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



Nye områder

- Forbedre eksisterende modeller med mer trening
- Sladding «on the fly»
- Tolke innhold og ikke bare ord
- Automatisk registrering av metadata ved digitalisering av historisk materiale
- Skanne filområder
- Automatisk matrikkelvask

SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



Kostnader/Gevinster

Kostnader

- Utvikling skjer i eget utviklingsmiljø

Innsats fra kunde

- Kunde må angi hva som ønskes rekategorisert arkiv/arkivdeler osv og evt tilføye fjerne ord/fraser fra ordliste som benyttes ved sensitivitetsanalyser

SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



Utfordringer/Sikkerhet

- Ingen 100% garanti
- Fortsatt behov for menneskelige vurderinger
- Det er fortsatt den enkelte etat/kommunes ansvar å overholde regelverket i forhold til hva som publiseres, men vi kan bistå.

SAMLE

- Digitalisering av papirarkiv
- Samling av elektroniske arkivkilder
- Sammenstilling av data
- Alt på et sted!



SPØRSMÅL?

Geir Harbak
918 58 480
geir.harbak@geomatikk-ikt.no