

Supporting documents

RESAW seminar, Aarhus 2-3 december 2013

This document is a collection of supporting documents for each item in the programme.

The supporting documents does not pretend to give a comprehensive overview of all relevant topics within RESAW. Rather they present each item in the programme, as well as they outline some of the main points to debate. Thus, the supporting documents are considered the starting point of a longer process where the first steps are to find out what we could put into RESAW, and the next step will be the actual drafting of an application (hopefully to Horizon 2020 in 2015/16).

The supporting documents for Thursday are based on:

- the proposal we submitted in October 2012
- the summaries of the two web discussions
- the document 'RESAW seminar and Horizon 2020' (June 2013)
- my own reflections about RESAW, some of which have been presented to some of you in a powerpoint presentation

However, the supporting documents do not take up all points in these documents.

It has to be emphasized that these supporting documents are not an exhaustive catalogue of what can be brought up and debated at the seminar. They are rather thought of as a list of what I had in mind we could discuss. Feel free to bring up other topics in relation to each item.

Niels Brügger, november 2013

Horizon 2020 — what we need to know

Adriënne Heijnen, Aarhus University, Research Support Office

This session aims at giving a short introduction to Horizon 2020 and research infrastructures. The status of the Horizon 2020 calls, the process within the coming years, and how to promote RESAW with a view to becoming part of the next call. The final call is not out before 11 December so the presentation is based on the latest draft.

15 min. presentation, 15 min. comments and QA.

Digital Humanities and Research Infrastructures

Niels Ole Finnemann, NetLab

The aim of this session is to set the scene for some of the general issues related to RESAW's topics. Focus will be on digital humanities and research infrastructure, and focus is on the scholars' point of view.

30 min. presentation, 15 min. comments and QA.

Web archives — interacting with scholars

Helen Hockx-Yu, the British Library

This session has the same aim as the previous, but with focus on how things look from a web archive's point of view.

30 min. presentation, 15 min. comments and QA.

Scholarly use of web archives and existing web archives

The aim of this session is to give an overview of what we already have in the RESAW group, as well as to have the presenters reflect over their possible role(s) in a project like RESAW. The session will consist of eight presentations from, on the one hand, research groups and individual scholars who have used (or are about to use) materials from web archives in their research, and, on the other hand, presentations of existing national web archives.

All presenters have been asked to address the following four issues:

- who am I/are we, and what am I/are we doing, in general?
- ongoing activities specifically related to RESAW's topics?
- what do I/we consider the main purpose, necessity and rationale of RESAW — what is the RESAW project in 10 years?
- what am I doing in RESAW?

15 min. for each presentation; four presentations in a row, then 15 min. for comments and QA. Then the next four presentations, followed by another 15 min. for comments and QA.

Other approaches to web archives and scholarly uses of web archives

The aim of this session is the same as in the previous session, but focus is on other approaches to web archives and scholarly uses of web archives. All presenters have been asked to address the same four issues as in the previous session.

15 min. for each presentation; three presentations in a row, then 10 min. for comments and QA. Then the last three presentations, followed by another 10 min. for comments and QA.

What do we consider the main aims of RESAW?

The aim of this session is to debate what could be the aims of RESAW. The format of the session will be as follows:

- short introduction to the main points below (Niels, 5 min.)
- short statements from Niels Ole Finnemann, NetLab & Bjarne Andersen, Netarkivet (max. 5 min. each)
- open discussion (30 min.)

I have divided my approach to what could be the aims of RESAW into two sections, first some general reflections, second more detailed points to be considered. Some of the points below draw on the proposal we submitted in October 2012 as well as on the 'Consultation on possible topics for future activities for integrating and opening existing national research infrastructures'. The relevant sections of these two documents have been pasted at the end of this document.

General reflections

The overall aim of RESAW is to establish a collaborative European research infrastructure for the study of archived web materials.

The following are my reflections about each semantic entity in the aim as formulated above as well as about the demarcation issues related to each of them.

