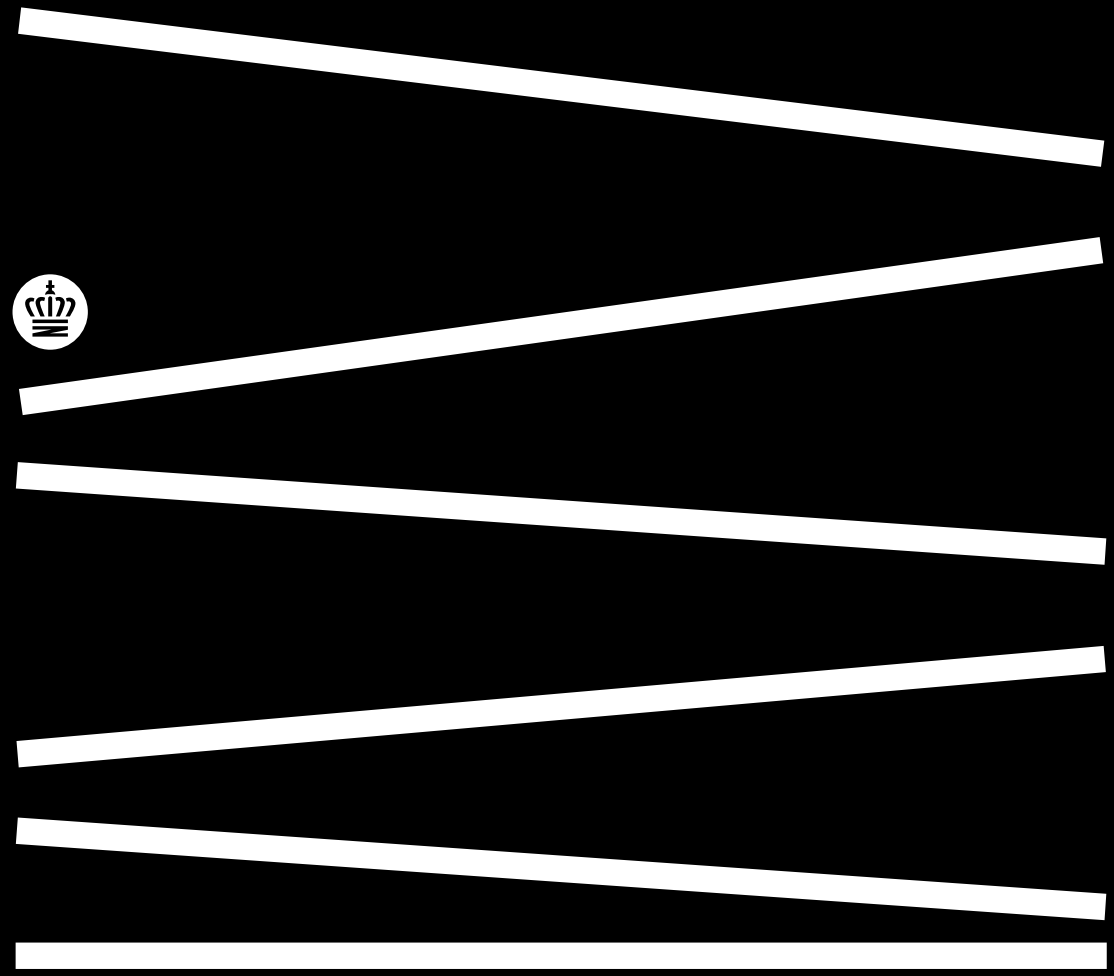
SolrWayback

Demo of the SolrWayback search interface, tools and playback engine for WARCs

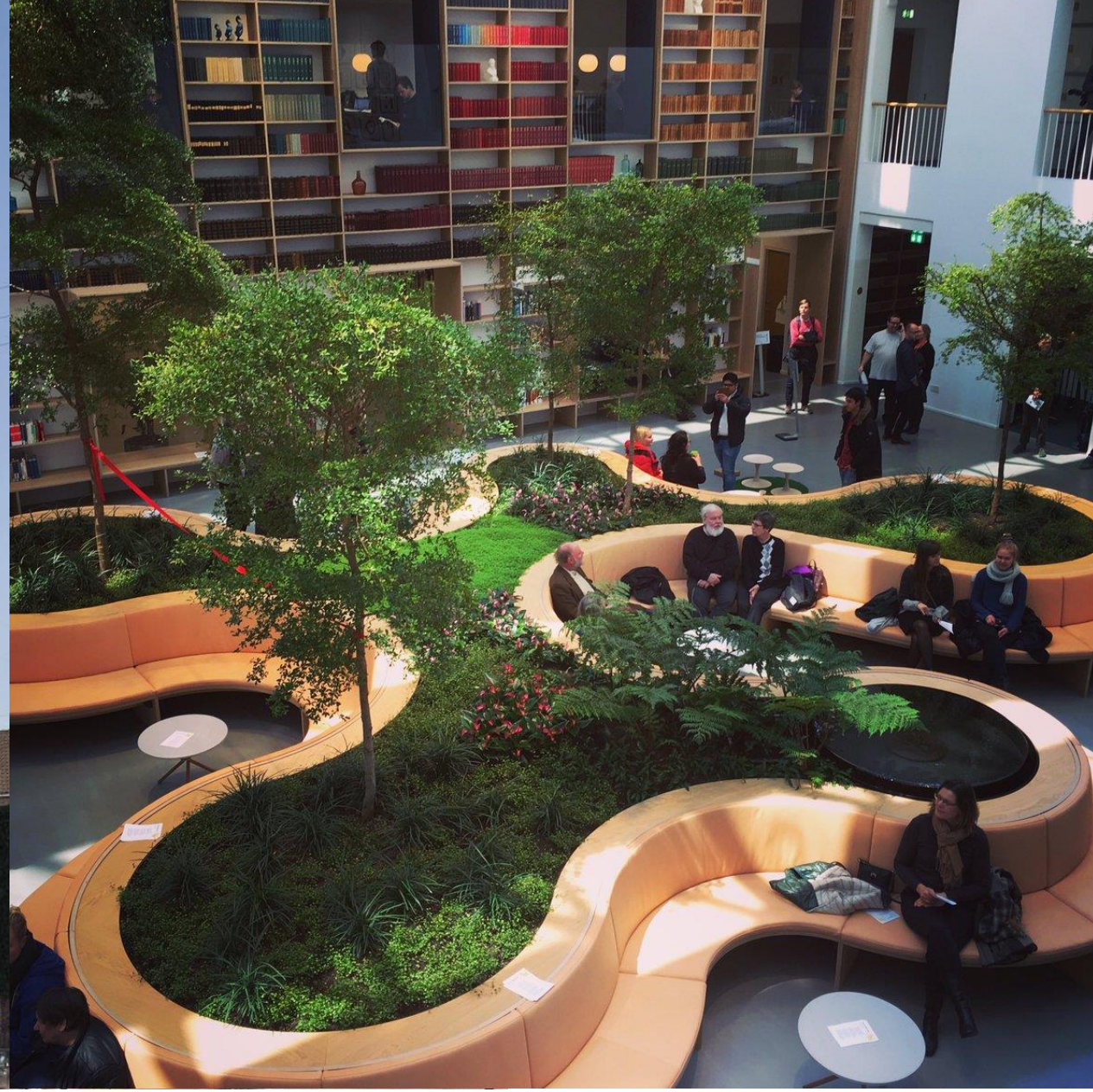
Thomas Egense
IT specialist

Anders Klindt Myrvoll
Programme Manager – the Danish web archive

IIPC 2018, Wellington



DET KGL. BIBLIOTEK
Royal Danish Library

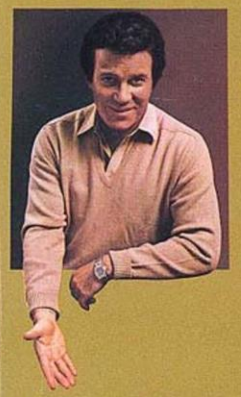


DET KGL. BIBLIOTEK
Royal Danish Library

https://da.aoab.dk/view/state-and-university-li
https://www.instagram.com/librarylovestories/ for more library pics!

