Taking stock, new ideas, next steps

WARCnet kickoff meeting
Aarhus, 4-6 May 2020

PROGRAMME

WARCnet
web arhice studies

Under the heading 'Taking stock, new ideas, next steps' the WARCnet kickoff meeting aims at:

o   Getting an overview of where we stand with web archive based research: What can web archives offer for the types of studies we would like to do? What do we have of studies (past, ongoing) about national webs, and events? Which tools are available? To what an extent can we manage the archived web material to prepare it for our studies, in each web archive and across web archives?

o   Having us think about new ideas about the points above, and about what would be the next steps, in general and within this project

Therefore, within each of the sessions there are two types of presentations:

o   Presentations from participants who have been within the field for a longer period of time, these presentations will mainly focus on the first point above

o   Brief presentations from participants who are new to the field, or external participants, and they will primarily focus on the second point

Each session is introduced by a brief overview of some of the topics that could be on the agenda in the session. Finally, each session is rounded off by a brief discussion of what the Working Group could be doing, based on the presentations and the already formulated tasks.

In this way the project will be launched by looking back on the state of the art, as well as having old and new eyes on what we should do.

Practicalities

Venue: online meeting in Zoom — when opening Zoom you will be prompted 'Join with computer audio'. Time zones (calculated by using worldtimebuddy.com): Aarhus (CEST) 12:00 — Dublin/London 11:00 Minnesota 05:00 — Waterloo, 06:00. All times in the program are CEST, please do the calculation yourself to match your time zone.

All presentations will be either a video recording or a slideshow with voice-over, except for the following sessions that will be live only: 'Introduction to the network and to how we can work together' (Monday 11:00 12:15), 'Business Meeting' and 'SG Meeting' (Wednesday 13:00-14.00 and 14:00-14:30). When presentations are shared from the hosts screen you will see this in full-screen mode. In case you want only a smaller window (e.g. if you want to use the chat feature), just press <esc> and the full-screen mode stops.

Chairing, the brief session introductions, and discussions/Q&A will also be live. If you wish to speak please use the chat to tell the chair, use an exclamation mark. Remember to mute your microphone when not speaking. All microphones will be muted by the host during presentations.

We will open for comments, Q&A, and discussion after every 2-3 presentations.

Feel free to use the chat during presentations and during the breaks; it is possible to have a one-to-one chat, but if you want to chat via video in Zoom during breaks, just ask the host to set up a breakout room with the people who should participate.

You may have your video camera on all the time, but if we experience that the video conference software becomes too slow we will ask you to switch it off.

Keynotes are open to a wider public, therefore people from outside the network will be joining these sessions.

All prerecorded presentations can be found in the network's Dropbox to which you have access.

We have a backup plan ready in case Zoom does not work well, you will receive an email with the relevant information in case we change platform.

If you experience problems with the video conference software we will do whatever we can to help you out.

You can contact us here:

Email: warcnet@cc.au.dk

Phone: +45 6015 9738 or +45 2893 0184

Skype: julle.05@hotmail.com or rosenbergkamilla@hotmail.com

Chat in Zoom

Monday 4 May

10:00-10:45
Help desk
Kamilla Rosenberg, Student Helper, Aarhus University
Julie Brøndum, Student Helper, Aarhus University
Niels Brügger, Aarhus University

See contact information above.

Please log on in good time.

10:45-11:00
Welcome
Niels Brügger, Aarhus University

11:00-12:15
Introduction to the network and to how we can work together
Niels Brügger and all participants

Presentation of network members ('knowledge exchange'), the General Guidelines explained, other practicalities, discussion of how we can work together

12:15-13.00
Lunch

13:00-14:15
Web archives

Chair: Valérie Schafer

o   Introduction: Niels Brügger, Aarhus University, 5 min.

o   Bibliothèque nationale de France: Archives de l'internet, the French web archive, Alexandre Chautemps, Bibliothèque nationale de France, France, 10 min.

o   Web archiving without a legal deposit: the Dutch webarchive of KB-NL, Kees Teszelszky, KB-NL, 10 min.

o   From linear to non-linear broadcast contents: considering an "augmented audiovisual archive", France, Eleonore Alquier, Institut Nationale de l'Audiovisuel, 10 min.

o   The Luxembourg Web Archive, Ben Els, Luxembourg Web Archive, Bibliothèque nationale du Luxembourg, Luxembourg, 10 min.

