

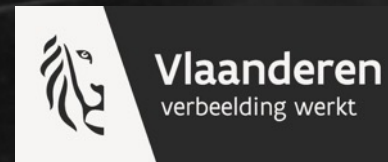
TEGENLICHT

Het digitaliseren, beschrijven en toegankelijk maken van zes journalistieke negatiefarchieven bewaard in ADVN, FOMU en KOERS.



FOMU foto
museum

KOERS
KONINKLIJK ORANJE-ROOD
INSTITUUT VOOR
SPORT



PROJECT

- inhaalbeweging digitale collectiedata
- ADVN, FOMU en KOERS
- voortbouwen op bestaande expertise
- consolideren in collectiewerking



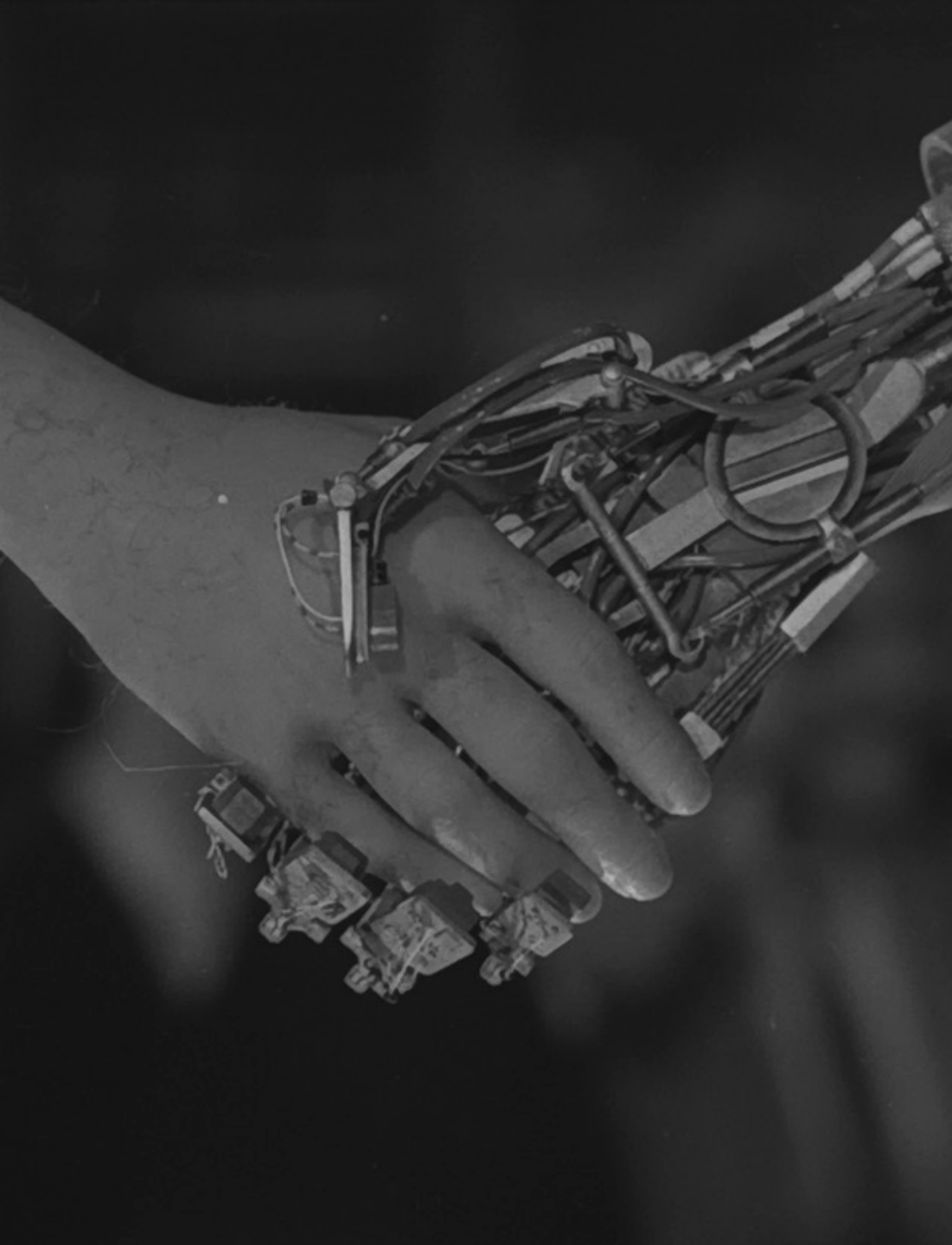


MATERIAAL

- > 400.000 negatieven, goede staat
- verschillende maatschappelijke aspecten in Vlaanderen van 1950 tot 2000:
- politiek, cultuur en media, sport