DET KGL.
BIBLIOTEK
Royal Danish Library

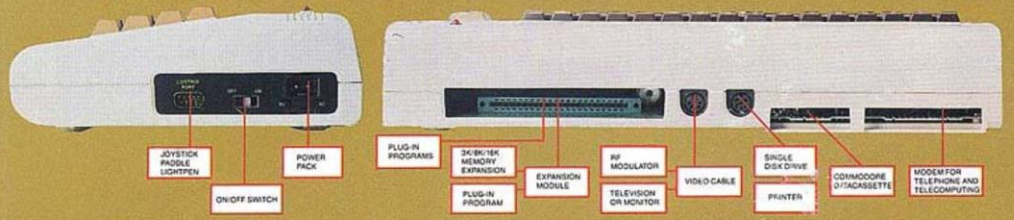
http://www.kb.dk/en/index.html



COMMODORE VIC-20®
"THE WONDER COMPUTER OF THE 1980s. UNDER $300."

—WILLIAM SHATNER

"The best computer value in the world today. The only computer you'll need for years to come."



VIC-20® VS. OTHER HOME COMPUTERS

Product Features	Commodore VIC-20	Atari® 400™	TI® 99/4A	TRS-80® Color Computer
Price*	$299.95	$458.95	$499.00	$399.00
Maximum RAM Memory	32K	16K	16K	32K
Keyboard Style	Full-Size Typewriter Style	Flat Plastic Membrane	Half-Size Typewriter Style	Calculator Style
Number of Keys	66	57	40	53
Programmable Function Keys	4	0	0	0
Graphic Symbols On Keyboard	62	0	0	0
Displayable Characters	512	256	64	256
Microprocessor	6502	6502	TI990	6809
Accessible Machine Language	YES	YES	NO	YES
Upper/Lower Case Characters	YES	YES	NO	NO
Operates with all Peripherals (Disk, Printer and Modem)	YES	NO	YES	YES
Full Screen Editor	YES	YES	NO	NO
Microsoft Basic	Standard	N/A	N/A	$ 99.00
Telephone Modem	$109.95	$399.95	$450.00	$154.95

*Manufacturer's suggested retail price Jan. 1, 1982. Includes BASIC cartridge required for programming.

Read the chart and see why COMPUTE! Magazine[1] calls the VIC-20 computer "an astounding machine for the price." Why BYTE[2] raves: "...the VIC-20 computer unit is unexcelled as a low-cost consumer computer." Why Popular Mechanics[3] says "...for the price of around $300, it's the only game in town that is more than just a game." And why ON COMPUTING INC.[4] exclaims: "What is inside is an electronic marvel... if it sounds as if I'm in love with my new possession, I am."

The wonder computer of the 1980s. The VIC-20 from Commodore, world's leading manufacturer of a full range of desktop computers. See the VIC-20 at your local Commodore dealer and selected stores.

1 April '81 issue 2 May '81 issue 3 November '81 issue 4 Fall '81 issue

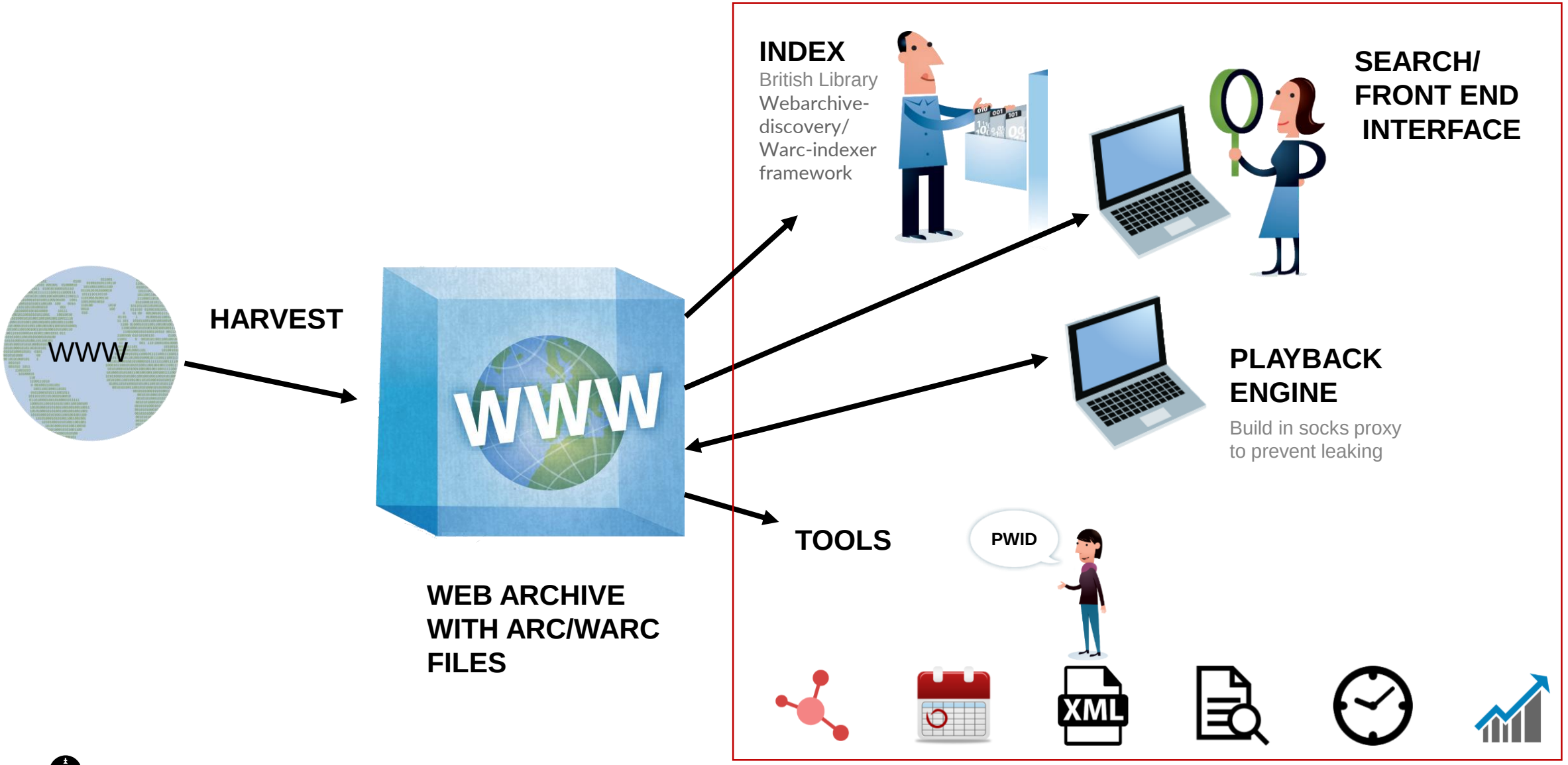
VIC-20 commodore COMPUTER

Commodore Computer Systems
681 Moore Rd., King of Prussia, PA 19406
Canadian Residents: Commodore Computer Systems
3370 Pharmacy Ave., Agincourt, Ont., Canada, M1W 2K4

Please send me more information on the VIC-20.