1) "research infrastructure"

What we are establishing is a research infrastructure (RI), and therefore it gives good sense to discuss what that could be. A good starting point is the definition formulated by ESFRI, the European Strategy Forum on Research Infrastructures. The following is the definition to be found in the EU call for "Integrating Activities" (a short version of ESFRI's definition):

Research infrastructures are defined here as facilities, resources, systems and related services that are used by research communities to conduct top level research in their respective fields.

This definition covers: major scientific equipment or sets of instruments, as well as knowledge-containing resources such as collections, archives and thematic data infrastructures, together with the associated human resources.

Research infrastructures may be “single-sited”, “distributed”, or “virtual” (the service being provided electronically).

Each of the points in this definition could, and should, be specified, but it is a good starting point.

Demarcation issue

Although RESAW is a RI in its own right for the study of archived web materials it gives good sense to investigate and debate if the RI could be associated or affiliated to other existing/coming RIs, for the benefit of both. Affiliating RESAW to the already existing RI DARIAH would be a possibility (DARIAH is an ESFRI RI, and has just achieved ERIC status (European Research Infrastructure Consortium), that is a specific legal form to facilitate the joint establishment and operation of research infrastructures of European interest).

2) "establish" and "archived web materials"

"Establish" could be many things but if we still aim at handing in an application for Horizon 2020 (and this is still my first priority) we have to remember the name of the initial consultation to which we handed in the proposal October 2012: "Consultation on possible topics for future activities for integrating and opening existing national research infrastructures". The important words here are 'integrating and opening existing national research infrastructures'. This means that what the EU supports is the integration of

already existing research infrastructures in each member state. Therefore "establish" means "integrate existing RIs", and "archived web materials" primarily means materials in existing web archives within each nation (national or other). Thus, RESAW is not an entirely new RI, but an integration of already existing RIs in each country with a view to facilitating cross-national research within what the EU system terms the ERA, the European Research Area.

"Establish" also raises another question: for how long? One of the main concerns within the EU system is the long term sustainability of RIs. The degree of commitment to long term funding is of relevance here, and this concerns web archives as well as scholarly communities and the will/ability to support the RI within each of the EU member states.

Demarcation issues

- how to integrate web archiving initiatives which are highly relevant but are not 'national', for instance the Internet Memory Foundation?
- how to integrate other national web archives such as collections in research libraries or micro archiving performed by research groups or individual scholars with a specific research purpose?
- how to handle, relate to, and integrate with other archived digital materials that are not web: a) the early pre-web internet (usenet, bbs, gopher, minitel...), b) the present post-web (mobile, tablets), c) other digitized media (newspapers, books, radio, television)?
- how to handle researcher generated collections accross web archives? This is a general issue relating to data management
- how can RESAW foster outreach to the industry and promote innovation (which companies would be interested? — the market is probably very limited)? In the latest draft of the call it reads: "Integrating Activities in particular should contribute to fostering the potential for innovation, including social innovation, of research infrastructures by reinforcing the partnership with industry, through e.g. transfer of knowledge and other dissemination activities, activities to promote the use of research infrastructures by industrial researchers, involvement of industrial associations in consortia or in advisory bodies. A specific work package on

innovation is therefore recommended in all Integrating Activity proposals”, p. 5

3) "study of"

RESAW is primarily aimed at studies performed by scholars. RESAW is based on the idea that the development of a RI should be research driven, that is take its point of departure in actual ongoing (or planned) studies. There are two reasons for this. First that no archiving of the web is neutral, and the complex impacts of the choices made by the archiving institutions should be debated with the scholars who are going to use the material with a view to improving the web archives and their possible integration. Second, that in the main access is in many countries restricted to researchers.

