

FAIR Photos

Transforming a collection of two million historical press photos into five-star data

Leon van Wissen (UvA)

Nico Vriend (NHA)

CLARIAH Tech & Data Day, 29 June 2023

Data

- 2 million photos of Photo Press Agency De Boer (1945-2005)
- Noord-Hollands Archief (NHA)
- Example of streetlife, popular culture, creative industries, occupations, canonization, etc.



Curation

Linking to URIs of photo contents:

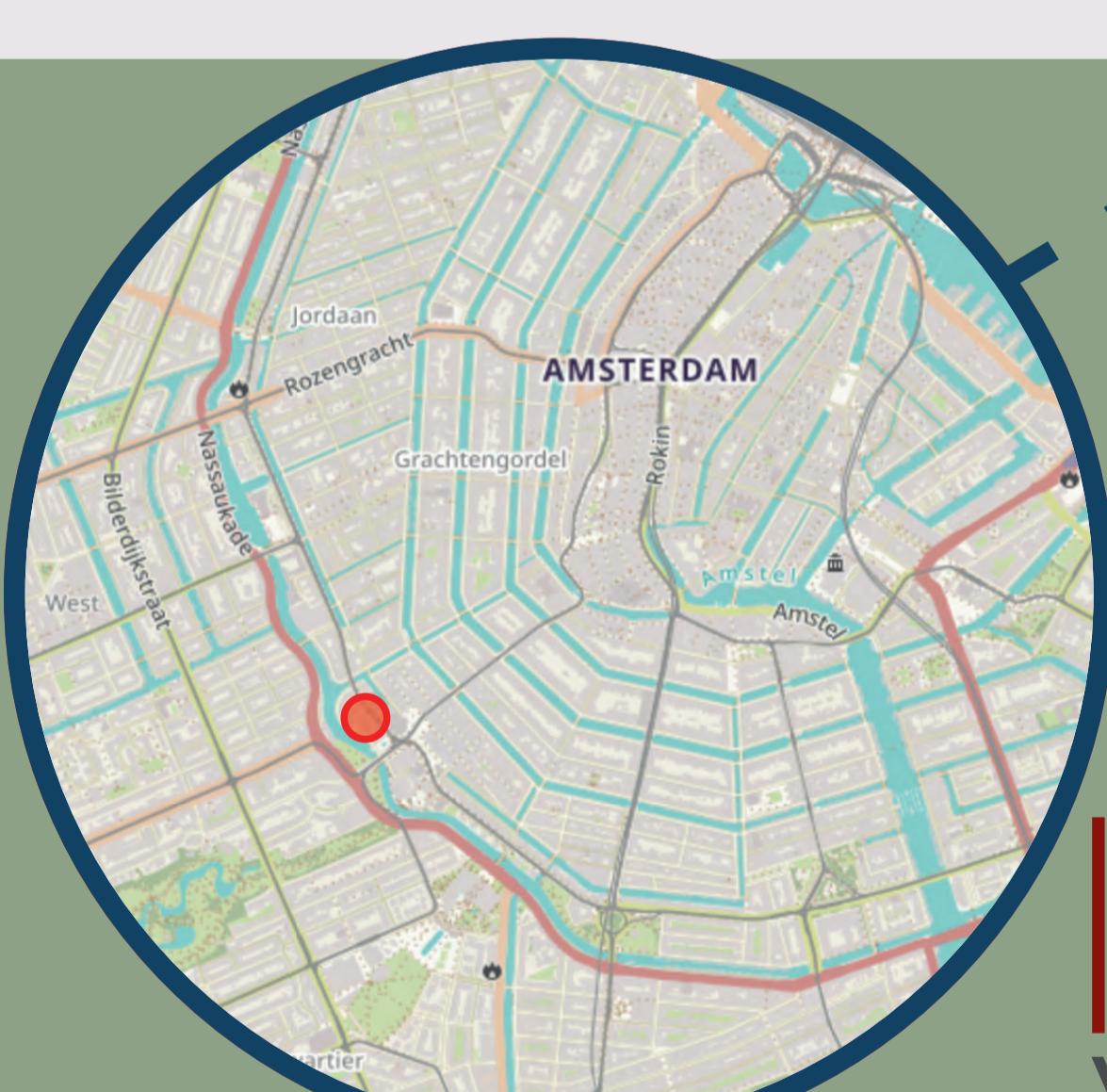
- Locations (Wikidata)
- Persons (Wikidata)
- Concepts in own SKOS thesaurus (RCE-infrastructure) adding hierarchy and links to other thesauri (AAT)

Publication

- Publishing data as CSV and RDF
- Storing all changes and enrichments in the archive's CMS
- Provenance (docs + code) in public git-repository

Relevance

- Corpus of 2 million tagged and well-described photos
- SSH, Cultural Heritage, Computer Vision
- More attention for open data in GLAM and academia
- Facilitate creative re-use (e.g. in Wikimedia)
- Knowledge sharing in GLAM



Locations

Jan Wolkers (Q379679)
Nederlands schrijver en beeldhouwer
Wolkers | Jan H
Hoofdpagina Gebruikersportaal De kroeg Nieuw item aanmaken Recent wijzigingen Willekeurig item Query Service In de buurt Random Primary Sources item Help ongeren
Hoofdpagina Gebruikersportaal De kroeg Nieuw item aanmaken Recent wijzigingen Willekeurig item Query Service In de buurt Random Primary Sources item Help ongeren
Taal Label Beschrijving
Nederlands Theo van Gogh Nederlands (1957–2004)
Engels Theo van Gogh Dutch film
British English Theo van Gogh geen beschrijving
American English Geen label ingesteld geen beschrijving
Alle ingevoerde talen



Persons

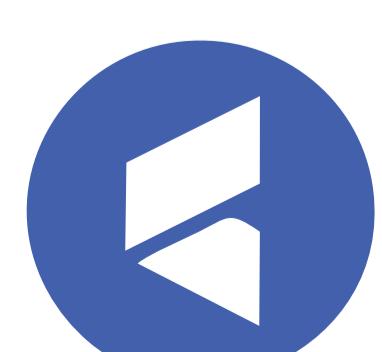
Theo van Gogh (Q205456)
Nederlands regisseur en schrijver (1957–2004)
Theodoor van Gogh | Theodoor "Theo" van Gogh
Hoofdpagina Gebruikersportaal De kroeg Nieuw item aanmaken Recent wijzigingen Willekeurig item Query Service In de buurt Random Primary Sources item Help ongeren
Hoofdpagina Gebruikersportaal De kroeg Nieuw item aanmaken Recent wijzigingen Willekeurig item Query Service In de buurt Random Primary Sources item Help ongeren
Taal Label Beschrijving
Nederlands Theo van Gogh Nederlands (1957–2004)
Engels Theo van Gogh Dutch film
British English Theo van Gogh geen beschrijving
American English Geen label ingesteld geen beschrijving
Alle ingevoerde talen



Concepts



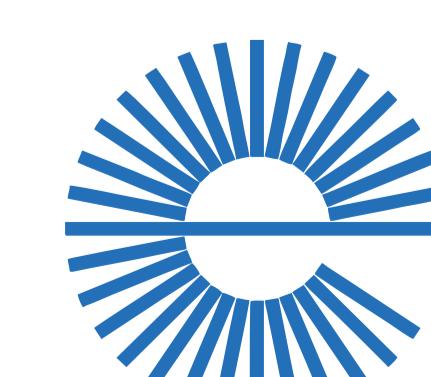
UvA



Clariah



noord-hollands archief



netwerk
digitaal
erfgoed