DIVE INTO THE EVENT-BASED BROWSING OF LINKED HISTORICAL MEDIA

VICTOR DE BOER, JOHAN OOMEN, OANA INEL, LORA AROYO, ELCO VAN STAVENEN, WERNER HELMICH AND DENNIS DE BEURS
DIGITAL HUMANITIES RESEARCHERS

Media researcher Lars Arve Røssland of the University of Bergen. (Photo: Andreas R. Graven)
EXPLORATIVE SEARCH

Erp, M. van; Oomen, J.; Segers, R.; Akker, C. van de; Aroyo, L.; Jacobs, G.; Legêne, S; Meij, L. van der; Ossenbruggen, J.R. van; Schreiber, G. Automatic Heritage Metadata Enrichment with Historic Events Museums and the Web 2011
http://www.museumsandtheweb.com/mw2011/papers/automatic_heritage_metadata_enrichment_with_hi
**USER EXPERIENCE**

- "DIGITAL SUBMARINE" UI PROVIDES EXPLORATIVE BROWSING
- SUPPORT FOR COLLECTION INTERPRETATION AND NAVIGATION
- TOUCH-BASED FIRST
- HISTORICAL EVENT-BASED RESEARCH AND EXPLORATION
- E.G. ART-HISTORY STUDENTS, HUMANITIES SCHOLARS

**API**

- SEARCHING AND RETRIEVAL OF EVENTS AND OBJECTS THROUGH SPARQL QUERIES
- SMART IMAGE CACHE PROVIDES VISUAL REPRESENTATIONS OF ENTITIES RETRIEVED FROM COMMONS

**TRIPLE STORE**

DIVE RDF'S DATAMODEL BASED ON SIMPLE EVENT MODEL (SEM), OPENANNOTATION (OA) AND SKOS

**ENTITY EXTRACTION**

CROWD-MACHINE PIPELINE FOR EVENT EXTRACTION AND ENTITY LINKING

**DATASETS**

MEDIA OBJECTS FROM OPEN API FROM CULTURAL HERITAGE PARTNERS: VIDEOS FROM NETHERLANDS INSTITUTE FOR SOUND AND VISION; RADIO NEWS BULLETIN SCANS FROM ROYAL LIBRARY
DATA: OPENIMAGES.EU

Open videos Netherlands Institute for Sound and Vision 3000, mostly news broadcasts
Scans of Radio bulletins (hand annotated)
1937 – 1984
1.5 Million OCR’ed and NErred
ENTITY EXTRACTION

EVENTS CROWDSOURCING AND LINKING TO CONCEPTS THROUGH CROWDTRUTH.ORG

SEGMENTATION & KEYFRAMES

LINKING EVENTS AND CONCEPTS TO KEYFRAMES

CROWDTRUTH.ORG
SIMPLE EVENT MODEL (SEM), OPENANNOTATION (OA) AND SKOS

DIVE:MEDIA OBJECT

SEM:EVENT

OA:ANNOTATION

SEM:PLACE

SKOS:CONCEPT

SEM:ACTOR

SEM:TIME

LINKS TO EUROPEANA (MULTILINGUAL)

LINKS TO DBPEDIA
CLIOPATRIA TRIPLE STORE

Welcome to ClioPatria, the SWI-Prolog Semantic Web Server

This server provides a Web frontend to the SWI-Prolog SesWen package supporting both the SPARQL and Sesame SeRQL query language.

About this server

This server is built on top of the SWI-Prolog Web infrastructure (PDF). It provides a web frontend to the SWI-Prolog Semantic Web libraries (PDF). Features:

- Fast in-memory RDF triple store, scalable to approx. 15 million triples on 32-bit and virtually unlimited on 64-bit hardware.
- Persistent store using journal files.
- Supports SPARQL and SeRQL 1.2 HTTP protocol for remote access.
- Full text search: prefixes of literals, words appearing in literals, stemming.

Further information and contact

- SWI-Prolog site and manual
- SPARQL Documentation
- Sesame and SeRQL site
- ClioPatria mailinglist
- SWI-Prolog mailinglist

130K TRIPLES (FOR NOW)
SPARQL ENDPOINT

HTTP://ECULTURE.CS.VU.NL:8877/DIVE/HOME
DIGITAL SUBMARINE UI

https://www.flickr.com/photos/benjcarson/245171885
INFINITY OF EXPLORATION

https://www.flickr.com/photos/mibuchat/2774251415
DEMO

DIVE.BEELDENGELEID.NL
THANK YOU

https://www.flickr.com/photos/robysaltori/

DIVE.BEELDENGELUID.NL
v.de.boer@vu.nl