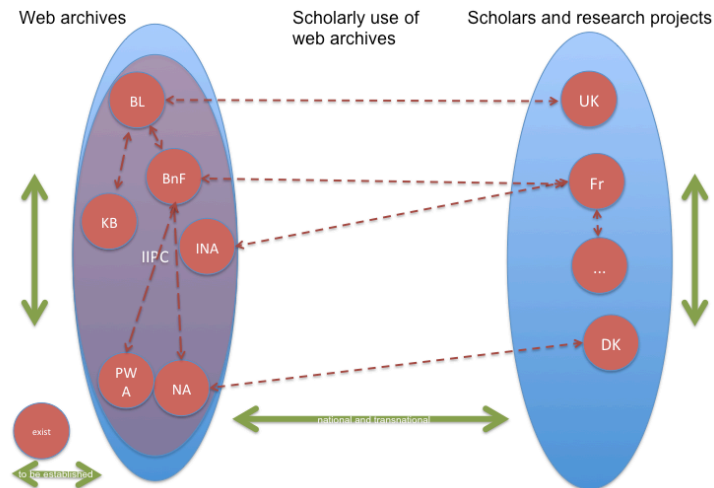
However, RESAW is not focusing on specific types of studies, methods or research communities. With a view to improving the RI for the benefit of as many scholars as possible RESAW aims at integrating as many disciplines and types of scholars and methods as possible.

Demarcation issue

Although web archives in most countries are only open for researchers, some countries have open access for everyone, for instance Portugal and Croatia; how to open RESAW to the wider public where that is possible?

4) "collaborative"

RESAW is unique since it aims at fostering cross-national integration in three dimensions: 1) between existing national web archives, 2) between existing national research clusters working with web archives and with existing/coming joint research projects, and 3) between web archives and research communities across borders.



What is important here is that all three forms of collaborations are there at the same time — none of them can stand alone — and that they ensure that RESAW can become a robust European research infrastructure which takes into account the specific nature of web archives just as it is tailored to fit actual researcher needs.

5) "European"

As the point of departure Horizon 2020 is primarily for EU member states and associated countries. Stakeholders outside Europe may be part of RESAW where they can contribute to the project, but how this can actually take place is not clear yet since the final call is not out yet.

Demarcation issue

- how to relate to the work done within already existing international associations such as the International Internet Preservation Consortium (IIPC)? — RESAW should not establish things that are already established elsewhere
- how to involve existing web archives outside Europe?
- how to integrate scholars outside Europe?

Detailed points

The following are some of the points I think should be taken into consideration when debating the aims of RESAW.

1) Access

Access to the different web archives is a key aim. But access can mean several things. The following figure may help frame the debate of access.

	Technological issues	Legal issues (copyright, privacy)
Locate an URL and see metadata		
Aggregated content		
See content on web pages		
Interact with viewed content (extraction and analytical tools)		

This grid illustrates how two different approaches to access are entangled. The two columns address the enabling/disabling factors for getting trans-national access — technology and legal issue — whereas the four rows identify what the researcher actually can get access to, ranging from a simple localisation of an URL to extracting and analysing the material.

A third approach to access cuts across columns as well as rows, namely the physical site of access: on site or online.

2) Workspace

The establishing of a common workspace could be an aim for RESAW. It could include cross-archive search facilities (URL, full text, other), information about each web archive, documentation (cf. below), possibilities of annotation (for individual scholars, groups; short/long term), analytical tools, help pages, display of use cases, discussion forum, and a 'market place' (looking for researchers in other countries for a cross-national project).

3) Documentation

In general, scholars would probably like to know as much as possible about what they find in a web archive: what can I expect to find/not find in each web archive? Where did it come from? How was it created? Etc. RESAW should aim at setting the standards for 'best practice' with regard to documentation of what is in the different web archives. Documentation can be technical information extracted from crawl logs and the like, as well as curatorial documentation, archiving strategies, known deficiencies, etc.

4) Researcher generated corpora

Studying all material in all European web archives will probably be the exception. Thus, RESAW should aim at establishing procedures as well as technological solutions which support the creation of cross-archival corpora. And a data management policy as to how to handle these corpora should be established (while analysing the material, when the results are under review, and in a long term perspective).

5) Analytical tools

Analytical tools to help analyse the material should be developed. Two issues are involved here.

First, that in some cases the point of departure can be taken in already existing software used for analysing online material, whereas in other cases new software has to be developed. It should be taken into account that for many scholars the website/URL is not necessarily the only analytical unit, this could (also) be smaller elements (images, video) or larger entities (a web sphere).

Second, that the possible applicability of the same analytical tools to different web archives may pose a challenge due to differences in archiving strategy, formats, etc.