Name _____
Address _____
City _____ State _____ Zip _____
Phone _____

CIRCLE 120 ON READER SERVICE CARD



DET KGL. BIBLIOTEK
Royal Danish Library

Out of the box, open source web-application for researchers to explore Arc/Warc files.

PWID/XML

web archive

content coverage

pwid:mia.oszk.hu:2018-04-24T09:06:21:page:https://mnm.h

time of archiving

archived U

Best poster iPRES 2018
Eld Zierau
- online soon

```
2 urn:pwid:mia.oszk.hu:2018-04-24T09:06:21Z:part:https://mnm.hu/en/museum
```

PWID-resolver

”Prototype **Resolver** for resolving PWIDs to entries (Netarkivet & open web archives),

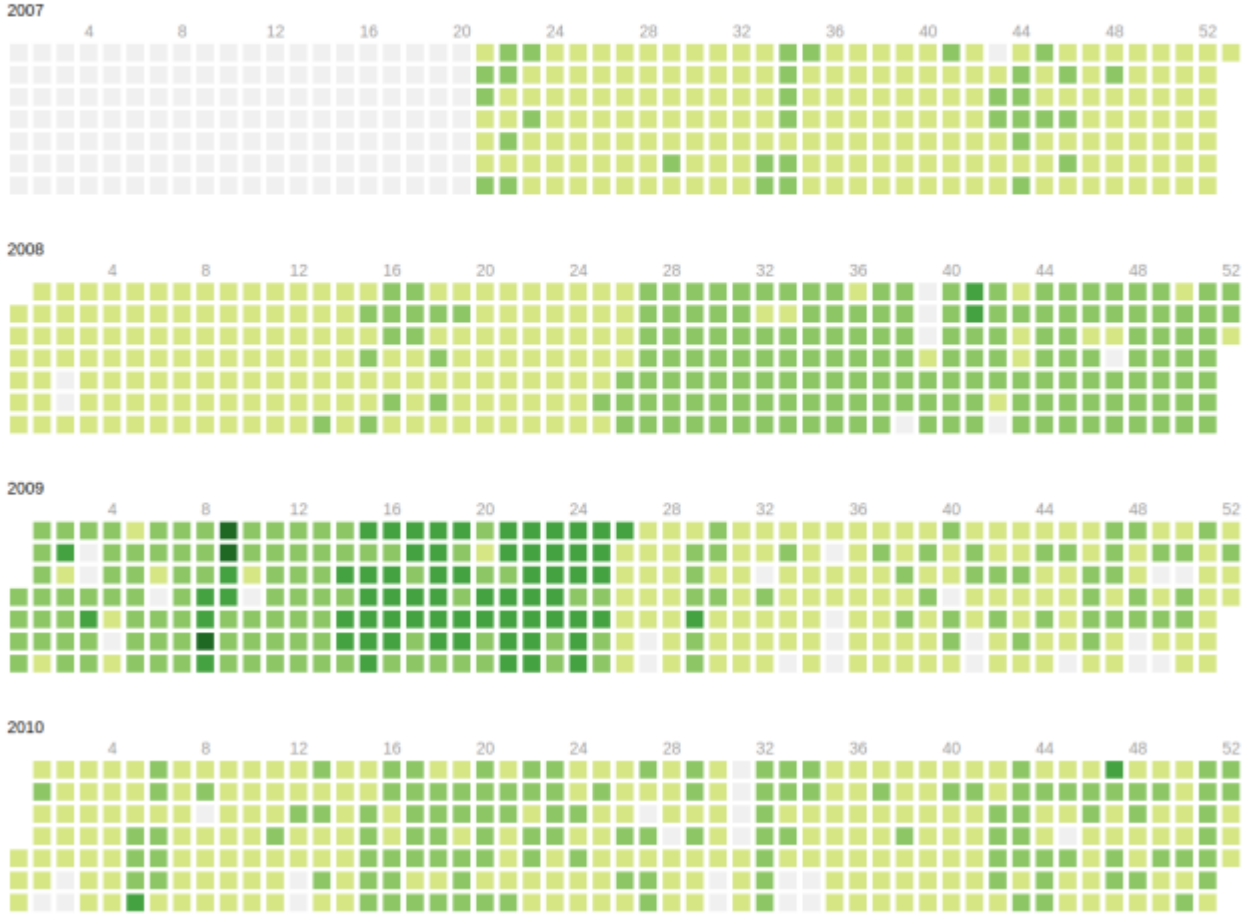
e.g. the PWID
urn:pwid:archive.org:2018-02-22T11:54:11Z:page:https://ipres2018.org/
resolves to:
http://web.archive.org/web/20180222115411/https://ipres2018.org/

Source at https://github.com/netarchivesuite/NAS-research/releases

Harvests

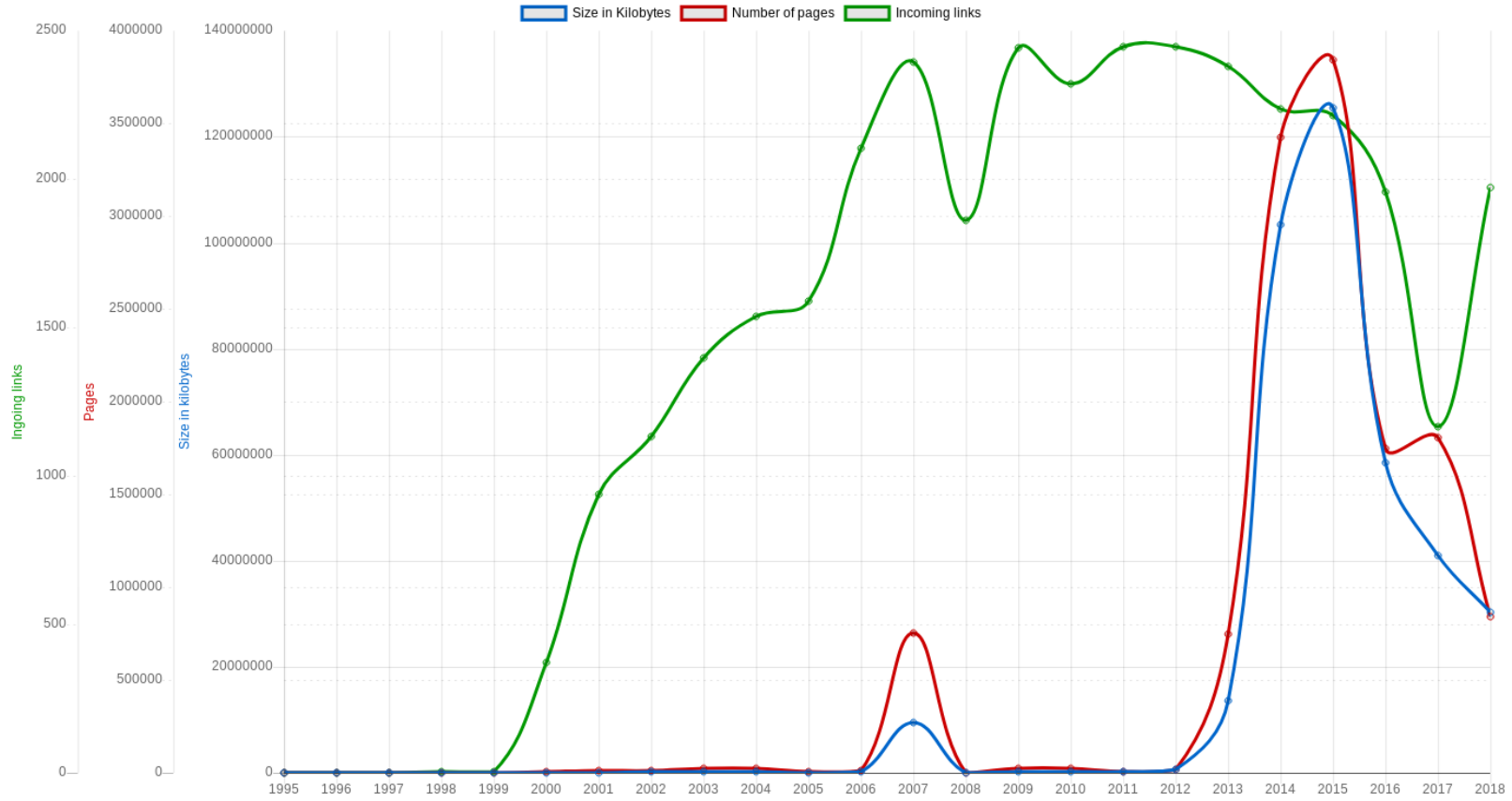
First harvest: **May 21, 2007**
Latest harvest: **September 17, 2015**
Total harvests: **19,794**

