

---

# Surveying the landscape of COVID-19 web collections in European GLAM institutions

Nicola Bingham<sup>\*1</sup>, Karin De Wild<sup>†2</sup>, Friedel Geeraert<sup>‡3</sup>, and Caroline Nyvang<sup>§4</sup>

<sup>1</sup>British Library – Boston Spa, West Yorkshire, United Kingdom

<sup>2</sup>Universiteit Leiden [Leiden] – P.O. Box 9500 2300 RA Leiden, Netherlands

<sup>3</sup>Royal Library of Belgium (KBR) – Belgium

<sup>4</sup>Royal Danish Library – Denmark

## Abstract

The aim of the WARCnet network (<https://cc.au.dk/en/warcnet/about>) 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. Within the context of this network, a survey was conducted to see how cultural heritage institutions are capturing the COVID-19 crisis for future generations. The aim of the survey was to map the scope and collection strategies of COVID-19 Web collections with a main focus on Europe. The survey was managed by the British Library and was conducted by means of the Snap survey platform. It circulated between June and September 2022 among mainly European GLAM institutions and 61 responses were obtained.

The purpose of this presentation is to show researchers which COVID-19 Web collections exist, what type of content the collections contain and how these collections can be accessed and used by researchers.

Respondents of the survey provided information about the different possibilities of accessing the collections and the research use of the collections. The scope of the collections was also studied, providing researchers further insight into the existing web archive collections of the COVID-19 crisis. More concretely, the presentation will provide further insight into the local, regional, national or global scopes of the different COVID-19 collections, the type of content that is included in the collections, the available metadata, the selection criteria that were used when curating the collections and the efforts that were made to create inclusive collections. The temporality of the collections will also be discussed by highlighting the start, and, if applicable, end dates of the collections and the capture frequency.

Further contextual information about the collections will also be presented. This covers the motivation behind the special COVID-19 collections, the key actors in initiating the collections, the actors who selected the content for these special collections and local, national or international partnerships that were set up.

---

\*Speaker

†Corresponding author: [k.de.wild@hum.leidenuniv.nl](mailto:k.de.wild@hum.leidenuniv.nl)

‡Corresponding author: [friedel.geeraert@kbr.be](mailto:friedel.geeraert@kbr.be)

§Corresponding author: [cany@kb.dk](mailto:cany@kb.dk)

The presentation will also highlight research use of these COVID-19 collections. One example is the work done by members of the WARCnet network. The researchers used the COVID-19 collaborative collection of the IIPC (International Internet Preservation Consortium) to study two themes: home schooling and feminism. For the former, a qualitative research approach was used and for the second a quantitative research approach. Specific services (SolrWayback (<https://netpreserve.org/event/iipc-rss-webinar-solrwayback4/>) and Linkgate (<https://linkgate.bibalex.org/>)) have also been created to offer alternative access routes into this particular collection to stimulate research use.

Revealing the scope of these Web collections will encourage humanists and data scientists to unlock the potential of these archived Web sources to further understand international developments on the Web during the COVID-19 pandemic. Mapping the COVID-19 web collections and making them more transparent will stimulate researchers to use multiple web archives, thereby generating new knowledge about this extraordinary event.

**Keywords:** GLAM sector, collecting strategies, researcher access, contextual information