As part of the data management policy mentioned above it should be debated if the analytical tools should be preserved as well, where, by whom, and for how long.

Excerpts

The following are excerpts about the aim of RESAW from the proposal we handed in October 2012.

The proposed research infrastructure, RESAW, aims to bridge the gaps between the fast growing number of web archiving institutions, who holds a still more significant part of 21th century cultural heritage, and a fast growing range of research communities and individual researchers for whom these archived materials are of increasing importance. It is proposed that such a research infrastructure should be build around general (not least) national Internet archives holding collections of materials primarily published on the web, in collaboration with an international community of scholars and researchers representing the widest possible range of disciplines and methodological approaches to the study of archived internet materials.

On a more detailed level the purpose is to 1) foster the collaboration and networking activities between the archival institutions and research communities on an European and global scale, 2) develop a coordinated, but distributed workspace for searching, analyzing, and presenting, 3) build the relevant skills for development and use of software supported methods in the study of internet materials across different national web archives, 4) give access to related research data storages, and 5) build strategies for archiving relevant web materials which are not taken care of within existing institutional frameworks, such as materials on the domain .eu and other non-national domains (.net .info .biz .mobi etc.).

The project should be driven by research questions which will lead to results also to be acknowledged by a wider community of scholars and researchers within the relevant fields who are not specialized in studying digital materials by help of software supported methods. Thus, the establishing of a sustainable web research infrastructure should be based on a close interplay between web archiving institutions, the relevant research communities, and the research questions which scholars want to inquire. The relevant research communities come from a wide range of fields (see 5 and 6b below) The great majority of

relevant archiving institutions are members of the “International Internet Preservation Consortium” (IIPC) who today counts more than thirty major and primarily national web archives as members. Twenty of these are national European archives.

RESAW will serve to 1) favour comparative approaches to different archives and interoperability across web archives in different countries, which use different principles for collection, and with holdings in different languages, 2) validate the quality of the various archives due to the different principles and combinations of methods used in building the archives, 3) establish a shared portal for accessing and documenting the web archives, 4) develop software supported methods allowing large scale European comparative research in web materials, 5) initiate investigations as to how other internet activities (e.g. email, apps to smart phones and tablets, facebook data etc.) can be integrated into general internet & web archives, and 6) facilitate online access to the archives and research data, and in so far online access is restricted due to national legislation to facilitate in situ access for researchers.

Furthermore the project should address the long term perspective of coordinating web archives with other sorts of digital collections, including digitized cultural heritage collections whether available online or not. The project is primarily, but not solely, aimed at researchers and scholars within the humanities and social sciences but is obviously also of high relevance for a wider public in modern societies.

The following are excerpts from “Consultation on possible topics for future activities for integrating and opening existing national research infrastructures”, the supporting document to the call for proposals, the section ‘What are Integrating Activities’.

Under FP7, each funded project (typically up to € 10 Million of EU contribution, over 4 years) is a combination of three mandatory types of activities:

a) **Trans-national access and/or service Activities**, to provide trans-national access of researchers or research teams to one or more infrastructures among those operated by participants, and/or to provide access to scientific services freely available through communication networks (e.g. databases available via Internet). Trans-national access may be made available to external users, either in person ("hands-on") or through the provision of remote scientific services, such as the provision of reference materials or samples or the performance of sample analysis. EU financial support should cover only part of the annual operating costs of the infrastructure to prevent it from becoming dependent on the EU contribution and should not include capital investments. This financial support serves to provide access "free of charge" to external users, including all the infrastructural, logistical, technological and scientific support (training courses for using the research infrastructures, and travel and subsistence for users, are supported to).

b) **Networking Activities**, to foster a culture of co-operation between research infrastructures, scientific communities and other key stakeholders, and help developing a more efficient and attractive European Research Area. Examples of activities are (non-exhaustive list): joint management of access; definition of common standards, protocols and interoperability; spreading of good practices, consultancy and training courses to new users; dissemination and/or exploitation of project results and knowledge, outreach toward industry, fostering and measuring contribution to socio-economic impacts, promotion of innovation; strengthening of virtual research communities; development and maintenance of common databases for the purpose of networking and management of the users and infrastructures; foresight studies for new instrumentation, methods, concepts and/or technologies; promotion of clustering and coordinated actions amongst related projects; coordination with national or international related initiatives and support to the deployment of global and sustainable approaches in the field; promotion of long term sustainability, including the involvement of funders and the preparation of a business plan beyond the end of the project; etc.

