2021-05-04 Statusmeeting

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


Participants

- BNF: Auriane, Clara, Sara
- ONB: Andreas
- KB/DK - Copenhagen: Tue, Stephen, Anders
- KB/DK - Aarhus: Colin
- BNE: José, Alicia
- KB/Sweden: Pär, Peter

Update on NAS latest tests and developments

Any feedback on NetarchiveSuite 7.0 (NetarchiveSuite 7.x Release Notes)

Status of the production sites

Netarkivet

- Broad crawl step 2 still progressing nicely - Fixed crawl time pr. job – 15 days!
- Kommuner og regioner-crawl progressing as well
- We are are doing a market analysis and business case regarding outsourcing the harvest part of web archiving
  - Have already had 2 SWOT-workshops.
  - MirrorWeb won Library of Congress tender for the next 5 years. approx 650 TB a year. A mix of Heritrix and browserbased harvesting
  - Still no conclusion but we are working on getting all questions answered
  - In terms of NetArchiveSuite -what are your main concerns? In Denmark we say we have a lot of technological debt regarding NAS, what are your views in technological debt?
- Bittmagasin-development/testing including NAS 7.0 will start soon
- Stephen back from paternity leave
- The new IT-specialist will start June 1st.
- We will also focus on strategy and procedures for
  - Getting Streaming content from providers like Netflix etc.
  - Youtube/video content
  - Podcasts

BnF

We are pleased to announce that the launch day of the ResPaDon project will take place on May, 17th 2021 and is entitled “Setting up a network about web archives, uses and opportunities”.

This project aims to develop and diversity uses, by researchers, concerning web archives which are crawled and preserved by the BnF. It is led by the University of Lille in partnership with the BnF, Sciences Po, the Condorcet Campus and the GERiiCO Laboratory.

Here is the link for online registration: https://www.collexpersee.eu/journee-de-lancement-du-projet-respadon-les-inscriptions-sont-ouvertes/

You can also find the program at this address: https://www.collexpersee.eu/wp-content/uploads/2021/04/Journee-ResPaDon_Programme-previsionnel120421.pdf

Our 7th Videos crawl, which has been launched on March 18th, has just finished. The crawl is made up of 81 Youtube channels and 92 633 videos. It should represent 5,4 TB of data.

In May, we will also launch a selective crawl for the regional and departmental elections in France, that will occur on June 20th and 27th. Two collection departments of the BnF - Law, Economics and Politics and Philosophy, History and Humanities - and 17 public libraries all over France are taking part in the selection process.

ONB

BNE
• Last month we started a “massive” crawl of periodicals in free access. We want to harvest more than 9000 websites that hosted electronic serials.
• We have been working these months on a collection about videogames, we will launch middle of May.
• We are on progress of testing SolrWayback with a our collection, but the process of indexing will be long.

Next meetings

• June 8th
• July 6th
• September 7th
• October 5th
• November 2nd
• December 14th
• January 11th, 2022

Any other business?