DOEL

- digitaliseringsgraad verhogen
- registratieachterstand verminderen
- expertise vergroten en delen





PLANNING

Fase 1: 2021

collectiebeschrijving

digitalisering

beeldherkenning vooronderzoek

Fase 2: 2022

beeldherkenning implementatie

inventarisatie

Fase 3: 2023

reconciliatie

open data

onderzoek referentieset en scanning on demand



DIGITALISERING

- Picturae
- oplevering december 2021

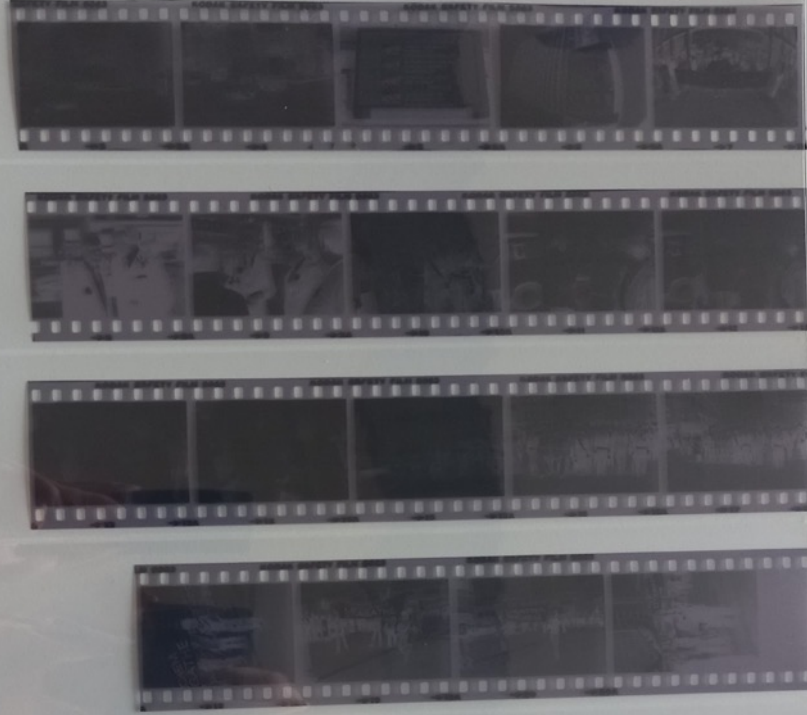
- ca. 13.000 contactvellen
- ca. 450.000 beelden

'A-S' Archival System
Sheet Ref No. AS20MS7SE
Thetford Norfolk U.K.

Secol



VNE_000639



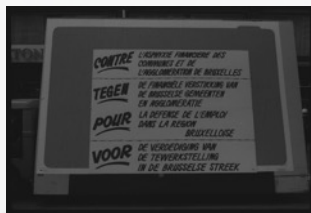
'A-S' Archival System
Sheet Ref No. AS20MS7SE
Thetford Norfolk U.K.

Secol



VNE_000639





VNE_000639

Secret



DIGITALISERING

- in bulk
- contactvelmethode + cropppen
- resolutie 800 ppi
- doel = beeldherkenning

- Picturae
- oplevering december 2021
- ca. 13.000 contactvellen
- ca. 450.000 beelden



BEELDHERKENNING VOORONDERZOEK

- Databank
- use cases:
 - instroom beelden
 - gezichtsdetectie
 - gezichtsherkenning
 - validatie
 - koppeling aan autoriteiten
 - terugstroom metadata naar collectiebeheersystemen
- modulaire architectuur uitwerken
- Knime

100%

3: Workshop Welcome to KNIME Analytics Platform 3: Building a Simple Classifier

KNIME Explorer

- My-KNIME-Hub (api.hub.knime.com)
- EXAMPLES (knime@api.hub.knime.com)
- LOCAL (Local Workspace)
 - Example Workflows
 - Basic Examples
 - Building a Simple Classifier
 - Data Blending
 - Simple Reporting Example
 - Visual Analysis of Sales Data

Workflow Coach

Node recommendations only available with usage d

Node Repository

- IO
- Manipulation
- Views
- Analytics
- DB
- Other Data Types
- Structured Data
- Scripting
- Tools & Services
- KNIME Labs
- Workflow Control
- Workflow Abstraction
- Reporting

Data Reading
Read the adult data set file. There is one row for each person, plus demographic info and the income group. The file is located in TheData/Basics/

Graphical Properties
Assign colors by income group.