Show: Months - **Days**



Domain developement for: kb.dk

kb.dk | Check domain



	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	
Size in KB	0	0	0	71	791	38,800	101,829	146,276	256,810	263,742	100,126	192,392	9,422,925	4,185	191,176	253,533	129,691	637,159	13,614,771	1(

Domain development over time

Installing SolrWayback

• Easy to install and use on Mac, Linux and Windows. Contains Webserver, Solr and warc-indexing tool. Just drop Arc/Warcs into a folder and start exploring the corpus.

•Github-link
https://github.com/netarchivesuite/solrwayback

More info

The National Széchényi Library - Hungary http://193.6.201.202/solrwayback/
Gabor Vitez re-wrote the geo search form google maps to open streetmaps.

Athens University of Economics and Business (older version of SolrWayback)
http://archive.aueb.gr/

- **Toke Eskildsen** has helped with sparring / performance tuning and a warc export.

- **Niels Gamborg** has made 75% of the front end search interface and tools. Retired now.

Abstract

DET KGL. BIBLIOTEK
Royal Danish Library

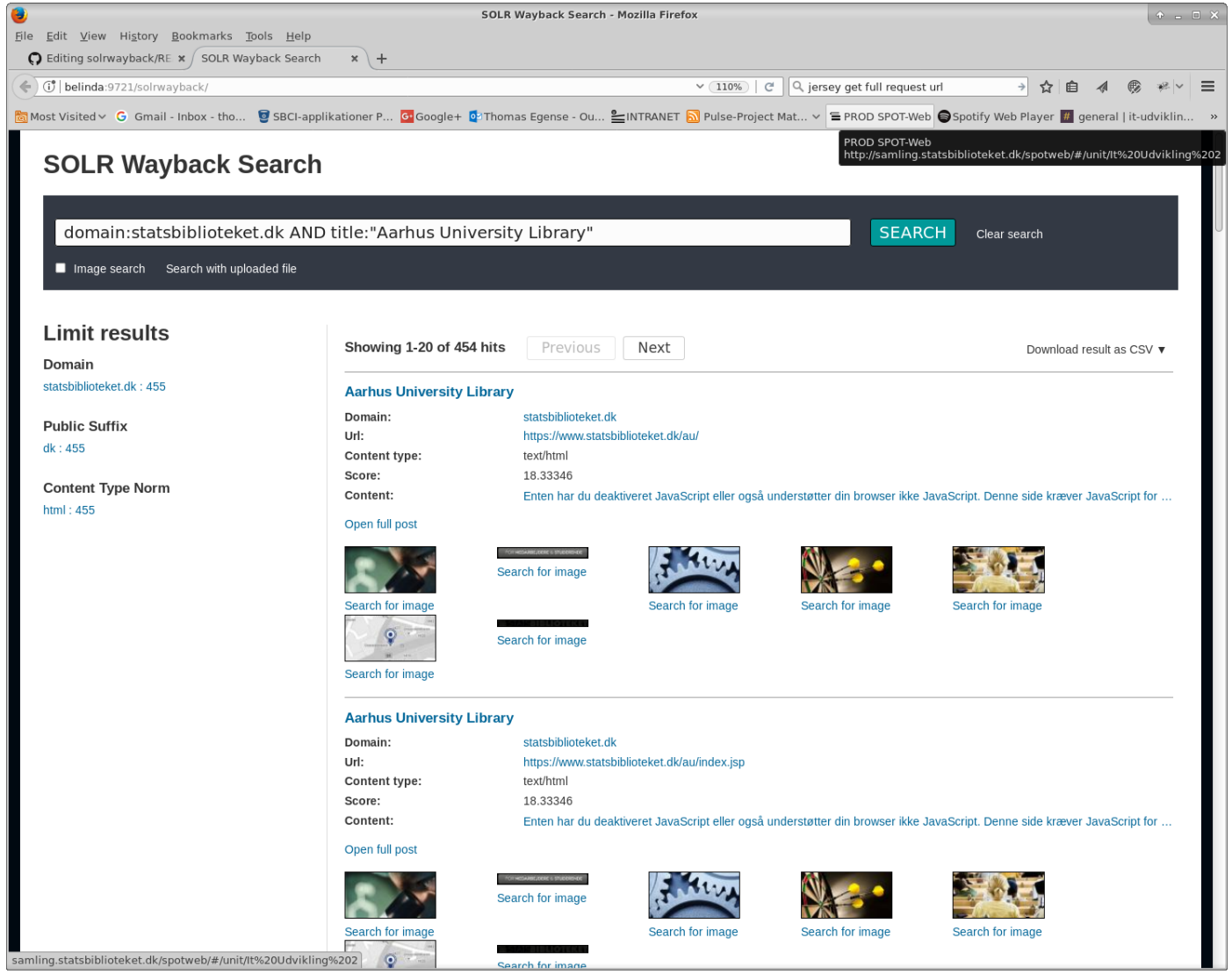
Contact

SolrWayback - Thomas Egense
teg@kb.dk @ThomasEgense

PWID - Eld Zierau
elzi@kb.dk @EldZierau

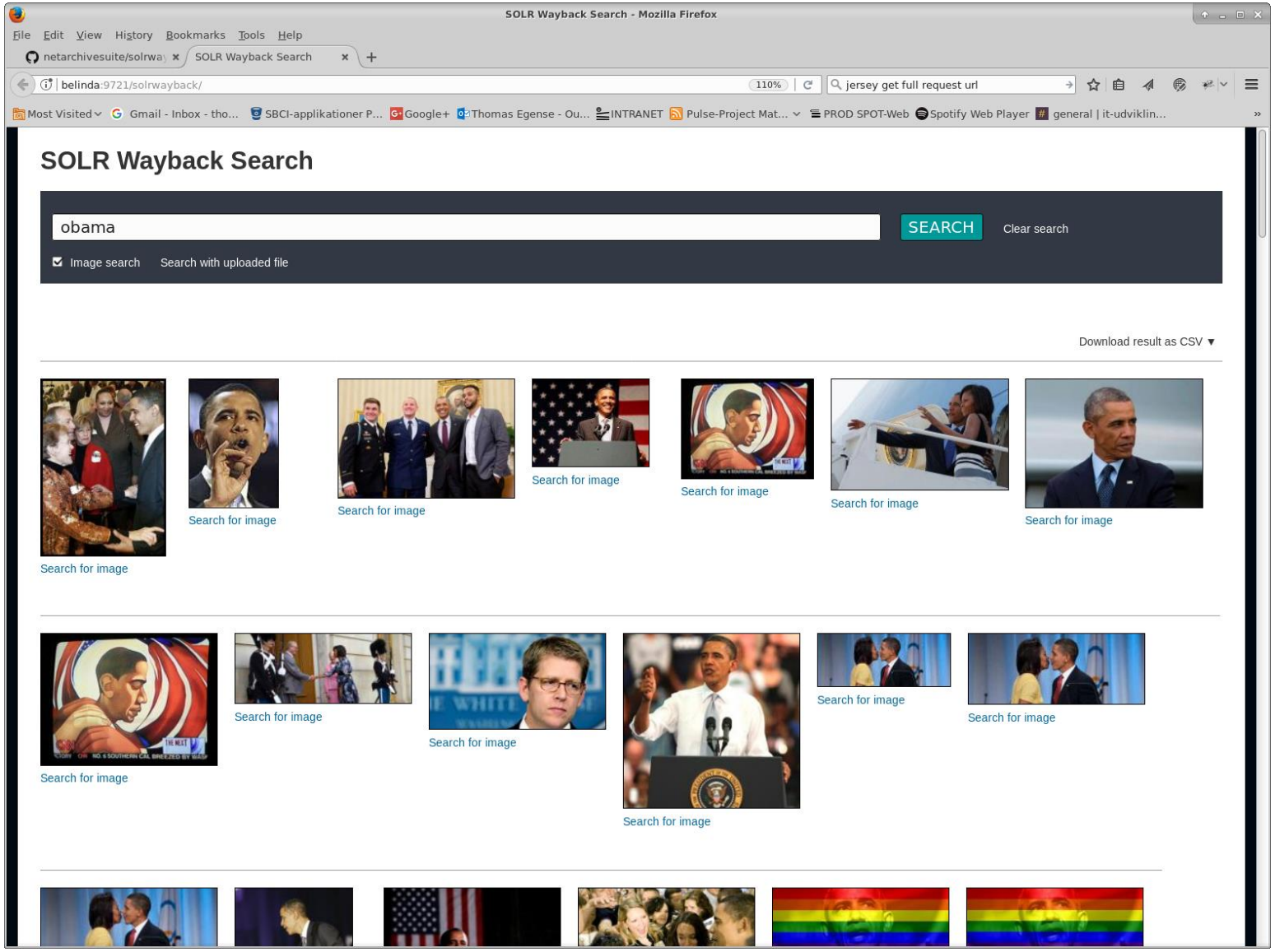
General inquires – Anders Klindt Myrvoll
ankm@kb.dk @AndersKlindt

