

Exploring special web archives collections related to COVID-19: The Case of the Icelandic web archive

Caroline Nyvang and Kristjana
Mjöll Jónsdóttir Hjörvar

WARCNET PAPERS

WARCnet
web archive studies

Exploring special web archives collections related to COVID-19: The Case of the Icelandic web archive

*An interview with Kristjana Mjöll Jónsdóttir Hjörvar (The National
and University Library of Iceland, Landsbókasafn Íslands -
Háskólabókasafn) conducted by Caroline Nyvang (The Royal
Danish Library)*

cany@kb.dk



WARCnet Papers
Aarhus, Denmark 2020

WARCnet Papers ISSN 2597-0615.

Caroline Nyvang and Kristjana Mjöll Jónsdóttir Hjörvar: Exploring special web archives collections related to COVID-19: The Case of the Icelandic web archive
© The authors, 2020

Published by the research network WARCnet, Aarhus, 2020.
Editors of WARCnet Papers: Niels Brügger, Jane Winters, Valérie Schafer, Kees Teszelszky, Peter Webster, Michael Kurzmeier.

Cover design: Julie Brøndum
ISBN: 978-87-94108-00-3

WARCnet
Department of Media and Journalism Studies
School of Communication and Culture
Aarhus University
Helsingforsgade 14
8200 Aarhus N
Denmark
warcnet.eu

The WARCnet network is funded by the Independent Research Fund Denmark | Humanities (grant no 9055-00005B).



WARCnet Papers

Niels Brügger: *Welcome to WARCnet* (May 2020)

Ian Milligan: *You shouldn't Need to be a Web Historian to Use Web Archives* (Aug 2020)

Valérie Schafer and Ben Els: *Exploring special web archive collections related to COVID-19: The case of the BnL* (Sep 2020)

Niels Brügger, Anders Klindt Myrvoll, Sabine Schostag, and Stephen Hunt: *Exploring special web archive collections related to COVID-19: The case Netarkivet* (Oct 2020)

Friedel Geeraert and Nicola Bingham: *Exploring special web archives collections related to COVID-19: The case of the UK Web Archive* (Nov 2020)

Friedel Geeraert and Barbara Signori: *Exploring special web archives collections related to COVID-19: The case of the Swiss National Library* (Nov 2020)

Valérie Schafer, Jérôme Thièvre and Boris Blanckemane: *Exploring special web archives collections related to COVID-19: The case of INA* (Aug 2020)

Niels Brügger, Valérie Schafer, Jane Winters (Eds.): *Perspectives on web archive studies: Taking stock, new ideas, next steps* (Sep 2020)

Friedel Geeraert and Márton Németh: *Exploring special web archives collections related to COVID-19: The case of the National Széchényi Library in Hungary* (Oct 2020)

Friedel Geeraert and Nicola Bingham: *Exploring special web archives collections related to COVID-19: The case of the IIPC Collaborative collection* (Nov 2020)

Caroline Nyvang and Kristjana Mjöll Jónsdóttir Hjörvar: *Exploring special web archives collections related to COVID-19: The Case of the Icelandic web archive* (Nov 2020)

All WARCnet Papers can be downloaded for free from the project website warcnet.eu.

Exploring special web archives collections related to COVID-19: The Case of the Icelandic web archive

An interview with Kristjana Mjöll Jónsdóttir Hjörvar (The National and University Library of Iceland, Landsbókasafn Íslands - Háskólabókasafn) conducted by Caroline Nyvang (The Royal Danish Library)

Abstract: This WARCnet paper is part of a series of interviews with European web archivists who have been involved in special collections related to COVID-19. The aim of the series is to provide a general overview of COVID-19 web archives. The paper is a transcript of an interview with Kristjana Mjöll Jónsdóttir Hjörvar, who is in charge of special collections with the Icelandic web archive at The National and University Library of Iceland.

