

WEB AS DATA

CHALLENGES AND TRIUMPHS OF CREATING AND WORKING WITH A DERIVED WEB CORPUS

Ditte Laursen, Ph.D, Head of Department, Royal Danish Library
Niels Brügger, Professor, Media Studies, Aarhus University



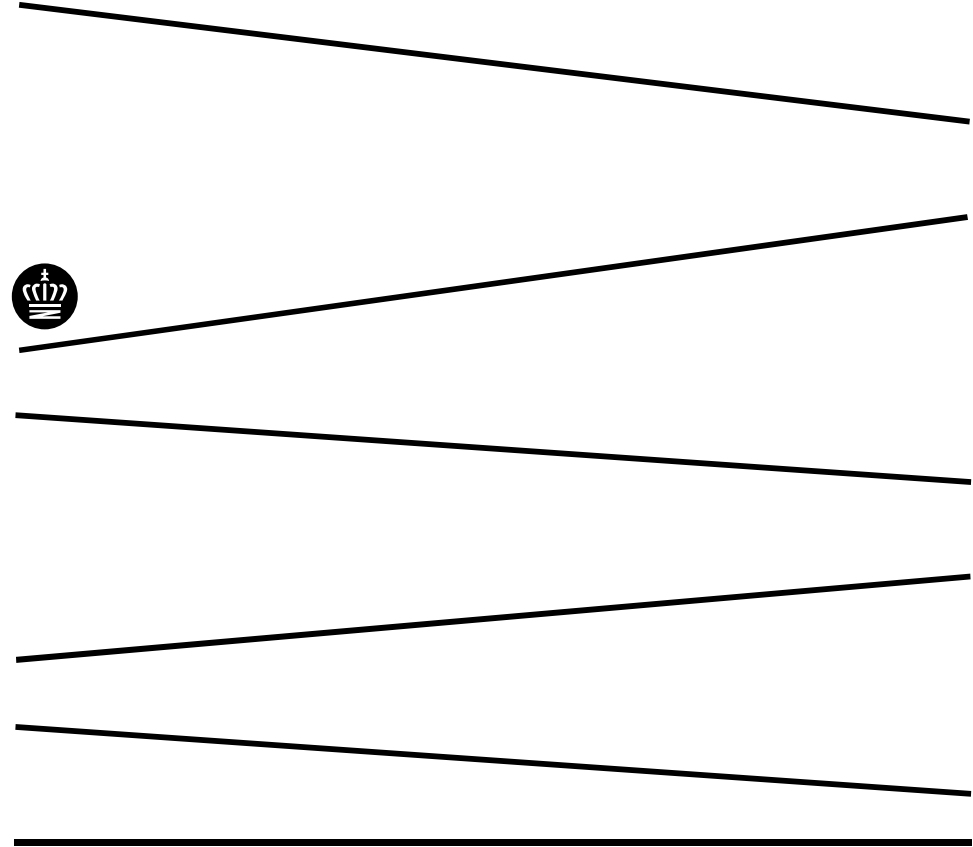
**DET KGL.
BIBLIOTEK**
Royal Danish Library



AARHUS UNIVERSITY



INTRODUCTION



**DET KGL.
BIBLIOTEK**
Royal Danish Library

 AARHUS
UNIVERSITY

Antoinette Fage-Butler & Loni Ledderer:
Working with the “mHealth in
Denmark” Web archive project: Perspectives
on challenges, solutions and
triumphs in a transdisciplinary study

Tanja Svarre: Working with the web archive
project: Digital connectivity and
terminology in the Danish public sector

Mette Skov: Characteristics of Danish
Museum Websites

INTRODUCTION



**DET KGL.
BIBLIOTEK**
Royal Danish Library

 AARHUS
UNIVERSITY

Antoinette Fage-Butler & Loni Ledderer:
Working with the “mHealth in
Denmark” Web archive project: Perspectives
on challenges, solutions and
triumphs in a transdisciplinary study