c) **Joint Research Activities**, to improve, in quality and/or quantity, the services provided by the infrastructures. Examples of activities are (non-exhaustive list): higher performance methodologies and protocols; higher performance instrumentation, including the testing of components, subsystems, materials, techniques and dedicated software; integration of installations and infrastructures into virtual facilities; innovative solutions for data collection, management, curation, annotation, deposition; etc.

As under FP7, future EU actions are expected to:

- Bring together and integrate, on a European scale, key research infrastructures, in order to promote their coordinated use and development;
- Provide the users of research infrastructures with a harmonised, improved and optimised access to the best research infrastructures in a given field, independent of where the research infrastructures are located and by whom they are operated;
- Create the basis for a more rapid advancement of science in Europe, enabling the development of new advanced technologies and the associated growth of the European technology market as well as the creation of a new generation of researchers, ready to exploit in the best way all the essential tools needed for their research;
- Harmonise and organise the continuous flux of data collected or produced, by integrating major scientific equipment (telescopes, synchrotrons, research vessels, etc.) or set of instruments (sensors, microscopes, radars, etc.), as well as knowledge based resources (collections, archives, structured scientific information, data infrastructures, etc.);
- Structure the European Research Area, by developing synergies and complementarities between infrastructure operators, as well as a more co-ordinated approach with users and public authorities;
- Increase the potential for innovation and technology transfer of the related research infrastructures, in particular by reinforcing the partnership with industry and the use of research infrastructures by industrial researchers;
- Contribute to developing appropriate skills in Europe.

Joint transnational research projects using web archives?

The aim of this session is to debate how a joint transnational research projects using web archives could be conceived. The format of the session will be as follows:

- short introduction to the main points below (Niels, 5 min.)
- short statements from Valérie Schafer, l'Institut des sciences de la communication du CNRS & Jane Winters, University of London (max. 5 min. each)
- breakout sessions (30 min.)
- discussion, all participants (30 min.)

One of the cornerstones of RESAW is the close cooperation between national European web archives and research communities using the web archives. The rationale for this is that a research infrastructure is best developed to fit the user's needs if the point of departure is taken in actual ongoing research projects.

In the following I outline some general criteria for a joint research project within RESAW, followed by a list of possible concrete projects.

General criteria

I suggest that a joint research projects within RESAW must meet the following five criteria:

1. the research topic must be transnational, not necessarily including all European nations (or all participating nations in RESAW), but it must include at least two
2. scholars and web archives from at least two countries must be involved
3. material in web archives must be used, not necessarily as the only source, but it must play an important role
4. there are no obligations as to research disciplines, types of research questions, theories, or methods

5. a research project must, on the one hand, be a research project in its own right, but, on the other hand, have a RI component, that is: any research project should contribute to the development of the RI (e.g. validation of archives used, testing, shortcomings of the RI, analyzing infrastructural needs, development of analytical tools (or other tools))

Possible concrete projects

In the second web discussion a number of concrete research projects were suggested. In the gray box I have pasted my summary regarding joint research projects. I have a few things to add myself, they can be found after the paste.

From the summary of the second web discussion session within the RESAW project (September 2013), pp. 5-7

1) Book on web histories

A collective book about web histories in Europe (Valérie).

2) National domains on the web

A study of what a number of national domains has looked like through the years, based on a number of shared measuring points: average size of a website, update frequency of websites, top 100 most linked to websites, number of images/videos/written text, etc., etc. And studies of a more qualitative nature could also be made (Niels).