Keywords: web archives, COVID-19, special collections, Iceland, Icelandic Web Archive

The interview was conducted on October 29, 2020 with Kristjana Mjöll Jónsdóttir Hjörvar. Kristjana is a librarian with the Legal Deposit Department at Landsbókasafn, the National and University Library of Iceland. She handles electronic materials that fall under the Icelandic legal deposit law, such as eBooks and archived web content. The National University Library of Iceland has archived web pages since 2004. However, the collection also includes Icelandic web material dating back to the 1990s, as The Internet Archive has donated material to the National and University Library of Iceland. Kristjana has worked with the library since 2016, and she is one of two employees working with the Icelandic Web Archive. She is solely in charge of special collections. When engaged with special collections, Kristjana uses an estimated 2-3 hours a week curating for the Icelandic web archive.

THE REASONS OF THE SPECIAL COLLECTION

Did you conduct a special COVID collection? And in case you did: why?

Kristjana: Yes, we did. We did it in order to preserve this part of the Icelandic culture for the future. In general, we collect Icelandic webpages, sites where Iceland is mentioned, and websites owned by Icelanders. It is important to collect this as previously, debates would appear mainly in the papers, but now it also takes place in blogs and on social media sites.

We have a long tradition for conducting special collections in relation to various elections in Iceland. We made our first special collection on something unrelated to elections in 2018 when Iceland went to the World Cup in soccer. I actually regret that we did not make a special collection during the European Championship. This, too, was a very important event in Iceland.

COVID-19 was the second event on which we decided to conduct a special collection. I would also like to do something in relation to the Eurovision Song Contest. Normally, Icelanders are not too keen on Twitter but during events such as the Eurovision Song Contest and sporting events many take to Social Media. It would be fruitful to be able to document this and to do collections on unexpected events such as political scandals in Iceland. It is difficult, however, since I am the only person who curates materials for the web archive.

THE SCOPE OF THE COVID-19 COLLECTION

What did you collect? Websites only? Social media (if so, which ones, specific hashtags, profiles, languages)?

Kristjana: We currently have two collections going. One which harvest sites once a week, and one that runs every day. We already harvest news media often during regular crawls, so our focus is on other websites. We mainly collect Icelandic websites but also some in English and the other Nordic languages. Not Finnish, though, as I do not understand Finnish. My main approach has been to search the web for "Iceland" and various COVID-related terms.

So, no Social Media?

Kristjana: No, we do not collect Social Media content due to technical difficulties. It really feels futile not to be able to do this, but it's just too difficult in terms of technology. We looked into web recording for some Social Media sites. But it was simply too time consuming.

Are you still adding URLs to your collection?

Kristjana: Yes, I am still actively adding URLs, though not as many as in the spring when COVID-19 initially started to affect Iceland. During the summer months, when COVID-19 was not on the rise, I did not add so many new URLs to the collection. Now, since the corona virus began popping up again especially in the news media, I have increased the number of new URLs added to the collection. I add maybe 5-10 new URLs a month. Right now, we harvest around 83 websites on a weekly basis and 5 on a daily basis.

How do you archive nationally something, which is fundamentally global?

The Internet is fundamentally a global phenomenon. There are ongoing crawls on Icelandic material outside .is, so we are used to collecting outside of the Icelandic sphere.

THE FRAME OF THIS SPECIAL COLLECTION

When did you start? When do you (plan to) stop?

Kristjana: We started in late March/early April. Most websites were added during these months. During the summer, I did not add that many new URLs. We do not have any plans of stopping before COVID-19 stops! I follow the news media and its coverage. We will do that until COVID-19 is over.

How was quality control done on the collection (if applicable)?

We do not perform a quality check, since we have no way of doing that. But I do know that at times something isn't harvested properly due to programming issues on a site or if the website or part of it is unreachable for some reason.

Did you encounter any issues, challenges, or limits related to the collecting activity?

Kristjana: I find it difficult to know what to collect. I do not have any colleagues with whom I can brainstorm.

So far, I have not encountered any technical problems. If possible, though, I would like to be able to collect Social Media since that is where the debate goes on. Regular people often do not get a say in ordinary media but in Social Media, everyone has a voice.

How easy was it to manage collecting activity from home, in case you did that?

Kristjana: I did work from home for a bit, but it was not a problem in terms of setting up the collection. I simply add the URLs to a seed list. It is quite easy to do from home. I use the same methods as I do when I am at work. It requires a minimum of technical knowhow.

