2022-04-12 Statusmeeting

- Update on NAS latest tests and developments
- Status of the production sites
- Next meetings
- Any other business?

Agenda for the joint NetarchiveSuite tele-conference 2022-04-12, 13:00-14:00.

Participants
- BNF: Sara
- ONB: Andreas
- KB/DK - Copenhagen: Anders, Thomas, Stephen, Tue
- KB/DK - Aarhus: Colin
- BNE: Alicia, Miguel, José
- KB/Sweden: Peter, Par, Jonas

Update on NAS latest tests and developments

Status of the production sites

Netarkivet

BnF

Last week, a Datasprint organized by the BnF and Sciences Po Paris took place over 5 days in the BnF DataLab, within the framework of the ResPaDon project. This event brought together researchers, engineers and library professionals around 4 topics concerning web archives.

As we did last December concerning the IA harvest and to celebrate 20 years of harvests about French elections, we opened a participation form until the end of June. The public is invited to take part in the selection of French websites, concerning the 2022 presidential and legislatives elections. It is possible to make suggestions at this address: [https://www.bnf.fr/fr/collecte-du-web-electoral-pour-les-elections-presidentielle-et-legislatives-de-2022](https://www.bnf.fr/fr/collecte-du-web-electoral-pour-les-elections-presidentielle-et-legislatives-de-2022)

For the third year, the Environmental Issues and Artificial Intelligence harvests have been prepared. The crawls have been launched at the end of March. More than 714 and 639 websites have been selected respectively.

Finally, we are going to launch our annual harvest. Nearly 8900 selections will be collected with 3 budget parameters (from 50 000 to 150 000 URL per domain) and 19 jobs.

ONB

BNE

We continue testing Solrwayback, lately we have indexed a closed collection about the terrorist gang ETA. This collection has 943 seeds and approximately 2 TB of information. We want to test different aspects of the search response in a specific collection, what level of relevance and accuracy have in the full-text search and way of Solrwayback tools represents information in graphics: wordcloud, link graph, domain stat or ngram Netarchive. Our goal is to improve the experience of the researcher by offering a better way of finding information, and new ways of export and work with this. All these tests will be very important to make decisions, how to show our collections by SolrWayback.

KB-Sweden

Next meetings

- May 10th
- June 7th
- July 5th
- September 6th
- October 4th
- November 8th
- December 6th
- January 10th, 2023
Any other business?