Discussion/Q&A after every second presentation

14:15-14:45
Coffee break

14:45-16:00
Keynote: Web archives as a critical method for the future of digital research
Matthew S. Weber, Associate Professor in the Hubbard School of Journalism and Mass
Communication at the University of Minnesota

Chair: Jane Winters

Discussion/Q&A after the keynote.

(Open for audience outside network)

16:00-
Unwind online

The is just an informal experiment, but I will be drinking a beer and having some snacks, and I suggest that
those of us who are willing can get together online to drink, chat and socialise for half an hour or so — I have
no idea if this works, and don't feel obliged to join in.

Tuesday 5 May

9.00-10:15

Web archives (continued)

Chair: Valérie Schafer

o Netarkivet, The Royal Danish Library, Ditte Laursen, The Royal Danish Library,10 min.

o The promise of web archiving in Belgium: a closer look at (research) user needs, Eveline Vlassenroot, imec-mict-UGent, Belgium, 10 min.

o The UK Web Archive, The British Library, Jason Webber, The British Library, 10 min.

o Web Archives: An overview of developments in Ireland, Sharon Healy & Michael Kurzmeier, Maynooth University, 5 min.

o Web Archives and Digital Humanities, Katharina Schmid, Bayerischer Staatsbibliothek, 5 min.

o Data Friction, Context and the 'Allure of Computation' in the UK Web Archive, Jessica Ogden, University of Southhampton, 5 min.

o April Fools in the Web, Caroline Nyvang, The Royal Danish Library, 5 min.

Discussion/Q&A after every second/third presentation

10:15-10:30

Coffee break

10:30-12:15

WG1: Comparing entire web domains

Chair: Anne Helmond

o Introduction: Niels Brügger, Aarhus University, 5 min.

o Beyond commensurability: developing a framework for comparative web archive research?, Anat Ben-David, The Open University of Israel, 10 min.

o Negotiating the archives of UK web space, Jane Winters, University of London, 10 min.

o Contextualizing and engaging with Web domains, Valérie Schafer, C2DH, University of Luxembourg, 10 min.

o Studies of the Danish web using Netarkivet, Janne Nielsen, Aarhus University, 10 min.

o Definition and Representation of National Web Domains across Web archives, Eld Zierau, The Royal Danish Library, 10 min.

o Preservation of EU websites: Web archiving at the Publications Office of the EU, Els Breedstraet, Publications Office of the EU, 5 min.

o Networks of Power: The Political Economy of Historical Web Infrastructures, Signe Sophus Lai & Sofie Flensburg, the University of Copenhagen, 5 min.

Discussion/Q&A after every second/third presentation

12:15-13.00
Lunch

13.00-14.15

WG2: Analysing transnational events

Chair: Jane Winters

o   Introduction: Niels Brügger, Aarhus University, 5 min.

o   Studying European events: feedback on a pedagogical experiment, Valérie Schafer, C2DH, University of Luxembourg, 10 min.

o   A multiscalar approach: the case of past uses online since the 2000s, Sophie Gebeil, Aix-Marseille Univ (TELEMME laboratory), 10 min.

o   IIPC activities, Olga Holownia, IIPC, 10 min.

o   The Shapes of Archives and Memory, Helle Strandgaard Jensen, Aarhus University, 5 min.

o   Museums on the Web, Karin de Wild, LUCAS, University of Leiden, The Netherlands, 5 min.