ACCESSIBILITY AND SEARCHABILITY

How is it (or will it be) possible to access and search these data?

Kristjana: The harvested sites will be accessible via the website [www.vefsafn.is] in the form of a basic URL search. We do not have any limitations on who can gain access – it is open to everyone.

Are researchers already asking you about these collections, waiting to analyse them?

Kristjana: No, people probably know that we do special collections but it's unlikely that they know of all of them.

How did you communicate about this special collection?

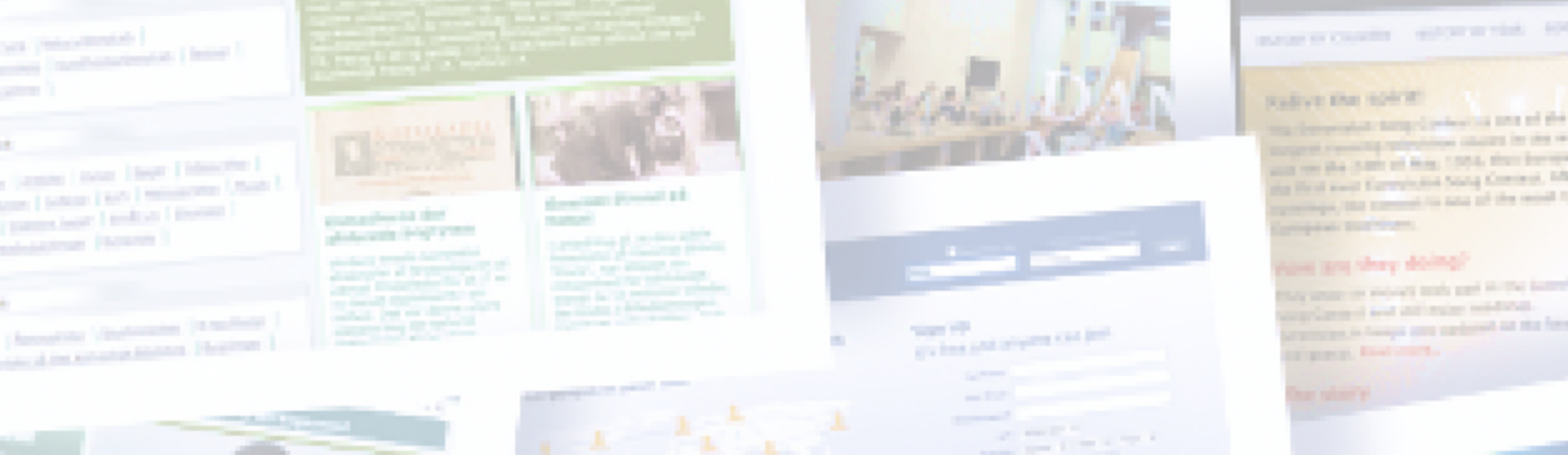
Kristjana: We have not advertised this special collection in any way. I think we just sent an email to some of our colleagues at the library asking them to contribute with URLs they found relevant.

Did you have any partnerships — with local stakeholders, Archive-It, the IIPC, etc. — during the collection process?

Kristjana: The library is a founding member of IIPC and our main web archiving technician is our contact there. He might have talked with them in relation to this special collection but I have not.

Do you have any advice for people, who would like to know more about the archival practices of your web archive? (Literature, websites etc. in English?)

Kristjana: The only information you will find in English is on our own website. It is, however, only partially translated.



WARCnet Papers is a series of papers related to the activities of the WARCnet network. WARCnet Papers publishes keynotes, interviews, round table discussions, presentations, extended minutes, reports, white papers, status reports, and similar. To ensure the relevance of the publications, WARCnet Papers strives to publish with a rapid turnover. The WARCnet Papers series is edited by Niels Brügger, Jane Winters, Valérie Schafer, Kees Teszelszky, Peter Webster and Michael Kurzmeier. In cases where a WARCnet Paper has gone through a process of single blind review, this is mentioned in the individual publication.

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 2020-22, 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).



WARCNET PAPERS



warcnet.eu

warcnet@cc.au.dk

twitter: @WARC_net

facebook: WARCnet

youtube: WARCnet Web Archive Studies

slideshare: WARCnetWebArchiveStu