Tanja Svarre: Working with the web archive
project: Digital connectivity and
terminology in the Danish public sector

Mette Skov: Characteristics of Danish
Museum Websites

INTRODUCTION



**DET KGL.
BIBLIOTEK**
Royal Danish Library

 AARHUS
UNIVERSITY

Antoinette Fage-Butler & Loni Ledderer:
Working with the “mHealth in
Denmark” Web archive project: Perspectives
on challenges, solutions and
triumphs in a transdisciplinary study

Tanja Svarre: Working with the web archive
project: Digital connectivity and
terminology in the Danish public sector

Mette Skov: Characteristics of Danish
Museum Websites

INTRODUCTION
3 PRESENTATIONS
DISCUSSION



**DET KGL.
BIBLIOTEK**
Royal Danish Library

 AARHUS
UNIVERSITY

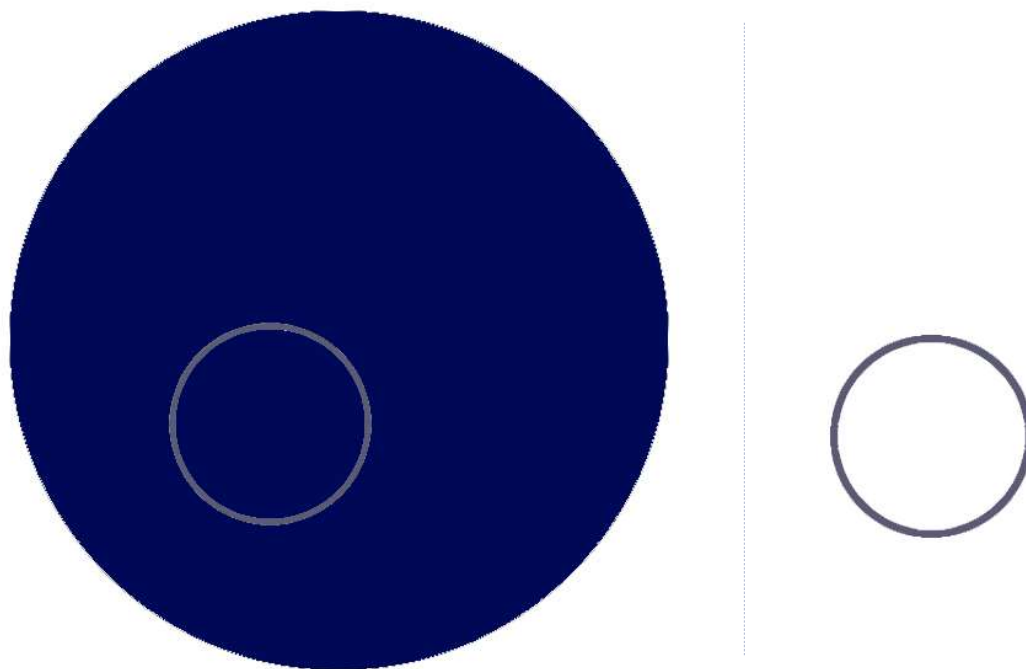
THE DANISH WEB ARCHIVE



- collect the web as part of the country's legal deposit legislation
 - public websites, also with payment or login
 - websites on the .dk TLD
 - websites minded on a Danish audience / written in Danish
 - websites about Danish people
 - more or less any site of interest to Denmark
- total 878 TB (2000-2021)



WEB CORPUS



RESEARCH INFRASTRUCTURES

Library

Search Results for <http://www.netlab.dk/>

Year Range	Pages
jan 1996 - dec 1997	0 pages
jan 1998 - dec 1999	0 pages
jan 2000 - dec 2001	0 pages
jan 2002 - dec 2003	0 pages
jan 2004 - dec 2005	1 page
jan 2006 - dec 2007	4 pages
jan 2008 - dec 2009	3 pages
jan 2010 - dec 2011	3 pages
jan 2012 - dec 2013	6 pages
jan 2014 - dec 2015	8 pages