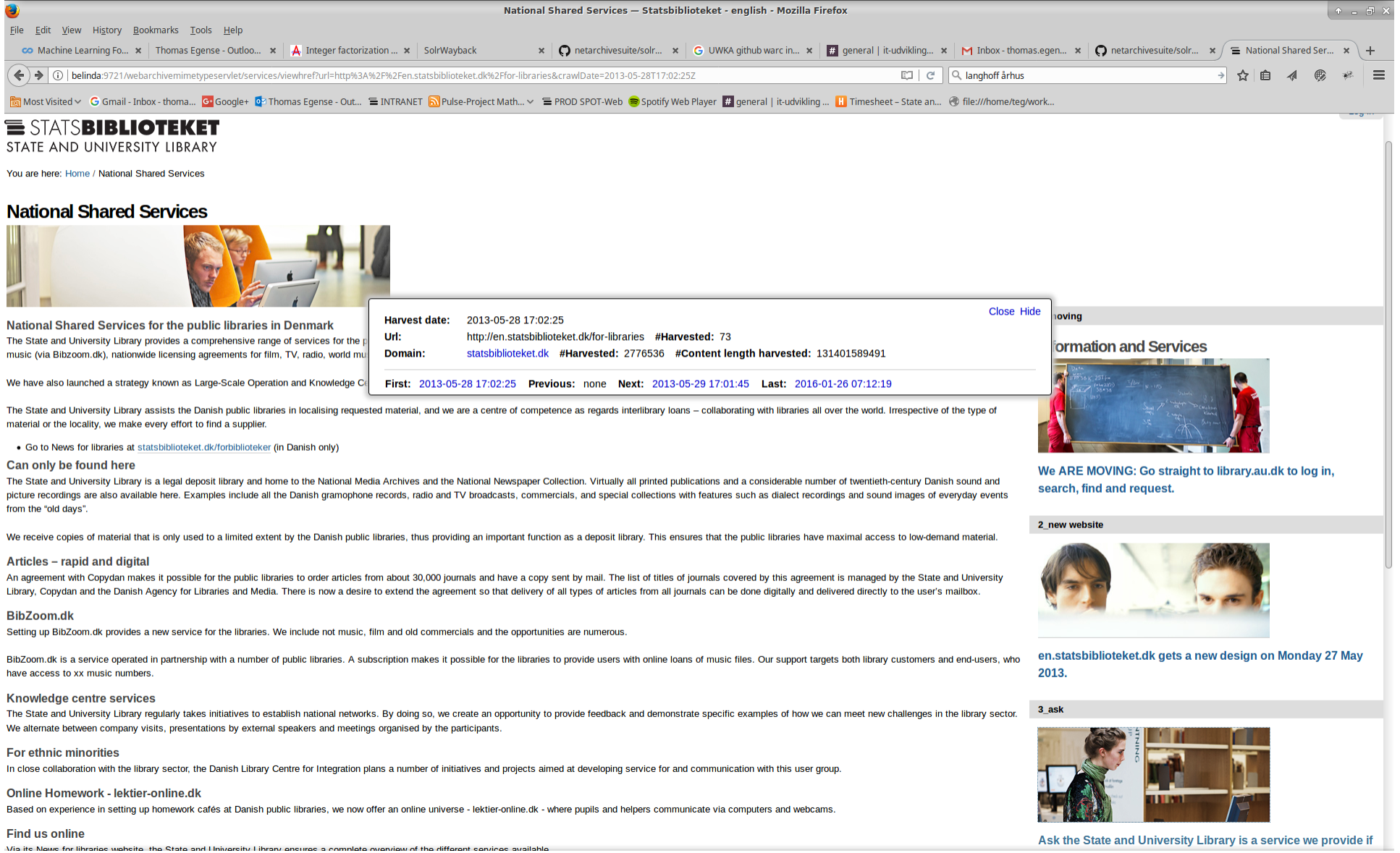
Questions and discussion



DET KGL. BIBLIOTEK
Royal Danish Library

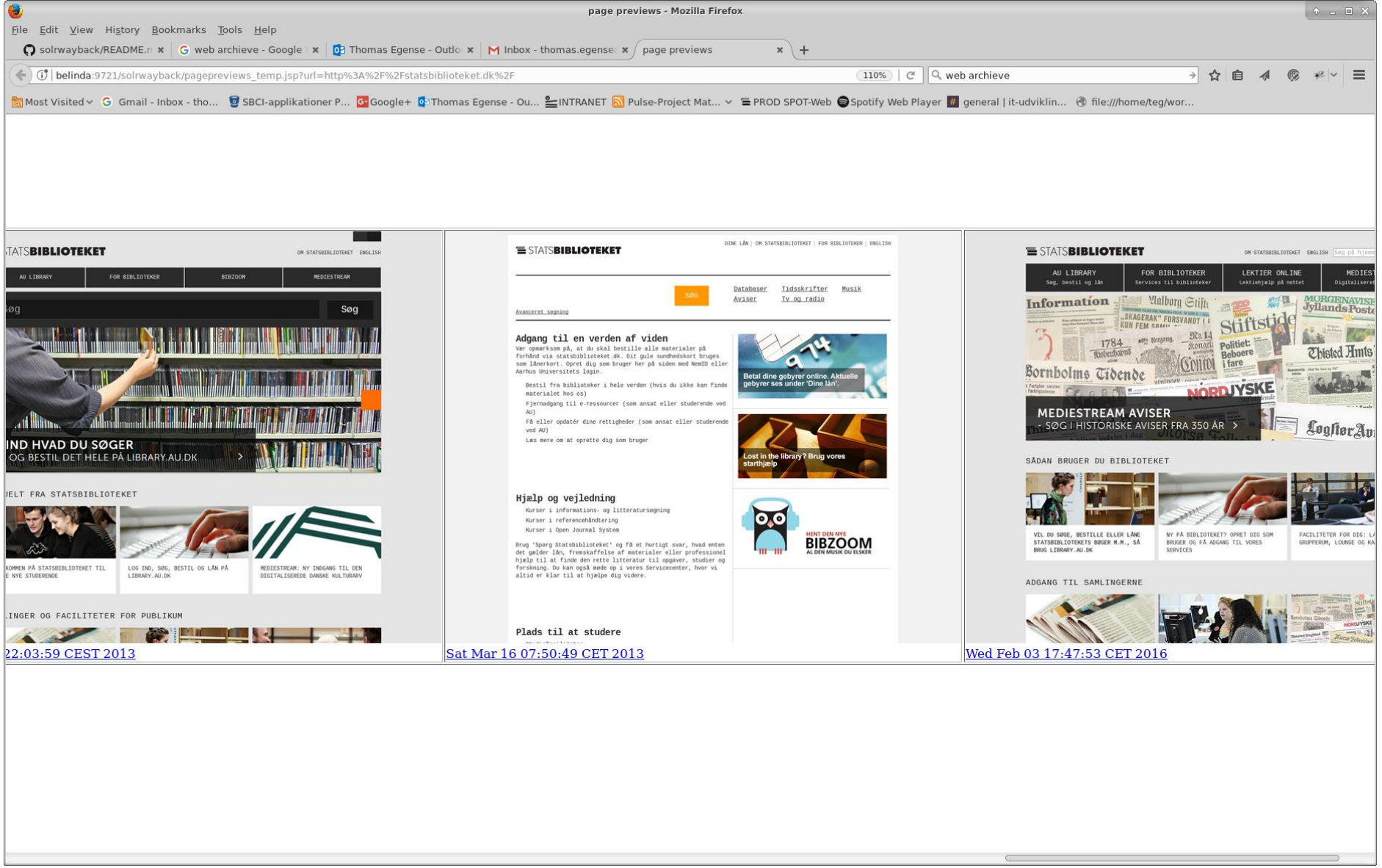
Search example showing hits. Images are shown in search-result.


SOLR Wayback Search - Mozilla Firefox
File Edit View History Bookmarks Tools Help
netarchivesuite/solrwa
SOLR Wayback Search
belinda:9721/solrwayback/
110%
jersey get full request url
Most Visited
Gmail - Inbox - tho...
