



UPPSALA
UNIVERSITET

Fackföreningsrörelsens digitala omvandling.

Att bevara organisationsmaterial i den digitala tidsåldern

Jenny Jansson
Katrin Uba

Statsvetenskapliga institutionen

RJ-finansierat
infrastrukturprojekt

Vad händer med material som skapats på internet?

Allt fler aktiviteter sker online och det finns inget system för arkivering av dessa aktiviteter.

Vårt mål är att arkivera material genererat av svenska fackföreningar på internet.

Vi laddar ner och indexera svenska fackföreningars hemsidor, Facebook-gruppsidor, Twitterflöden, youtubefilmer osv.



UPPSALA
UNIVERSITET

Vad gör vi som andra inte gör?

- KulturArw3, Internet Archive etc.
- Regelbunden nedladdning (jämna intervall)
- Hamnar hos organisationernas egna arkiv
- Sökfunktion
- Sociala medier – finns det någon som arkiverar flöden?



UPPSALA
UNIVERSITET

Varför fackföreningar?

- Gammal rörelse med exceptionellt bra (pappers)arkiv
- En rörelse som är lätt att identifiera





UPPSALA
UNIVERSITET

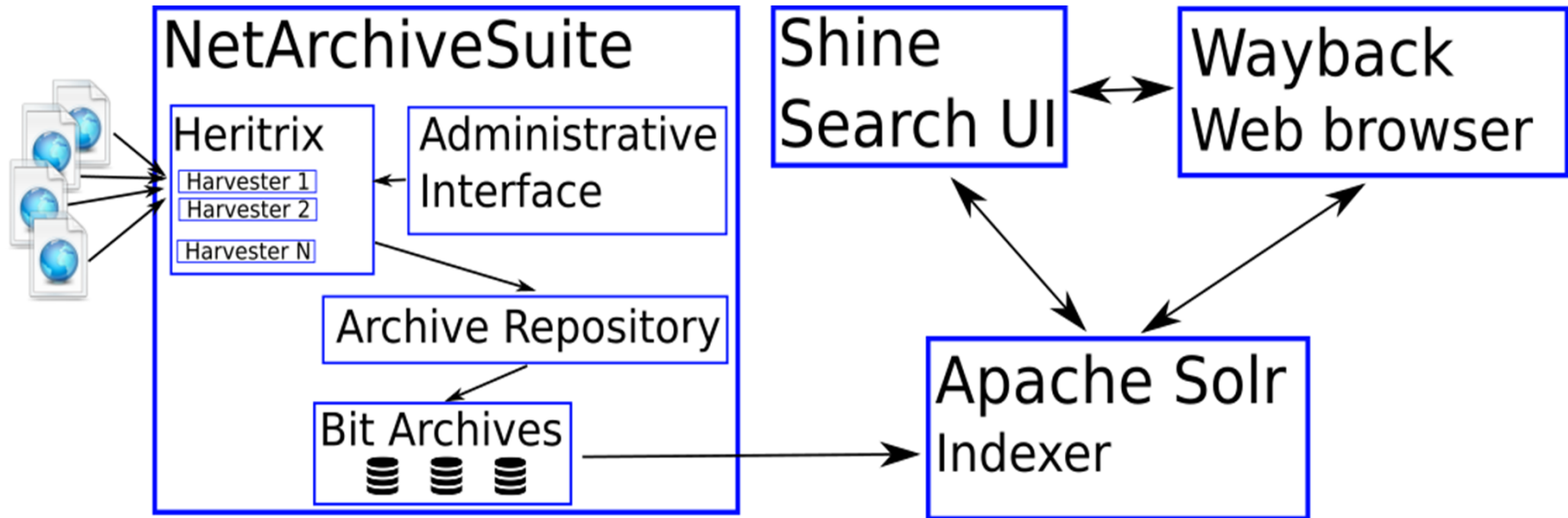
Vad gör vi?

Fas 1: Förarbete

1. Samla in adresser till Facebook, Twitter, Instagram, YouTube
2. Intervjuer med sociala medieansvariga på förbunden
3. Samlat in samtycke



Fas 2: Hemsidor





UPPSALA
UNIVERSITET

Nedladdning

- NetArchiveSuite, heritrix3
 - Danska KB
- Hela webbplatsen: en gång i månaden
- Förstasidan: en gång per dag





UPPSALA
UNIVERSITET

Indexering

- Solr
 - Fulltextsökning
 - Indexering





Gränssnitt 1: SHINE

Search [Advanced Search](#)

General Content Type 10

- html 59,748,477
- other 28,845,532
- text 3,294,963
- Show more...

Crawl Years 10

- 2005 14,190,001
- 2004 12,301,855
- 2012 11,622,915
- 2008 8,655,255
- 2011 8,200,379
- Show more...

Postcode District 10

- L70 547,773
- NW10 356,023
- GU46 311,442
- NW1 306,109
- BT15 221,923
- Show more...

Public Suffix 10

- co.uk 85,340,734
- org.uk 3,203,334
- ac.uk 1,583,614
- gov.uk 1,204,978
- com 260,408
- Show more...

