

NetarchiveSuite

The NetarchiveSuite software is developed and maintained by The Royal Danish Library (previously two separate institutions - The Royal Library Copenhagen and the State and University Library Aarhus) together with the National Library of France ([BnF](#)) and the National Library of Austria ([ONB](#)). The National Library of Spain ([BNE](#)) and National Library of Sweden ([KB-SE](#)) are also active members of the development project.

The NetarchiveSuite is a complete web archiving software package, developed from 2004 and onwards, and used in production to harvest the Danish web since 2005. The primary function of the NetarchiveSuite is to plan, schedule and run web harvests of parts of the Internet. It scales to a wide range of tasks, from small, thematic harvests (e.g. related to special events, or special domains) to harvesting and archiving the content of an entire national domain. The software has built-in bit preservation functionality. The systems architecture allows for the software to be distributed among several machines, possibly on more than one geographical location. The NetarchiveSuite is built around the Heritrix web crawler, which it uses to harvest the web.

To get started with NetarchiveSuite, [download it](#) and try it out with our [Quick Start installation setup](#), which only requires one standard Linux machine.

The software is released with full source under the [LGPL license](#).

- [Documentation](#) — This page provides links to all the NetarchiveSuite documentation.
- [Releases and downloads](#) — You can find the list of of NetarchiveSuite releases here.
- [Curators](#) — Information on curator related NetarchiveSuite issues.
- [Development](#) — Development related information
- [Project](#) — Project information
- [Usages](#) —
- [Subprojects](#) — List of the subprojects run in parallel to the core project:
- [Free text search](#) — You can find information on how free text search is done in NetarchiveSuite.
- [Discussion](#)

Mailing lists

[Announcements](#)
[Users](#)
[Development](#)
[Curator](#)

Development

[Issue tracker](#)
[Continuous integration](#)
[Source code](#)
[Reviews](#)

Want to join the NetarchiveSuite project?

Sign up to SBForge (top left corner) and send csr@kb.dk a request for participation.

Blog Posts

[NetarchiveSuite 5.0-RC1 ready](#) created by
Mikis Seth Sørensen Jun 09, 2015
NetarchiveSuite

[Finished moving source code to GitHub](#) created by
Mikis Seth Sørensen Jun 11, 2014
NetarchiveSuite












[NetarchiveSuite 4.4 is now ready](#) created by
Mikis Seth Sørensen May 12, 2014
NetarchiveSuite

[Initial proposal for agenda ready](#) created by

Final date set to 24-25 november created by

Updates

Recently Updated

-  [Managing Database Schema Changes](#)
 - updated Jul 20, 2018
 - [view change](#)
-  [Detailed Architecture](#)
 - updated Jul 20, 2018
 - [view change](#)
-  [2018-07-17 Statusmeeting](#)
 - updated Jul 17, 2018
 - [view change](#)
-  [2018-07-17 Statusmeeting](#)
 - updated Jul 17, 2018
 - [view change](#)
-  [2018-07-17 Statusmeeting](#)
 - updated Jul 17, 2018
 - [view change](#)
-  [2018-07-17 Statusmeeting](#)
 - updated Jul 17, 2018
 - [view change](#)
-  [2018-07-17 Statusmeeting](#)
 - updated Jul 17, 2018
 - [view change](#)
-  [Docker-based Test System](#)
 - created Jul 16, 2018
-  [Browser Based Harvesting with Umbra](#)
 - updated Jul 16, 2018
 - [view change](#)
-  [Overall Architecture for Browser-based Harvesting](#)
 - updated Jul 16, 2018
 - [view change](#)
-  [Heritrix+Umbra install guide](#)
 - updated Jul 09, 2018
 - [view change](#)