SBCI-applikationer P...
Google+
Thomas Egense - Ou...
INTRANET
Pulse-Project Mat...
PROD SPOT-Web
Spotify Web Player
general | it-udviklin...
SOLR Wayback Search
obama
SEARCH
Clear search
Image search
Search with uploaded file
Download result as CSV
Search for image



DET KGL. BIBLIOTEK
Royal Danish Library

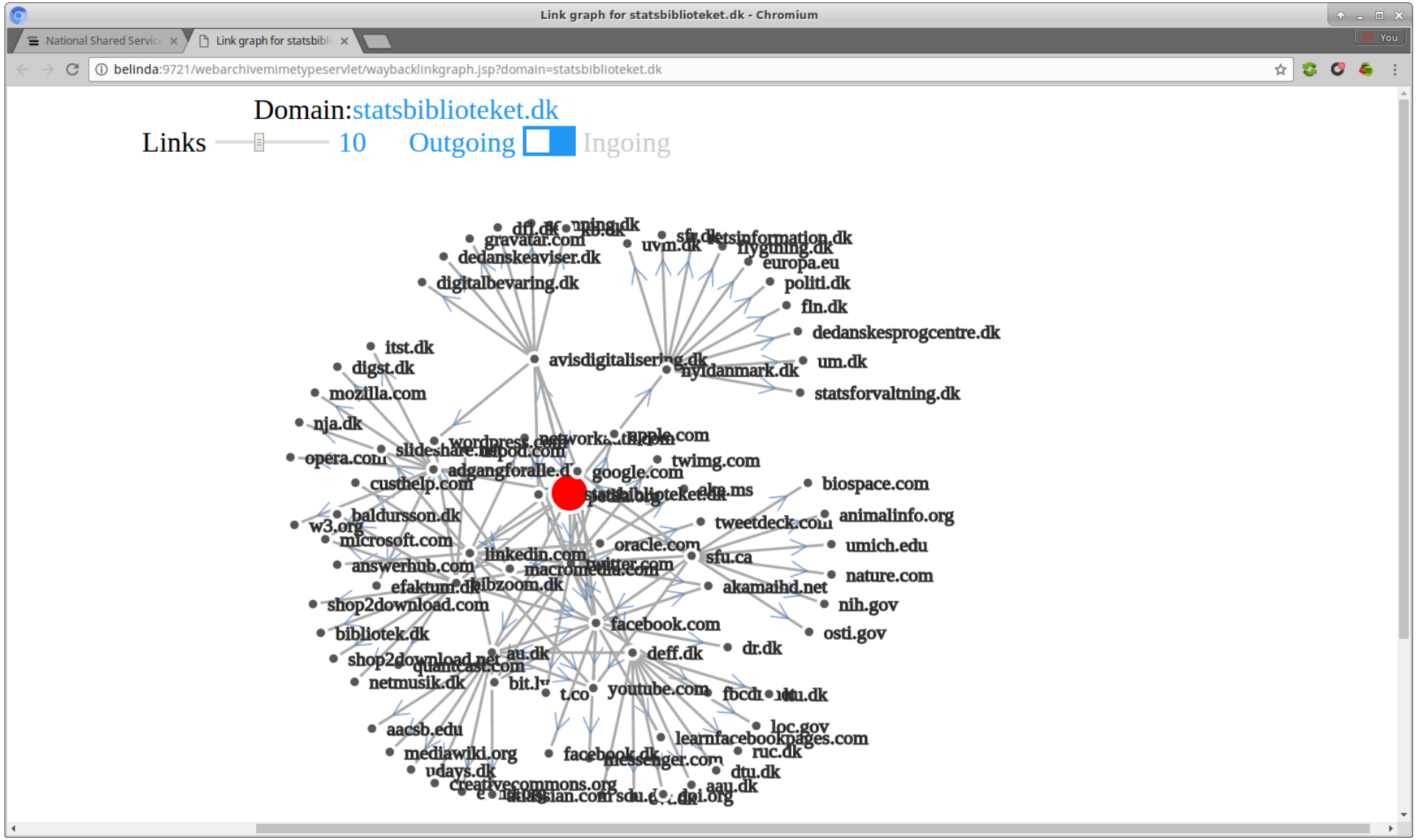
SOLRWayback showing an archived webpage with an overlay statistics and further navigation options.