3) The web in planned transnational events

A study of the web as it is used in relation to transnational events, within each nation state, and on web domains which are not related to a nation (for instance .eu). Two events were suggested: EU Parliamentary elections, and the Olympics. Both events are of great political and cultural importance to many people, they are national as well as transnational, they are recurrent every fifth/fourth year which makes historical studies possible, and they have been archived in many national web archives, and even across web archives with the 2012 London Olympics where a transnational web archiving was initiated, but apparently this material has never been studied. These studies could use

the same measure points as in relation to the project 'National domains on the web', but any other kind of study of life on the web in relation to these events could be imagined, from almost all disciplines within the humanities and the social sciences (Niels).

Regarding the EU Parliamentary elections it was added that these elections would be an ideal joint research project since they are a repeated event, they take place in all countries at the same time, and they can be studied in a longitudinal way (2004 if there are archives-2009-2014). In addition, in relation to this event social media such as Facebook and Twitter will become very important in 2014. However, it has to be taken into consideration that the status of the election varies from country to country, and it must be established if most countries have archived or plan to archive a number of websites/social media in 2014 (Fabienne).

It was added that the longitudinal/historical perspective is very important since it highlights what we as scholars can actually do with existing web archives where we cannot go back and change the archiving since the material is gone. And if the project runs until the next elections in 2019 web archives and scholars can together plan archiving in relation to these elections. This would give a nice test and development of the nexus between web archives and scholars: a) web archiving done in the past without scholars involved, b) web archiving done with scholars semi-involved (2014), and c) web archiving done based on close collaboration (2019) (Niels).

4) The web in unplanned transnational events

In addition to a planned event such as EU elections or Olympics it was suggested to try to establish a kind of 'contingency plan' to be used when an unexpected event happens (9/11, tsunami...). Such a contingency plan could be established as a collaboration between web archiving institutions and scholars, and it should set up a number of procedures to secure as much being archived of the event as possible – content disappears or changes rapidly during events; some web archives already have some kind of plans, but a more systematic approach as well as collaborations with scholars could be developed (Niels).

5) The period before national web archives were established

The question was raised if joint research projects could begin before the national web archives were established, that is going back to 1996 when the Internet Archive was founded (Valérie).

It was noted that this would require an agreement with the Internet Archive allowing scholars to get – of to get access to – national domains in their archive. This is probably doable – more of the web archives in RESAW have experience with this (Niels).

6) Contextual studies

Within a historical approach it was suggested that contextual issues that web archives cannot inform us explicitly about are included in web studies, that is such things as traffic, speed, users and uses, technical history and web technology. This kind of contextual information could be of great help to anyone studying web archives, especially those who are not familiar with the historical context of the web, its use, etc. (Valérie).

It was added that the combination of web archived material and contextual information constitutes a very interesting new research issue (Niels).

7) A country's 'presence' in other national web archives

One could examine the 'presence' of other countries in each national web archive (Anat).

To this I would like to add the following three points:

1) National domains on the web

As can be seen above I suggested a project about 'National domains on the web'. Since I suggested this project in the web discussion I have developed it further in a paper entitled "Probing a nation's web sphere: A new approach to web history and a new kind of historical source" (submitted to the ICA conference 2014), and a reviewed and edited version will be included in the *Routledge Companion to Comparative Internet Histories* (eds. M. McLelland, G. Goggin, 2014/15). The aim of the project is, on the one hand, to identify a

nation's web sphere as a new kind of historical source, and, on the other hand, to establish a European 'web trend index' based on national studies. The question under scrutiny is: How to study a nation's web sphere and its developments over time?

There is a great deal of overlaps between this project and the UK project 'Big UK Domain Data for the Arts and Humanities (BUDDAH)' which has just received funding from the Arts & Humanities Research Council. BUDDAH is headed by Jane Winters, University of London, and participants are the Oxford Internet Institute (Eric Meyer, Ralph Schroeder), the British Library (Helen Hockx-Yu), and I am associated the project as an academic consultant.