Discussion/Q&A after every second/third presentation

14:15-14:45
Coffee break

14:45-16:00

Keynote: You shouldn't need to be a web historian to use web archives: Lowering barriers to access through community and infrastructure
Ian Milligan, Associate Professor, Department of History at University of Waterloo

Chair: Valérie Schafer

Discussion/Q&A after the keynote

(Open for audience outside network)

16.00-
Unwind online — if this worked yesterday

Wednesday 6 May

9:15-11:00

WG3: Digital research methods and tools

Chair: Sally Chambers

o   Introduction: Niels Brügger, Aarhus University, 5 min.

o   Ina and BnF web archive tools for researchers, Sophie Gebeil, Aix-Marseille Univ (TELEMME laboratory) (on BNF and INA tools), 10 min.

o   Digital Methods for Web History, Anne Helmond, University of Amsterdam, 10 min.

o   Large-scale studies using Netarkivet's HPC 'Deic National Cultural Heritage Cluster', Janne Nielsen, Aarhus University, 10 min.

o   Methods for Retrospective Web Archiving, Anat Ben-David, The Open University of Israel, 10 min.

o   Deriving Specific Small-Scale Collections from Web Archives, Federico Nanni, The Alan Turing Institute, 10 min.

o   Hand me the data! What you should know as a humanities researcher before asking for data from a web archive, Ulrich Have, NetLab/DIGHUMLAB, Aarhus University, 10 min.

o   Overwhelmed by Code: studying professional web designers through web archives, Derren Wilson, Manchester Metropolitan University, 5 min.

o   Engaging Students in web archival research, Max Odsbjerg Pedersen, The Royal Danish Library, Denmark, 5 min.

Discussion/Q&A after every second/third presentation

11:00-11:15
Coffee break

11:15-12:15

WG4: Research Data Management across borders

Chair: Ditte Laursen

o   Introduction: Niels Brügger, Aarhus University, 5 min.

o   When the left hand does not know what the right hand is doing: research data management for Web archive studies, Peter Webster, Webster Research and Consulting, 10 min.

o   Web-archives for Open Science: How FAIR can we go? Sally Chambers, Ghent Centre for Digital Humanities, 10 min.

o   Research data Management across borders: Challenges and possible solutions, Niels Brügger, Aarhus University, 10 min.

o   All participants

Discussion/Q&A

12:15-13.00
Lunch

13:00-14.00

Business meeting

Chair: Niels Brügger, Aarhus University

Electing SG members, WG Chairs, and members of Short-Term Network Stays Review Committee.

Looking ahead to the next project meeting

14:00-14:30
SG meeting (separate meeting)

Abstracts

Monday 4 May

13:00-14:15 ⊙

Web archives

Bibliothèque nationale de France: Archives de l'internet, the French web archive, Bibliothèque nationale de France, France, Alexandre Chautemps
The BnF has performed experimental web crawlings since the early 2000s and has been harvesting the French web in the scope of the legal deposit since 2006. Our collection's size is about 1.2 peta octets. Our most ancient archived web pages date back to 1996, and the collection grows daily. The French web archive is constituted by both annual domain crawls and focused crawls of sites selected by librarians and partners. BnF has been involved in several research projects making use of web archive collection and is attentive to meet the scholar's needs. International collaboration is also one of the BnF's priorities, at once in IIPC, RESAW and WARCnet frameworks.

Web archiving without a legal deposit: the Dutch webarchive of KB-NL, Kees Teszelszky
KB-NL has collected born digital material from the web since 2007 through web archiving. It makes a selection of websites with cultural and academic content from the Dutch national web. Most of the sites were harvested because of their value as cultural heritage of the Netherlands, including Frisian heritage, language and culture. I will introduce the content of the KB-NL web archive and will present the preliminary results and the research potential of mapping, harvesting and creating a web data set out of the Dutch national domain by KB-NL.

From linear to non-linear broadcast contents: considering an "augmented audiovisual archive", Institut Nationale de l'Audiovisuel, France, Eleonore Alquier
The French Audiovisual Institute (INA) has mission to collect, preserve, restore and communicate France's radio and television heritage, which was extended in 2006 to online public web contents. This submission aims to present the evolution of (traditionally linear) audiovisual archiving methods, granting the necessity to consider related web contents: why and how broadcasters tend to create non-linear contents, impacts of these new practices on collecting, documenting, curating and accessing this so-called "augmented archive".