Data Partitioning
Create two separate partitions from original data set: training set (80%) and test set (20%).

Train a Model
This node builds a decision tree. Other Learner nodes train other models. Most Learner nodes output a PMML model (blue square output port).

Apply the Model
Predictor nodes apply a specific model to a data set and append the model predictions.

Score the Model
Compute a confusion matrix between real and predicted class values and calculate the related accuracy measures.

Descriptive Statistics
Calculate the statistical properties of the data set attributes.

Try this:
KNIME's Interactive Visualizations:
1) Execute the workflow
2) Open the Scorer node view
3) Highlight a cell in the confusion matrix
4) Open the Interactive Table view
5) Select "Highlight" -> "Filter" -> "Show Highlighted Only"
This shows only the misclassified data rows.

Interactive Table
Display table of the entire data

Visualize
Create interactive scatter plot.

Outline

Console Node Monitor

Node:

State: no node selected

Port Output:

INGEST

- via externe harde schijven
- workflow MAM meemoo
- kopie naar tijdelijke drager
- bewerkt met XNView MP
- beeldmetadata gecreëerd en toegevoegd





GEZICHTSDETECTIE & -HERKENNING

- Google Vision API
- Google Cloud + Mongo DB
- faceID's in W3C Annotation formaat
- ~~clusteren~~
- referentieset
 - naam
 - ID
 - 5-tal afbeeldingen
- Azure



VALIDATIE

- LabelStudio
- medewerkers en vrijwilligers

Submit

#290198



Admin created, less than a minute ago

#8n4eh 2/2



- checkbox 1 image
- checkbox 1 image
- checkbox 1 image
- checkbox 1 image
- checkbox 1 image
- checkbox 1 image
- checkbox 1 image
- checkbox 1 image
- checkbox 1 image
- checkbox 1 image

Vic Anciaux

0.95428

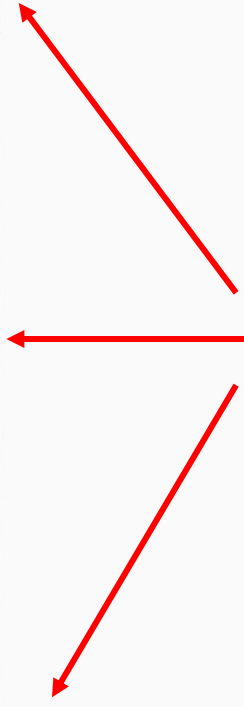
Is deze persoon juist geïdentificeerd? Vergelijk met de gevalideerde foto(s) rechts

Juist^[1] Fout^[2] onduidelijk^[3]
 onbekend^[4]

andere id



Vertical toolbar with icons for zoom, pan, and other image manipulation tools.