Home | Help

All Fields

"Metroen giver overskud"

Limit your search

Crawl Year

Domain

Content Type

Public Suffix

You searched for: "Metroen giver overskud"

« Previous | 1 - 25 of 4,530 | Next »

- [http://epn.dk/handel_service/transport_logistik/article1054414.ece.html,2007-09-01T20:45:48Z,index,epn.dk - Metroen giver overskud,](http://epn.dk/handel_service/transport_logistik/article1054414.ece.html,2007-09-01T20:45:48Z,index,epn.dk-Metroen%20giver%20overskud)
- [http://epn.dk/handel_service/transport_logistik/article1054414.ece.html,2007-08-25T21:42:44Z,index,epn.dk - Metroen giver overskud,](http://epn.dk/handel_service/transport_logistik/article1054414.ece.html,2007-08-25T21:42:44Z,index,epn.dk-Metroen%20giver%20overskud)
- http://epn.dk/handel_service/transport_logistik/article1054414.ece

University



CHALLENGES



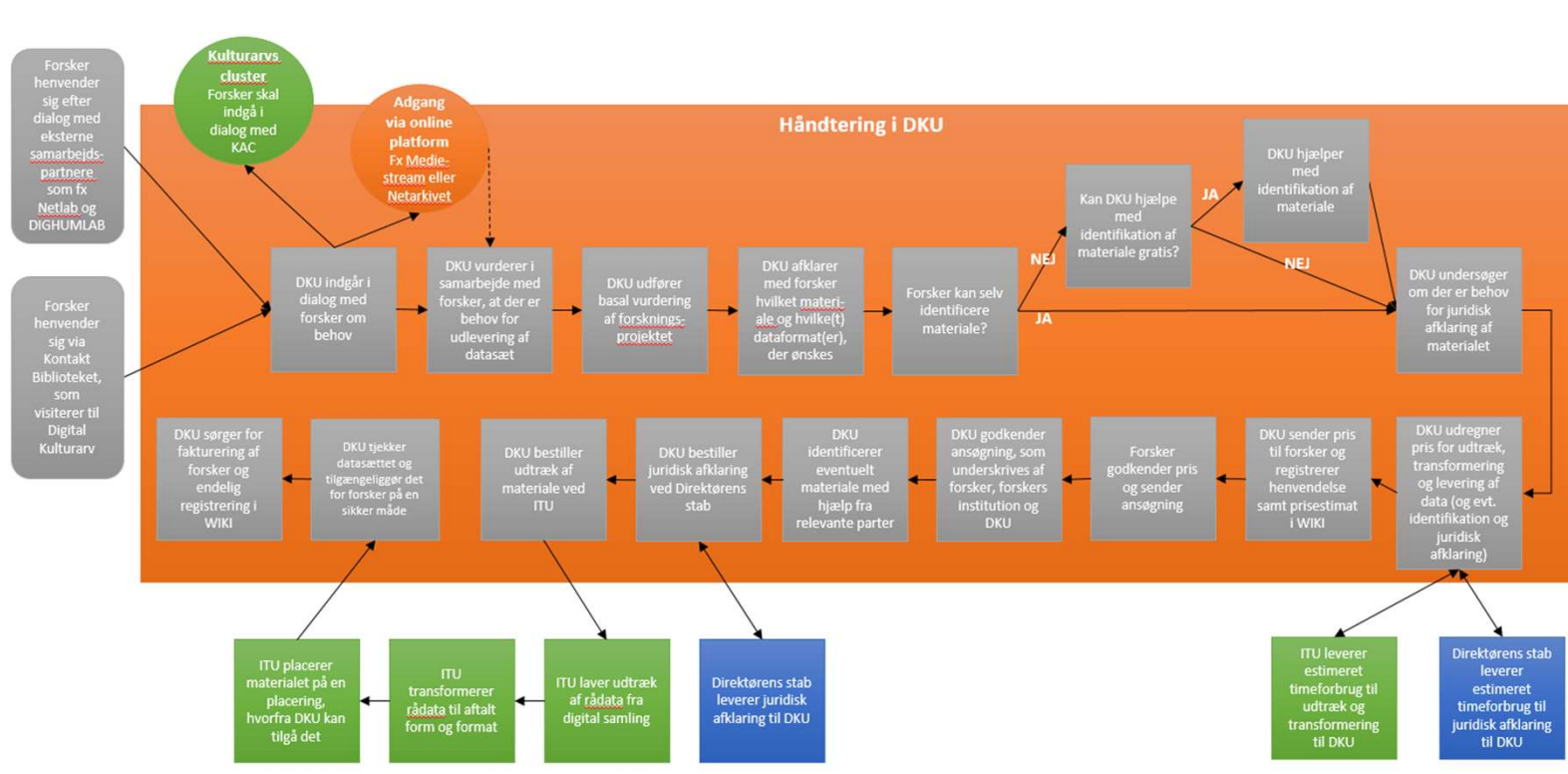
Librarian



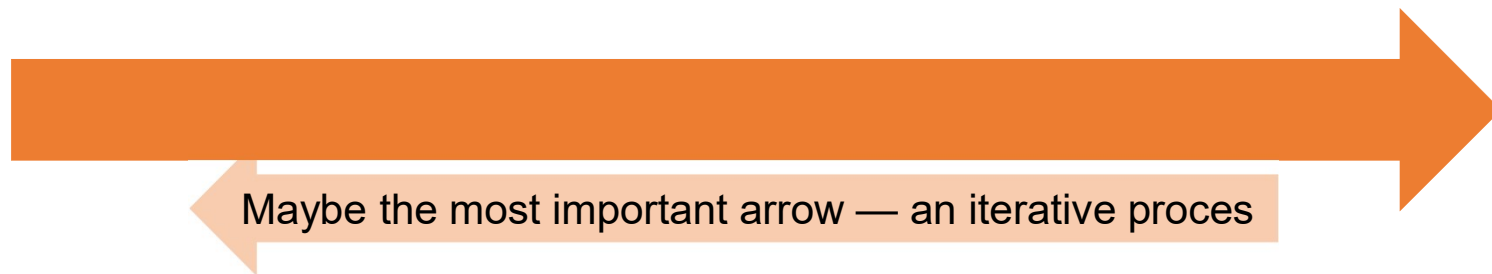
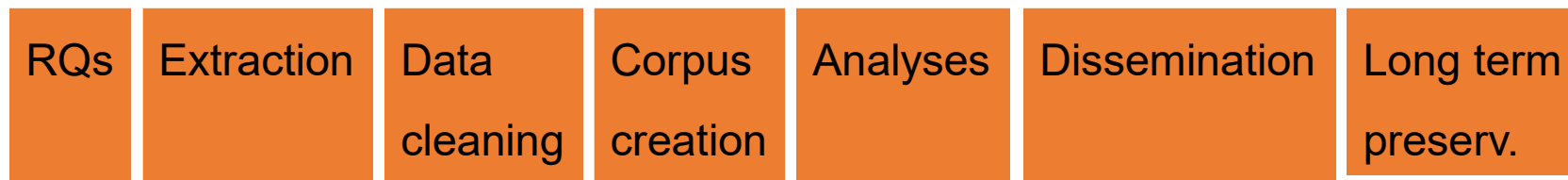
Researcher



LIBRARY WORKFLOW



RESEARCH WORKFLOW



CHALLENGES

Service level?

Competences?

Extraction tools?



Librarian

Research question?

Storage?

Requirement specification?



Researcher



COLLABORATION



WEB AS DATA

CHALLENGES AND TRIUMPHS OF CREATING AND WORKING WITH A DERIVED WEB CORPUS

Ditte Laursen, Ph.D, Head of Department, Royal Danish Library
Niels Brügger, Professor, Media Studies, Aarhus University



**DET KGL.
BIBLIOTEK**
Royal Danish Library



AARHUS UNIVERSITY