The Luxembourg Web Archive, Luxembourg Web Archive, Bibliothèque nationale du Luxembourg, Luxembourg, Ben Els
The presentation will introduce some basic information and key data about the Luxembourg Web Archive at the National Library of Luxembourg: Our team, when and how the program was started, the Luxembourg legal deposit and our collection scope, what we have collected so far and our plans for the near future.

14:45-16:00 ⏲

Keynote: Web archives as a critical method for the future of digital research

Matthew S. Weber, Associate Professor in the Hubbard School of Journalism and Mass Communication at the University of Minnesota

Web archiving is often viewed as a passive activity for researchers. The normative workflow sees researchers turning to libraries and repositories for access on an as needed basis and then delving deep into the world of analysis. There is a lot that can be learned from this type of scholarship, and yet the idea of web archives is increasingly a misnomer in that you do not need to be looking backwards or focusing solely on past events. This talk focuses on the ways in which web archives are being utilized to develop new streams of research, including the use of web archives as an observatory for studying modern social phenomena. I will focus on a number of specific use cases highlighting the implementation of web archives to study local news media and modern political campaigns, and will conclude by discussing new workflows and processes for using web archives in digital humanities and social science research.

Tuesday 5 May

9.00-10:15 ⏱

Web archives (continued)

Netarkivet, The Royal Danish Library, Ditte Laursen

The Danish national Web archive Netarkivet was established in 2005, and Denmark became one of the first countries in the world that mandated legal deposit of dynamic web materials. This presentation will give a historical overview of the types of research conducted in the archive, advanced over time in parallel with the archive's continued effort to widen access and to develop new tools and services. The importance of user-driven and experience-led service developments will be emphasized as well as the archive's challenges in balancing generic services with individual research needs.

The promise of web archiving in Belgium: a closer look at (research) user needs, imec-mict-UGent, Eveline Vlassenroot, Belgium

Despite its long history, the Belgian web is currently not systematically archived. Without a Belgian web archive there is a significant risk that essential born digital resources for historians will be lost forever. PROMISE, a research project, which investigates the feasibility of developing a sustainable web archiving service for Belgium, is introduced. The results of this web archive pilot are presented and the importance of research use of the future Belgian web archive is stressed.

The UK Web Archive, The British Library, Jason Webber

This presentation will introduce the British Library's UK Web Archive, which harvests and preserves the .uk CCTLD under non-print legal deposit legislation. At present the vast majority of the archived websites can only be accessed on-site in the UK's six legal deposit libraries, but parts of the collection are available on an open-access basis. Current modes of access to the archive, and its potential for research, will be briefly considered.

Web Archives: An overview of developments in Ireland, Sharon Healy & Michael Kurzmeier, Maynooth University

This presentation provides an overview of the development of web archives in Ireland, and offers a glimpse of levels of awareness and engagement with web archives in Irish academic institutions.

Web Archives and Digital Humanities, Katharina Schmid, Bayerischer Staatsbibliothek

In a project supported by the German Research Foundation, the Munich DigitiZation Center at the Bavarian State Library and the University of Passau are exploring how methods and tools from the Digital Humanities can be applied to web archives. We have conducted an event crawl for the European elections and extracted data from the collection with the Archives Unleashed Toolkit and a new tool, warc2corpus. Based on this data, we aim to perform analyses like Latent Dirichlet Allocation and Structural Topic Modeling to approach research questions from the political sciences.

Data Friction, Context and the 'Allure of Computation' in the UK Web Archive, Jessica Ogden, University of Southhampton

This presentation is based on a three-month placement at the British Library researching the challenges of using the UK Web Archive for social science research. Using the Hyperlocal News collection, I map curator/crawling activities and develop a conceptual framework that charts researcher practices associated with using the UKWA. Finally, three observations are made to reflect the ways that the web archives present epistemological challenges for social science researchers.