Since my two texts mentioned above are not public yet, I cannot send them to the RESAW group, but, in brief, they are not an empirical study, but rather a discussion of some of the methodological challenges that this new type of historical study and historical source material puts on the research agenda, and an attempt to outline a relevant and useful analytical design. Both aims are debated by using Denmark and the material in the national Danish web archive Netarkivet. However, all the debated methodological challenges are generic and will therefore in the main apply to any country. The following topics are discussed:

Introduction

- A nation's web sphere
- The necessity of web sphere historiography

Constructing a nation's web sphere

- Archived web
- Spatial delimitation: The nation's web sphere is not a mirror image
- Temporal delimitation: The nation's web is not a clear-cut point in time
- Netarkivet, the Danish internet archive

Characterizing the development of a nation's web sphere: Analytical design

- Size: Bytes
- Space: Geolocation

- Structure: Networks of hyperlinks
- Vivacity: Domain names and updating
- Content: closedness, file and software types, language, and semantics
- Comparative studies

Based on these points continued analyses of the nation's web sphere's developments could be summed up in a national 'web trend index' to be used by scholars as well as by the users of the web, especially the web producers. With the above items as a starting point the objective is to establish a pan-European research design within RESAW where the same approach is used to study each nation's web sphere with a view to comparing the national developments on a European scale thus making a European 'web trend index'.

I would like to form a RESAW-based task force to work with this project as a kind of pilot project which in itself is relevant, but which could also be used as a show case as to what could be done with a RI such as RESAW. I intend to apply for funding in 2014 at the Aarhus University Research Foundation (for seminars, catering, IT-support, etc.).

2) Eurovision Song Contest

Two concrete topics for joint research projects were mentioned in the second web discussion – the Olympics, and EU Parliamentary Elections. A third pan-European event could be added: the Eurovision Song Contest. It is a truly European event, it is a national event, but it also spans all EU countries, it is a recurrent event which makes historical studies possible, it could be studied from a variety of scholarly disciplines (from musicology to cultural studies, internet studies, sociology, and media studies, just to mention a few), and to my knowledge several web archives have already made themed collections about this event.

3) Wikipedia

The suggested joint research projects focus on large entities such as a national web or an event, but focus could also be on one individual website. And Wikipedia would be worth considering: it is transnational, but it has national versions in each country, it is international, but also European.

The main challenges for integrating web archives — possibilities of collaboration?

The aim of this session is to debate the possibilities of collaboration among web archives. The format of the session will be as follows:

- short introduction to the main points below (Niels, 5 min.)
- short statements from Claude Mussou, Institut National de l'Audiovisuel & Bjarne Andersen, Netarkivet (max. 5 min. each)
- breakout sessions (30 min.)
- discussion, all participants (30 min.)

Some of the concrete forms of collaboration could include the points mentioned above regarding the main aims of RESAW, that is access, workspace, documentation, researcher generated corpora, and analytical tools. What could be debated in relation to this is, on the one hand, what can easily be done — the low-hanging fruits that can easily be picked — and, on the other hand, what are the major obstacles (regarding for instance technical, legal, and organisational issues).

European web archives already have a strong tradition for collaborating within the International Internet Preservation Consortium (IIPC). In relation to this the relation to IIPC could be debated.

Debating the possibilities of collaboration among web archives with regard to archiving may also be relevant. A possible concrete project could be to set up an archiving strategy for the archiving of the domain name .eu, and to start using this strategy. This task was mentioned in the proposal from October 2012, and some of the issues of interest here are to clarify the possible technical and legal implications (copyright, privacy, access), as well as the organisational and curatorial issues. Archiving of the .eu domain could, for instance, be performed by a web archive with open access such as the

Portuguese Web Archive; this would allow for researcher access to this material, and it could be used as an easily accessible training collection for scholars in other countries.

In line with this the possible collaborations with the Internet Archive should also be examined since this is also an open web archive.

Organising the coming one year and a half: Possible activities and plans

The aim of this session is to find a way of organising the coming one year and a half until the next Horizon 2020 call will be out.

I will give a short introduction, followed by an open discussion, and possibly groups centered around specific tasks may be formed.

I suggest that in the period 2014-2015 RESAW is organised as a network centered around some joint activities. If the group agrees I will offer to act as the facilitator for the network, that is keeping track of planned activities, being an information hub, and ensuring the progress.

I consider the main tasks to be:

1. Keeping the network alive
2. Managing the contact to the national members of EU's RI programme committee

Keeping the network alive

I suggest that the backbone of the network will be 2-3 seminars and a conference. A timeline with seminars and a conference could look like this:



In addition to this individual participants/groups can organise a variety of collaborative activities on an ad hoc basis.