Sample Mode Search Reset

Search Term(s): cat

Results [Concordance](#)

Results 1 to 10 of 92,718,387 CSV Asc

« 1 2 3 4 5 6 » Action

1 ""
other

2 "Connections to the Outside World"
Mar 15 1995 11:53:12 GMT
html
bssv01.lancs.ac.uk





Gränssnitt för hemsidorna

Organisation

Tidpunkt (nedladd.)

Dokumenttyp

FRITEXTSÖKNING

SAKREGISTER

Placering på hemsidan

Filstorlek

Filtyp



Gränssnitt 2: Wayback Machine



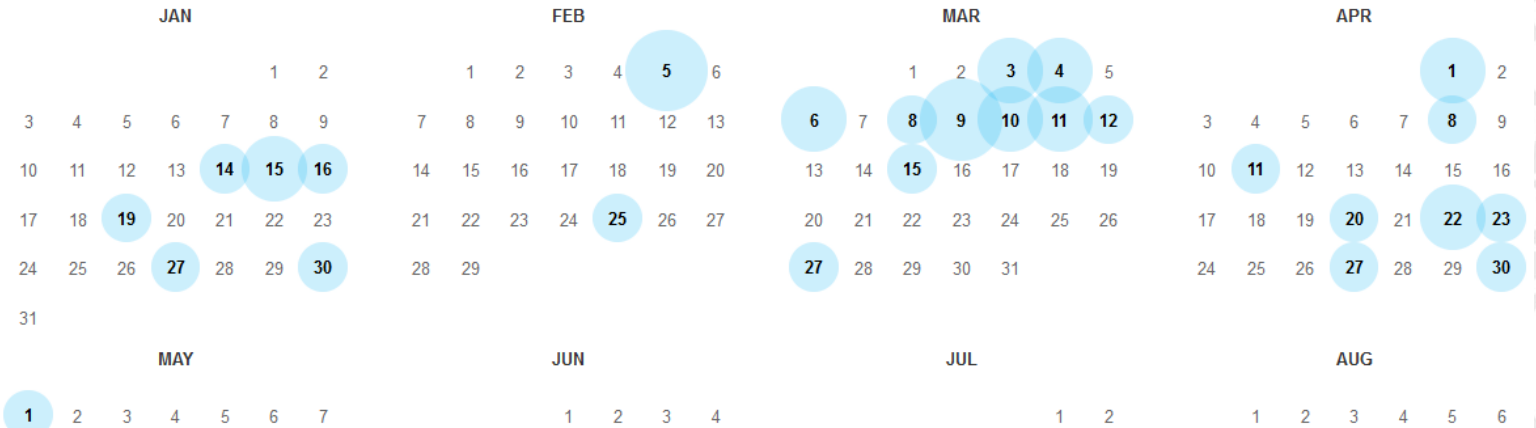
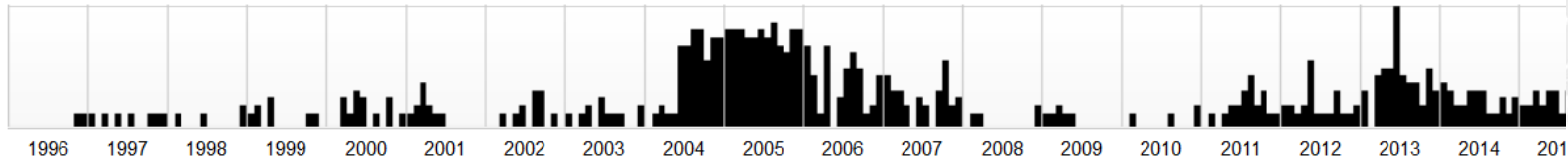
http://www.lo.se

BROWSE HISTORY

<http://www.lo.se>

Saved **917 times** between **November 9, 1996** and **November 17, 2016**.

PLEASE DONATE TODAY. Your generosity preserves knowledge for future generations. Thank you.





UPPSALA
UNIVERSITET

Sammanfattning hemsidorna

- Beprövad mjukvara för harvesting
- Beprövad mjukvara för indexering och användargränssnitt
- Allt funkar typ





UPPSALA
UNIVERSITET

Fas 3: Sociala medier

- Nedladdning: Olika program för olika sociala medier
- Indexering: Solr
- Användargränssnitt: SHINE





UPPSALA
UNIVERSITET

Twitter

- Tweets samt kommentarer, likes/hjärtan, retweets, lista över följare och följer
- Har gjorts förut
- Twitter hyfsat öppna för nedladdning





UPPSALA
UNIVERSITET

Facebook

- Statusuppdateringar, kommentarer, likes, delningar
- Facebook gör vad de kan för att förhindra nedladdningar





UPPSALA
UNIVERSITET

YouTube

- Vi har vid två tillfällen laddat ned allt som finns på YouTube-kanalerna
- Metadata:
 - Views
 - Likes
 - Kommentarer
 - Längd
 - datum
- Ännu inget script för regelbunden nedladdning



UPPSALA
UNIVERSITET

Fas 4: Indexering

- Indexeringen bygger på content analysis/thesaurus/ordlistor





UPPSALA
UNIVERSITET

Alla kommentarer är välkomna!

www.statsvet.uu.se/digifacket