April Fools in the Web, Caroline Nyvang, The Royal Danish Library

A brief introduction to a special harvest of web-based April Fools jokes. The harvest was conducted in 2019 and 2020 at The Royal Danish Library by means of the BCweb application (developed by BnF)

10:30-12:15 🕐

WG1: Comparing entire web domains

Beyond commensurability: developing a framework for comparative web archive research?, Anat Ben-David, The Open University of Israel

Entire archives of national web domains differ in their scope, depth and time span. Quantitative analyses of entire national web domains may facilitate the characterization of each domain, but these structural differences raise commensurability questions for cross-domain analyses and historiographies. Studies in comparative histories address the problem of commensurability when comparing the histories of entire societies. This presentation asks what lessons can be learned from comparative history for developing a theoretical framework for comparative web archive research?

Negotiating the archives of UK web space, Jane Winters, University of London

Web archives, like their more traditional counterparts, reflect the social, political, cultural and economic contexts within which they are formed. And these contexts are neither stable nor purely national; they are influenced over time by changes in legislation, by shifting organisational priorities, by fluctuations in funding, by new governments, even by the interests of individuals. This presentation will discuss what this has meant, and continues to mean, for the archives of UK web space.

Contextualizing and engaging with Web domains, Valérie Schafer, C2DH, University of Luxembourg

This presentation will suggest concrete avenues both to contextualize research on domain names (e.g. collaboration with the Afnic in France and use of its web archives to retrieve the price of a domain name, its policy for allocating domain names, the number of domain names it allocated with regards to the domain names retrieved in web archives, etc.) but also to create public engagement with this topic (identifying the first web page in each country, the most represented twitter accounts in web archives, etc.).

Studies of the Danish web using Netarkivet, Janne Nielsen, Aarhus University

The presentation will introduce the project Probing a nation's web domain about the historical development of the Danish web as an example of how a national web domain can be studied. The presentation will also briefly address some methodological aspects of the study and highlight some possibilities and challenges if such a project should be done

comparatively across different national webs.

Definition and Representation of National Web Domains across Web archives, Eld Zierau, The Royal Danish Library

This presentation is based on results presented at RESAW 2015 and RESAW 2017 about challenges and solution of how to define national web domains and how we can represent them in a standardized reusable and preservable way and with possibility to include elements from several web archives.

Preservation of EU websites: Web archiving at the Publications Office of the EU, Els Breedstraet, Publications Office of the EU

This presentation gives in one slide a short overview of the Publications Office of the EU (OP)'s web preservation service. In the slides following this overview, more information on the OP and its archiving and long-term preservation service can be found, as well as contact details of the author and of the OP's web archiving team. Feel also free to contact the author via twitter @breedel_els or Linkedin: https://www.linkedin.com/in/elsbreedstraet

Networks of Power: The Political Economy of Historical Web Infrastructures, Signe Sophus Lai & Sofie Flensburg, the University of Copenhagen

The paper studies the evolution of the web infrastructure in Denmark and assesses the changing ownership structures of websites and third-party trackers across three years (2002, 2012, and 2019). We build on and develop Thomas P. Hughes' momentum-theory (1994) and argue that the internet, as other large technological systems, has evolved in different phases reflecting a shift from being mainly influenced by socio-cultural conditions to having a determining influence on the development of societal structures.

13.00-14:15 ⏲

WG2: Analysing transnational events

Studying European events: feedback on a pedagogical experiment, Valérie Schafer, C2DH, University of Luxembourg

Based on the experience of the Winter School I conducted for 2 years on Web archives at the university of Luxembourg and which involved research on European events and Europan history by Master's students, I will draw some methodological conclusions (interesting topics, difficulties in grasping participation, typology of transnational events with regard to Web archives, issues of visual studies and transmedia studies...).