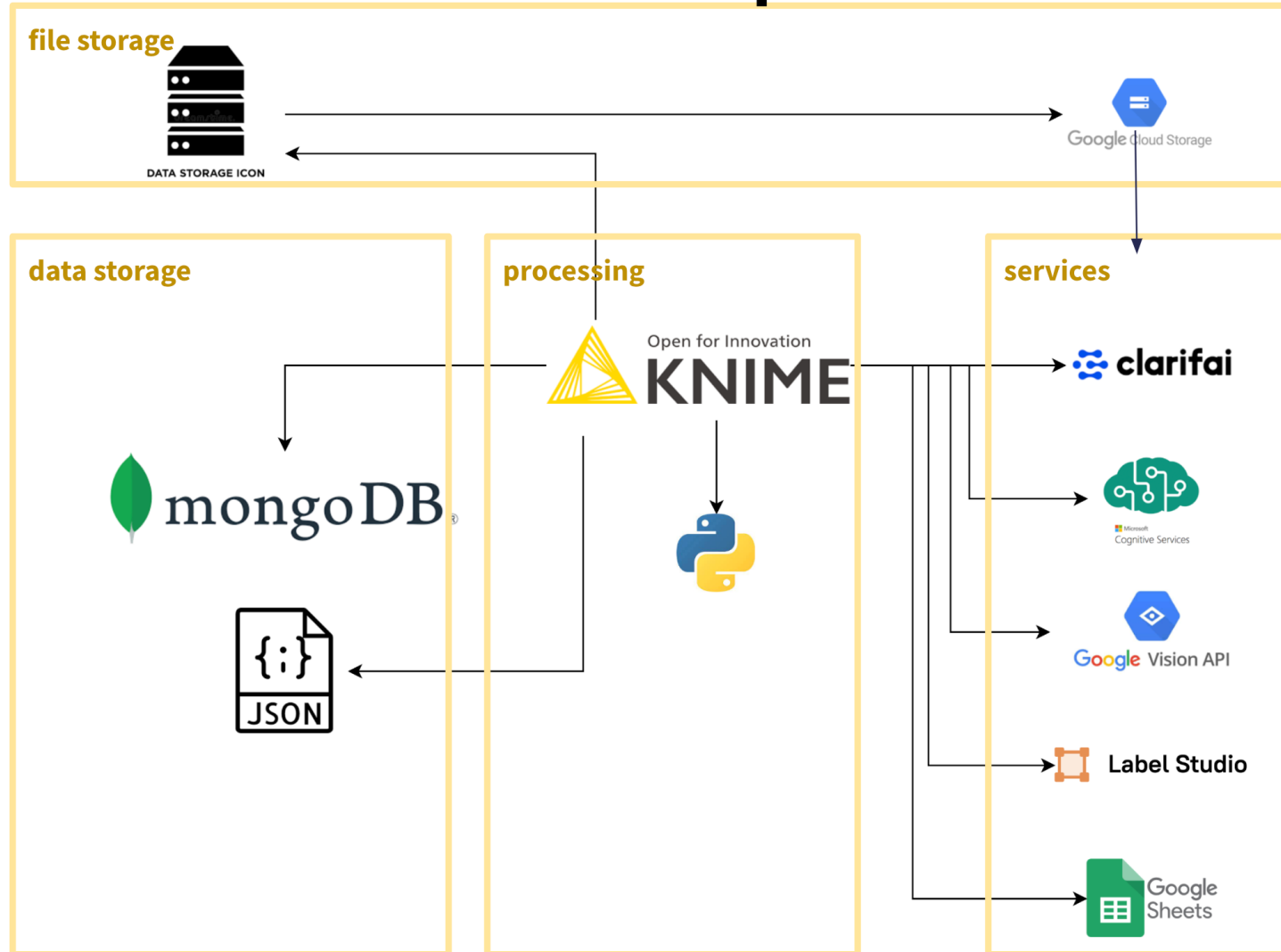
Task #290198

VALIDATIE



- LabelStudio
- medewerkers en vrijwilligers
- juist / fout / onduidelijk / onbekend
- beelden uit referentieset
- waarschijnlijkheidsscore
- ID van andere persoon toevoegen
- ondergrens op 0,50

Workflow componenten



Submit

#290198



Admin created, less than a minute ago

#8n4eh 2/2



- image
- 1
- 1
- 1
- 1
- 1
- 1
- 1
- 1
- 1
- 1

Vic Anciaux

0.95428

Is deze persoon juist geïdentificeerd? Vergelijk met de gevalideerde foto(s) rechts

Juist^[1] Fout^[2] onduidelijk^[3]

onbekend^[4]

andere id



Task #290198

BEVINDINGEN



- Vóór validatie:
 - 1.325 unieke personen herkend
 - 168.580 gezichten herkend
- Maar:
 - 980 personen hebben gemiddelde confidence score onder 0,50
 - 57.707 taken in LabelStudio dus herkende gezichten met score hoger dan 0,50

BEVINDINGEN

- Azure levert goede resultaten op
- niet alleen frontale portretten
- ook in massa, gezicht gedraaid of niet helemaal scherp