And the website which we already have at resaw.eu could be used as a joint platform for information sharing and collaboration. All three types of activities needs some kind of funding which remains to be found.

Seminars and conference

With a view to having a recurrent event at fixed intervals I suggest that this first seminar in Aarhus is followed by 2-3 seminars in 2014-15, and that if possible each seminar is organised by one or more of the participants and hosted by their institution. Each seminar could focus on a theme of relevance for RESAW.

Possible funding opportunities:

- If it is possible to use Danish funding outside Denmark NetLab may consider to fund the expenses for catering for these seminars; other expenses (if any) have to be funded by the organising institution

With a view to boosting an interest in the field I would like to organise an international conference in Aarhus in the Spring 2015 with the working title "The scholarly use of web archives". An organizing committee with those who want to participate in the organisation will be formed. The format will be decided by the organising committee, but the following are my suggestions:

- the conference will be aimed at scholars, web archiving institutions, web archivists, curators, companies, and public institutions interested in web archiving
- a variety of presentation formats could be used (short/long papers, poster sessions, demonstrations of software)
- acceptance should be based on blind peer review

Possible funding opportunities:

- I will make an application to the Danish Digital Humanities Lab, the umbrella organisation of which NetLab is part; other possible funding opportunities are the Danish research Council, Aarhus University Research Foundation, Aarhus University, the Department of Aesthetics and Communication, and the Centre for Internet Studies

Collaborative activities on an ad hoc basis

One could imagine a number of possible activities, depending on the level of engagement of the participants as well as on the funding opportunities. This could for instance be:

- joint papers/roundtables/panels or pre-conferences on international conferences (AoIR, ECREA, ICA, IIPC, WebSci...)
- joint PhD courses
- joint Summer/Winter schools or workshops
- visiting each other and giving lectures
- joint publications (articles, book chapters), or jointly edited books, for instance high quality contributions from the conference mentioned above or from one/more of the seminars could be compiled to an edited volume; Routledge has a book series entitled 'the Routledge Studies in European Communication Research and Education Series', connected to ECREA
- organising of national/international events to promote the scholarly use of web archives, either in general or within specific scholarly domains; participants could get inspired and/or exchange presenters if case studies and other approaches are needed (already existing initiatives: INA's [Atelier DL web Ina](#), 2009-, and a workshop day about '[Les sources nativement numériques pour la recherche en histoire et patrimoine : pratiques et méthodes](#)' (9 December, Paris, organised by Camille Paloque-Berges))
- joint research projects or national projects including other participants from the network (HERA, ESF, ERC...)

Possible funding opportunities:

Some of the activities are part of our ordinary work, others have to be funded on an ad hoc basis.

The website [resaw.eu](#)

The website on [resaw.eu](#) was initially created as an internal working tool to support the discussion phase before the preparation of the application. Thus, we did not plan it to continue in this form after the application was handed in. However, since we are not making the application now we could decide to use [resaw.eu](#) as the main portal to the network, and the internal working tool with its forums could be supplementet with a public part of the website.

The public version could inform about activities and be the dissemination platform for the results we may generate within the period 2014-15. The following could be part of the public version:

- list of the participants in RESAW
- a brief overview of the field of the scholarly use of web archives (web archives, scholarly projects, etc.)
- advertising of relevant events
- open forum for people wanting to ask/interact with our community

However, it has to be emphasised that the quality of the public version is dependant on updated input from the participants.

We could also consider establishing an open email list.

Possible funding opportunities:

If I get support from my faculty I will offer to edit the website. NetLab's IT-developer maintains the technical side of the website.

Managing the contact to the national members of EU's RI programme committee

We were close to getting RESAW's topic in the first Horizon 2020 call. With a view to increasing the probability of getting the topic in the next call we should keep the contact to the different national members of EU's RI programme committee.

I will draft a memorandum about RESAW, and make a plan about whom to contact when about what.

Evaluation of the seminar

A brief session about what worked well and what could be changed for what is hopefully going to be the next seminar.