A multiscalar approach: the case of past uses online since the 2000s, Sophie Gebeil, Aix-Marseille Univ (TELEMME laboratory)

In this short paper, I would like to defend the fact that a corpus belonging to a national web sphere can be studied according to a multiscalar approach, questioning the phenomenon in the light of local, even micro-local, national and global dynamics. We will thus see that, if online memory uses are, like any other practice or event observable online, historically and socially situated, they are also geographically situated. Thus, cross-referencing these different scales forces the researcher to vary the methodologies while sometimes being confronted with a lack of sources depending on web archiving.

IIPC activities, Olga Holownia, IIPC

Since its foundation in 2003, one of the goals of the International Internet Preservation Consortium (IIPC) has been to provide a forum for the sharing of knowledge about web archiving. In addition to the annual conference and general assembly, the IIPC community has been actively collaborating through working groups, projects and hackathons. This presentation will provide an update on current and planned IIPC activities with a special focus on funded projects, research initiatives, training materials, and the ongoing COVID-19 collection led by the IIPC Content Development Working Group and organised in partnership with Archive-It.

The Shapes of Archives and Memory, Helle Strandgaard Jensen, Aarhus University

I would like to compare the history of two archival institutions, The National Archives (UK) and the Danish National Archives, c. 1995 2020 as it played out on the Web. Both institutions have developed a vast verity of online activities for their users since the late 1990s. One part of the study will compare the history of two institutions' websites and their different activities. The other will focus on how the ways the websites were archived (by the British Library, the Danish Royal Library and the Internet Archive) shape the possibilities for researching them, thereby speaking to a broader debate about how archival practices influence understandings of the past.

Museums on the Web, Karin de Wild, LUCAS, University of Leiden, The Netherlands

This presentation will introduce my upcoming research about the history of museums on the Web, from the rise of pioneering virtual museums and their online collections, to how the museums' presence on the Web evolved over time increasingly tying connections on a global scale. It will combine methods from history and data science and draw on a computational analysis of Web archives and digitized museum collections.

14:45-16:00 ⏱

Keynote: You shouldn't need to be a web historian to use web archives: Lowering barriers to access through community and infrastructure

Ian Milligan, Associate Professor, Department of History at University of Waterloo

If a researcher currently wants to use web archives at scale, they need to be a significant focus of their research activity --- in other words, a web historian really can't just dabble with these sources. Yet historians need to be able to draw on wide varieties of sources for their projects (from archival records to newspaper records to oral histories and beyond). In this talk, I explore how we are developing tools to lower these barriers so that a historian really could use web history data without a significant investment of time. I'll begin by discussing the current situation of working with web archives, and then highlight the Archives Unleashed project, datathons, and other ways in which cultural organizations are making web archives accessible.

Wednesday 6 May

9:15-11:00 🕐

WG3: Digital research methods and tools

Ina and BnF web archive tools for researchers, Sophie Gebeil, Aix-Marseille Univ (TELEMME laboratory) (on BNF and INA tools)

This short presentation will focus on how the BnF and Ina web archive tools can help researchers in their exploration and their analysis. I will first present the tools to search information inside the web archived. Then I will present the innovative tools that can be mobilized by our project in connection with the two institutions.

Digital Methods for Web History, Anne Helmond, University of Amsterdam

This talk will provide a short overview of the tools that have been developed at the Digital Methods Initiative at the University of Amsterdam. In addition, it will reflect on best practices for developing tools in the future by considering 1) the purpose and audience of the tools, 2) the technical environment of the tools, and the 3) use case scenarios of the tools. Finally, I would like to raise the idea of creating a tool repository, to consider applying for grants for the development and maintenance of tools, and possibly create worksheets for students and researchers to work with tools in the repository.

Large-scale studies using Netarkivet's HPC 'Deic National Cultural Heritage Cluster', Janne Nielsen, Aarhus University

The presentation will outline the approaches and tools used in the project Probing a nation's web domain, including points about collaboration (IT-developers, curators and researchers), infrastructure and access, corpus creation, scripts and possibilities for sharing tools/methods.