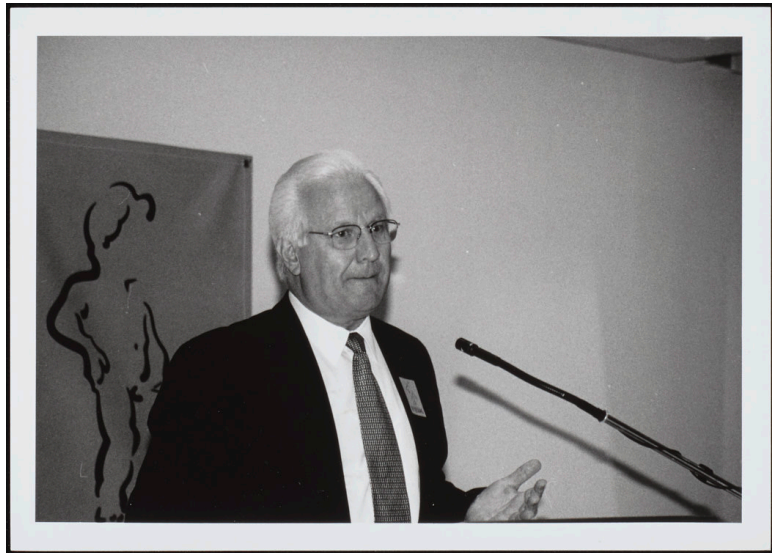




BEVINDINGEN BEELDHERKENNING

- Azure levert goede resultaten op
- niet alleen frontale portretten
- ook in massa, gezicht gedraaid of niet helemaal scherp
- belang referentieset!









person_name	imageFB (aantal)	aantal gevonden faces	confidence Level (gemiddeld%)	confidence Level (bin 10%)
Vic Anciaux	7	1.941	86	80-90
Gaston Geens	5	1.722	83	80-90
Hugo Schiltz	6	1.210	83	80-90
Jaak Gabriëls	6	1.036	80	71-80
Annie Pottie-Kindt	1	950	39	31-40
Lieve Favoreel	2	833	43	41-50
Lionel Vandenberghe	4	824	80	71-80
Jetje Claessens	3	818	45	41-50
Guy Verhofstadt	4	805	78	71-80
Bert Anciaux	6	729	82	80-90
Jo Dries	1	691	41	41-50
Paula D'Hondt	3	663	50	41-50
Johan Sauwens	7	656	81	80-90
Rika Steyaert	2	640	53	51-60
Willy Kuijpers	7	579	82	80-90
Leonie Verstraeten	1	560	45	41-50
Yvonne Publie	1	551	43	41-50
Wilfried Martens	5	546	77	71-80
Paula Marckx	2	540	42	41-50

- image
- 0
- 0
- 0
- 0
- 0
- 0
- 0
- 0
- 0
- 0

#289298 Admin #eAs0q 2/2 created, less than a minute ago

Submit

Annie Pottie-Kindt

0.5001

Is deze persoon juist geïdentificeerd? Vergelijk met de gevalideerde foto(s) rechts



- Juist^[1]
 Fout^[2]
 onduidelijk^[3]
 onbekend^[4]

andere id





person_name	imageFB (aantal)	aantal gevonden faces	confidence Level (gemiddeld%)	confidence Level (bin 10%)
Kamiel Goossens	1	97	94	90-100
Vic Anciaux	7	1.941	86	80-90
Gaston Geens	5	1.722	83	80-90
Hugo Schiltz	6	1.210	83	80-90
Bert Anciaux	6	729	82	80-90
Willy Kuijpers	7	579	82	80-90
Johan Sauwens	7	656	81	80-90
Paul Van Grembergen	4	514	81	80-90
Filip De Winter	5	514	80	71-80
Jaak Gabriëls	6	1.036	80	71-80
Jef Valkeniers	5	384	80	71-80
Lionel Vandenbergh	4	824	80	71-80
Guy Verhofstadt	4	805	78	71-80
Walter Baeten	1	186	78	71-80
Luc Vandenbrande	4	354	77	71-80
Wilfried Martens	5	546	77	71-80
André Denys	5	243	76	71-80

☐ image

#284275

PH phaedra.claeys #2703
created, 7 months ago

2/2



Update

☐ 0



☐ 1



☐ 1



☐ 1



☐ 1



☐ 1



☐ 1



☐ 1



☐ 0



☐ 1



☐ 1



Kamiel Goossens

0.98524

Is deze persoon juist geïdentificeerd? Vergelijk met de gevalideerde foto(s) rechts

Juist^[1] Fout^[2] onduidelijk^[3] onbekend^[4]

andere id





BEVINDINGEN & TOEKOMST



- referentieset is cruciaal
- liever niet dan slecht
- onder- en bovengrens bepalen
- afweging hoeveelheid / accuraatheid een hoeveelheid / manuren
- werkgroep aanbevelingen

- referentieset verder uitbouwen
- documentatie en workflows in toekomst inzetten