## Labor gone digital?

The trade union movements digital transformation: the preservation of organization material in the digital age



## About the project

Department of Government at Uppsala University
Researchers: Jenny Jansson and Katrin Uba
Collaboration with TAM and ARAB



## Problem and goal

 More and more activities takes place online through websites and social media and there is no system for archiving these activities

 Our goal is to build a system for archiving web pages and feeds from social media from the Swedish unions



## Why?

• In order to allow research on our contemporary unions in the future.

 The database will allow for comparisons over time and between different unions in Sweden.



#### How?

- Identify all organizations trade unions (federal, regional and local departments)
- Information from websites, facebook, twitter, instagram, youtube. Download, the first time everything, then at regular intervals, ie, once a month, but the front page and social media daily.

Indexing, primarily automatically by Content Analysis (Word combinations)

e material will be handed over to TAM and ARAB

#### Technical aspects

- One program for down loading
- One program to detect what's new since the last time
- One program for idexing
- Freeware



## Indexing

To make the database useful to as many people as possible the material will be indexed.

- Index of ordinary activities of Swedish unions
- Software that calculates the most common combination of words



# Who will have access to the database?

The website will be open to all (if not the unions wishes to restrict usage)

 Social media feeds will be limited because they contain personal information



#### Harvesting

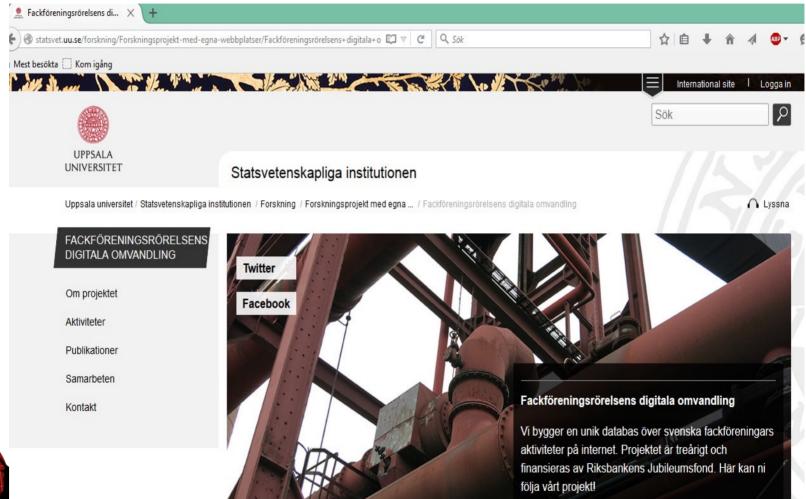
- The NetarchiveSuite software was originally developed by the two national deposit libraries in Denmark, The Royal Library
   (KB) and The State and University Library (SB)
- The NetarchiveSuite is a complete web archiving software package developed from 2004 and onwards.
- The primary function of the NetarchiveSuite is to plan, schedule and run web harvests of parts of the Internet.
- The NetarchiveSuite is built around the Heritrix web crawler, which it uses to harvest the web.

#### ARAB:s role

- Help and support the research team in the connections with the trade unions
- Long time preservation of the database
- Provide the database to researchers



#### www.statsvet.uu.se





#### Contact information

Jenny jansson@statsvet.uu.se