22:03:59 CEST 2013 | Sat Mar 16 07:50:49 CET 2013 | Wed Feb 03 17:47:53 CET 2016

DET KGL. BIBLIOTEK
Royal Danish Library

Page previews for different harvest times of a given url. Images are generated real-time and uses the build in socks proxy to prevent leaking to the live web.



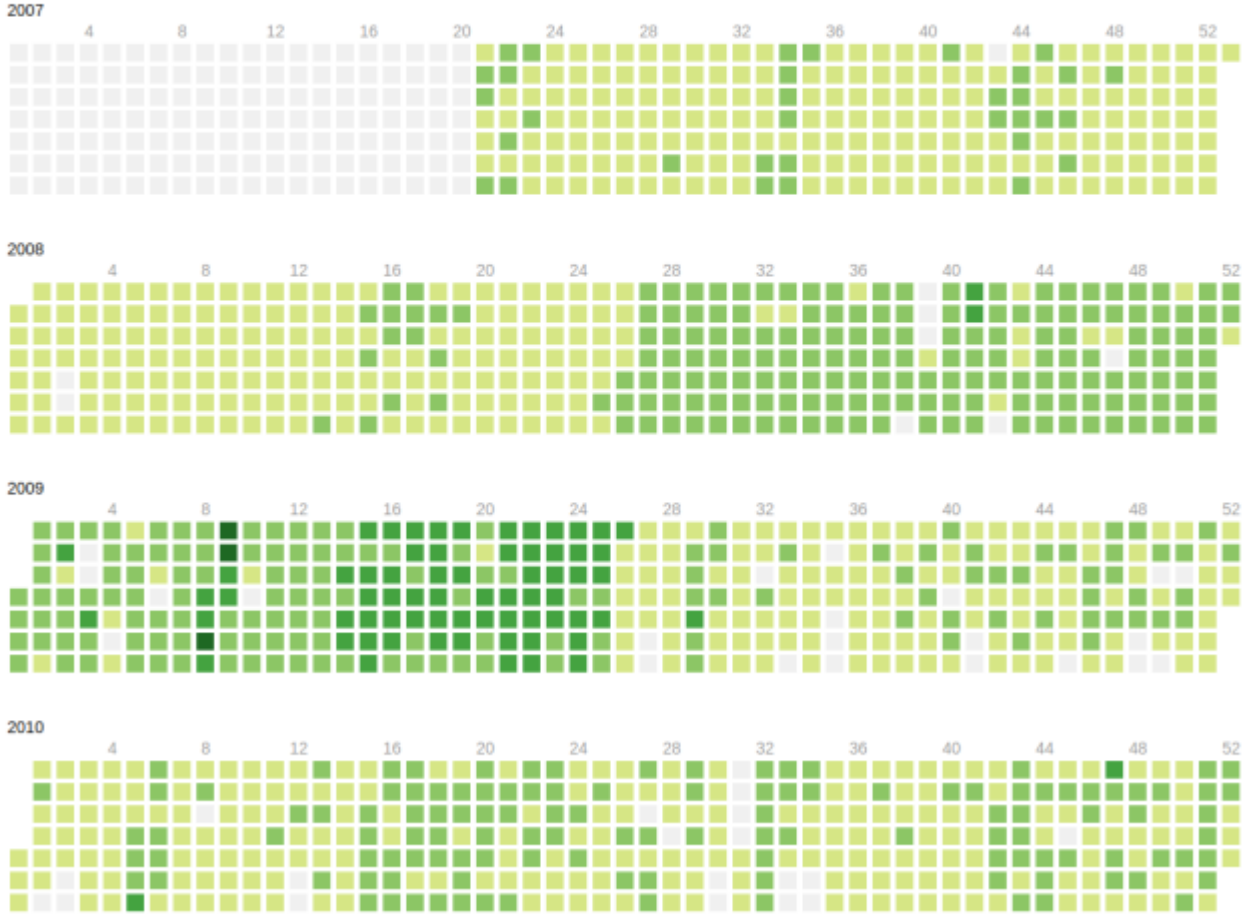
DET KGL. BIBLIOTEK
Royal Danish Library

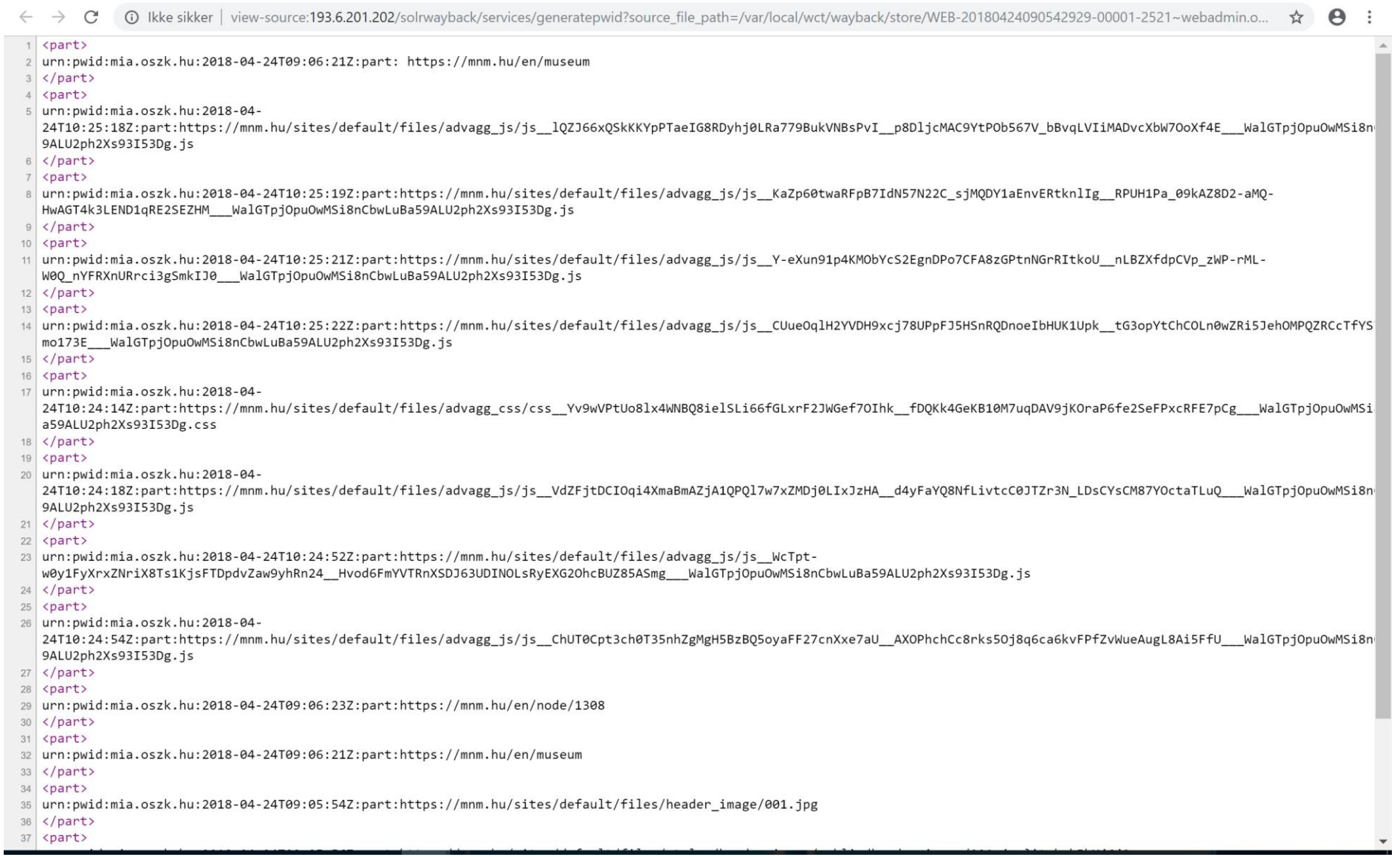
Interactive domain link graph

Harvests

First harvest: **May 21, 2007**
Latest harvest: **September 17, 2015**
Total harvests: **19,794**

