Moderna Tider 1936

A Data Management Plan created using DMPRoadmap

Creators: Mathias Johansson, Pelle Snickars

Affiliation: Lund University

Template: RJ

ORCID iD: https://orcid.org/0000-0001-5122-1549

Project abstract:
What is it that software sees, hears and perceives when technologies for pattern recognition are applied to media historical sources? All historical work requires interpretation, but what kind of algorithmic interpretations of modernity does software yield from historical archives? MODERN-36 is empirically committed to everyday experiences and sets out to study how machines interpret symbols of modernity in media from 1936. By utilising photographic and audiovisual collections, MODERN-36 seeks to analyse how modern Sweden was, while also exploring how computational methods can help us understand modernity in new ways. MODERN-36 will explore how artificial intelligence and machine learning methods can foster new knowledge about the history of Swedish modernity—while at the same time critically scrutinizing algorithmic toolboxes for the study of the past. MODERN-36 will use three datasets from 1936: 15,000 digitized photographs from DigitaltMuseum, all surviving radio programs from Swedish Radio, and all weekly newsreels and short films produced by Svensk Filmindustri. The research focuses on modernity in relation to gender, urbanity and industrialization, and will: (1.) to examine how software can assist historians in discerning new historical knowledge, (2.) to construct midsize and curated datasets that increase the scholarly capacity of media historical sources, and (3.) to interrogate algorithmic detection by evaluating what machines can—or cannot—notice in historic data.

Last modified: 31-08-2023

Moderna Tider 1936

---

1. Data description

- using existing data (archival data, data from previous project – your own or other researchers')
- collecting material or generate/produce data (e.g. interviews, measurement data)

The collected data is publicly available, published online under CC licenses, and will be downloaded for local processing.
The downloaded material will be used as a base to generate new content, such as upscaled images which get exponentially larger every time they are processed, the originals and the generated files will be processed with different AI tools such as object, detection classification etc.

Digitized and publicly available photograph, video, audio with metadata.

- .png
- .jpeg
- .jpg
- .zip
- .json
- .txt
- .sqlite3
- .tsv
- .csv
- .mp4
- .webm
- .3gpp
- etc.

- >1 TB

2. Documentation and data quality

We will only use non-proprietory software and file-formats.
Data pipelines will be constructed from open-source tools, and will be published at the end of the project.

N/A

3. Storage and backup

After the data has been collected, it will be strictly read-only, ergo they only need to be backed up once.

We are only working with publicly open data – hence security and sensitivity is not an issue.

4. Legal and ethical requirements

- No

We are only working with publicly open data – hence security, sensitivity and personal data is not an issue.

See 4.1

See 4.1

- No, LU is the sole party.

LU is the only part handling data.

5. Data sharing and long-term preservation

- Yes, but only metadata

Metadata will be provided through Zenodo.

DOI through Zenodo.

- It is not possible to assess this at the present time.

6. Responsibilities and resources

Lund University, project manager (Snickars) and systems developer (Johansson).

The small amount of funding needed for data management will be provided by the project.