Methods for Retrospective Web Archiving, Anat Ben-David, The Open University of Israel

This presentation introduces Retrospective Web Archiving as a method that allows web archiving institutions and researchers to archive past events that have not been archived in real-time. Apart from filling-in existing gaps in web archives, RWA allows expanding the boundaries of web archiving from the open web to cross-platform curation, and from national to international perspectives. The presentation outlines the theoretical principles and the methodological pipeline of RWA, using examples from a study on the 2014 War in Gaza.

Deriving Specific Small-Scale Collections from Web Archives, Federico Nanni, The Alan Turing Institute

I will offer an overview of the research I conducted during the last few years at the intersection of information retrieval and digital libraries. In particular, I will focus on approaches for the automatic creation of fine-grained collections from large-scale web archives, in order to support research in digital humanities and computational social sciences concerning specific events, entities and topics.

Hand me the data! What you should know as a humanities researcher before asking for data from a web archive, Ulrich Have, NetLab/DIGHUMLAB, Aarhus University
The mHealth project was among the first projects to have data delivered from the Danish Web Archive to University researchers for exploratory analysis. Which challenges and opportunities did the project face? How could other humanities scholars be inspired by this account?

Overwhelmed by Code: studying professional web designers through web archives, Derren Wilson, Manchester Metropolitan University
Responsive Web Design (RWD) is a web design process created after platformised touch screen internet devices became available. Qualitative data tracing this change is present in archives from 2007, but telling the full story of this period seems to require quantitative data and born-digital artefacts which may not be currently collected. Obsolete industry practices also complicate studying archived websites from this period.

Engaging Students in web archival research, Max Odsbjerg Pedersen, The Royal Danish Library, Denmark
A lot of national web archives and a lot of the methods employed in web research is out of reach for most students. The Archives themselves because of legal restrictions and the methods because the gap between the technical requirements of web research methods and the student's level of technical capacity. How do we ensure that students are intrigued by web research and eventual advance as web researches? Do the students need to reach the fields technical level or should the field try to meet the students at their level?

11:15-12:15 ⏱

WG4: Research Data Management across borders

When the left hand does not know what the right hand is doing: research data management for Web archive studies, Peter Webster, Webster Research and Consulting
Based on observations of the situation in the UK, I examine briefly the options for RDM for Web archives in particular. I suggest that there is a significant disjuncture between several different areas of policy-making and infrastructure provision: research data repositories for the hard sciences; publication repositories and data journals for the humanities and social sciences; the rapidly expanding provision of data from within GLAM institutions, including those that traditionally manage Web archives. I propose no easy answers, but point out the risk of an unintegrated approach, both within each individual country and worldwide.

Web-archives for Open Science: How FAIR can we go? Sally Chambers, Ghent Centre for Digital Humanities
Ensuring our research data is Findable, Accessible, Interoperable and Reusable is a core principle of Open Science. Yet, when it comes to research access to archived web resources and the resulting research data, how FAIR is our data? This presentation will provide an initial assessment of the FAIRness of our data, along with some initial suggestions as to how we can embed the FAIR principles of Research Data Management into the everyday work of WARCnet.

Research data Management across borders: Challenges and possible solutions, Niels Brügger, Aarhus University
On a national level Research Data Management (RDM) within web archive studies is still in its infancy, let alone transnational RDM. This presentation will outline a systematic approach to discussions of RDM by using an overall framework with the research phases (corpus creation, extraction, data cleaning, analysis, dissemination, and long-term preservation) and three transversal topics (legal

issues, research ethics, and RDM). This framework will be used to ask: how can we manage the data we intend to study in WARCnet across borders?

The aim of the **WARCnet network** is to promote high-quality national and transnational research that will help us to understand the history of (trans)national web domains and of transnational events on the web, drawing on the increasingly important digital cultural heritage held in national web archives. The network activities run in 2020-21, hosted by the School of Communication and Culture at Aarhus University, and are funded by the Independent Research Fund Denmark | Humanities (grant no 9055-00005B).

PROGRAMME

INDEPENDENT
RESEARCH FUND
DENMARK