Show: Months - **Days**



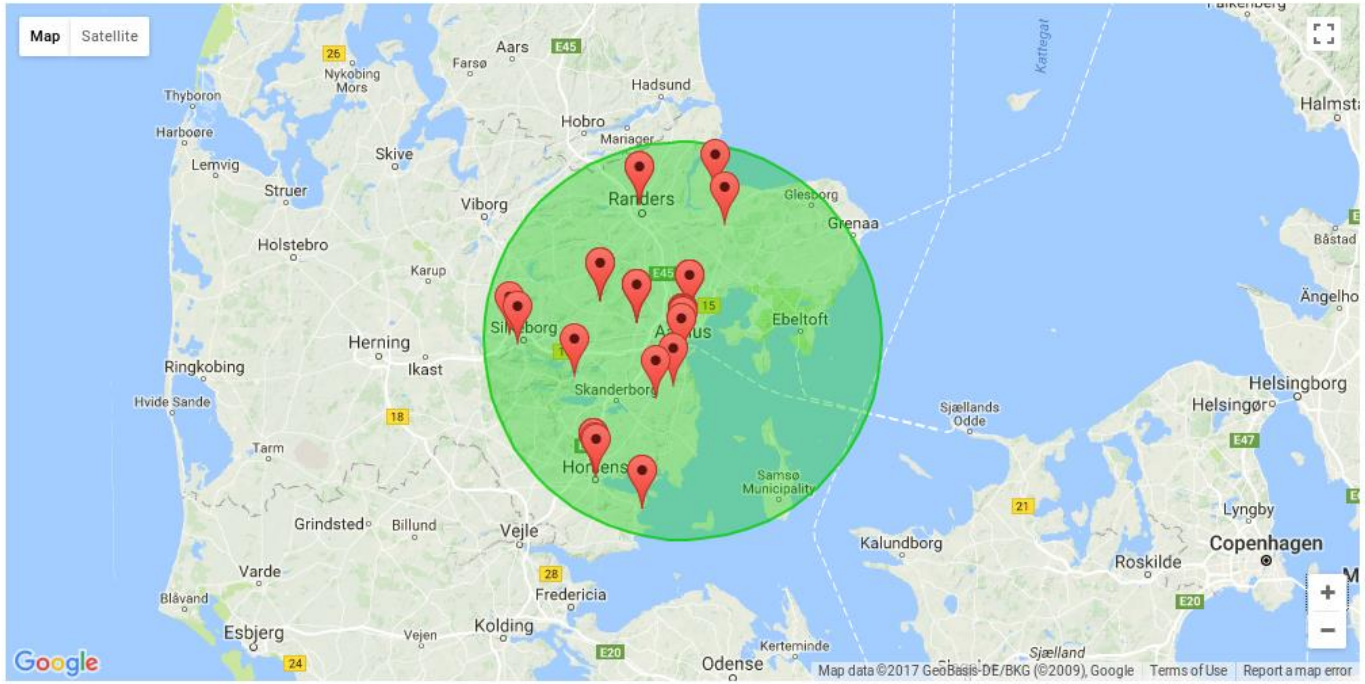

Ikke sikker | view-source:193.6.201.202/solrwayback/services/generatepwid?source_file_path=/var/local/wct/wayback/store/WEB-20180424090542929-00001-2521~webadmin.o...
<part>
urn:pwid:mia.oszk.hu:2018-04-24T09:06:21Z:part: https://mnm.hu/en/museum
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T10:25:18Z:part:https://mnm.hu/sites/default/files/advagg_js/js__lQZJ66xQSkKKYpPTaeIG8RDyhj0LRa779BukVNBsPvI__p8DljcMAC9YtPOb567V_bBvqLVIiMADvcXbW7OoXf4E___WalGTpjOpuOwMSi8n
9ALU2ph2Xs93I53Dg.js
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T10:25:19Z:part:https://mnm.hu/sites/default/files/advagg_js/js__KaZp60twaRFpB7IdN57N22C_sjMQDY1aEnvERtknlIg__RPUH1Pa_09kAZ8D2-aMQ-HwAGT4k3LEND1qRE2SEZHM___WalGTpjOpuOwMSi8nCbwLuBa59ALU2ph2Xs93I53Dg.js
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T10:25:21Z:part:https://mnm.hu/sites/default/files/advagg_js/js__Y-eXun91p4KMObYcS2EgnDPo7CFA8zGPtnNGrRItkoU__nLBZXfdpCVp_zWP-rML-W0Q_nYFRXnURrci3gSmkIJ0___WalGTpjOpuOwMSi8nCbwLuBa59ALU2ph2Xs93I53Dg.js
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T10:25:22Z:part:https://mnm.hu/sites/default/files/advagg_js/js__CUueOqlH2YVDH9xcj78UPpFJ5HSnRQDnoeIbHUK1Upk__tG3opYtChCOLn0wZRi5JehOMPQZRCcTfYSmo173E___WalGTpjOpuOwMSi8nCbwLuBa59ALU2ph2Xs93I53Dg.js
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T10:24:14Z:part:https://mnm.hu/sites/default/files/advagg_css/css__Yv9wVPtUo8lx4WNBQ8ielSLi66fGLxrF2JWGef7OIhk__fDQKk4GeKB10M7uqDAV9jKOraP6fe2SeFPxcRFE7pCg___WalGTpjOpuOwMSia59ALU2ph2Xs93I53Dg.css
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T10:24:18Z:part:https://mnm.hu/sites/default/files/advagg_js/js__VdZFjtDCIOqi4XmaBmAZjA1QPQl7w7xZMDj0LIxJzHA__d4yFaYQ8NfLivtcC0JTZr3N_LDsCYsCM87YOctaTLuQ___WalGTpjOpuOwMSi8n
9ALU2ph2Xs93I53Dg.js
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T10:24:52Z:part:https://mnm.hu/sites/default/files/advagg_js/js__WcTpt-w0y1FyXrxZNriX8Ts1KjsFTDpdvZaw9yhRn24__Hvod6FmYVTRnXSDJ63UDINOLsRyEXG2OhcBUZ85ASmg___WalGTpjOpuOwMSi8nCbwLuBa59ALU2ph2Xs93I53Dg.js
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T10:24:54Z:part:https://mnm.hu/sites/default/files/advagg_js/js__ChUT0Cpt3ch0T35nhZgMgH5BzBQ5oyaFF27cnXxe7aU__AXOPhchCc8rks5Oj8q6ca6kvFPfZvWueAugL8Ai5FfU___WalGTpjOpuOwMSi8n
9ALU2ph2Xs93I53Dg.js
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T09:06:23Z:part:https://mnm.hu/en/node/1308
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T09:06:21Z:part:https://mnm.hu/en/museum
</part>
<part>
urn:pwid:mia.oszk.hu:2018-04-24T09:05:54Z:part:https://mnm.hu/sites/default/files/header_image/001.jpg
</part>
<part>

SOLR Wayback Search

:

SEARCH Clear search

Image search Geo search Search with uploaded file



Adjust radius

Radius in km: 50

Center for this search

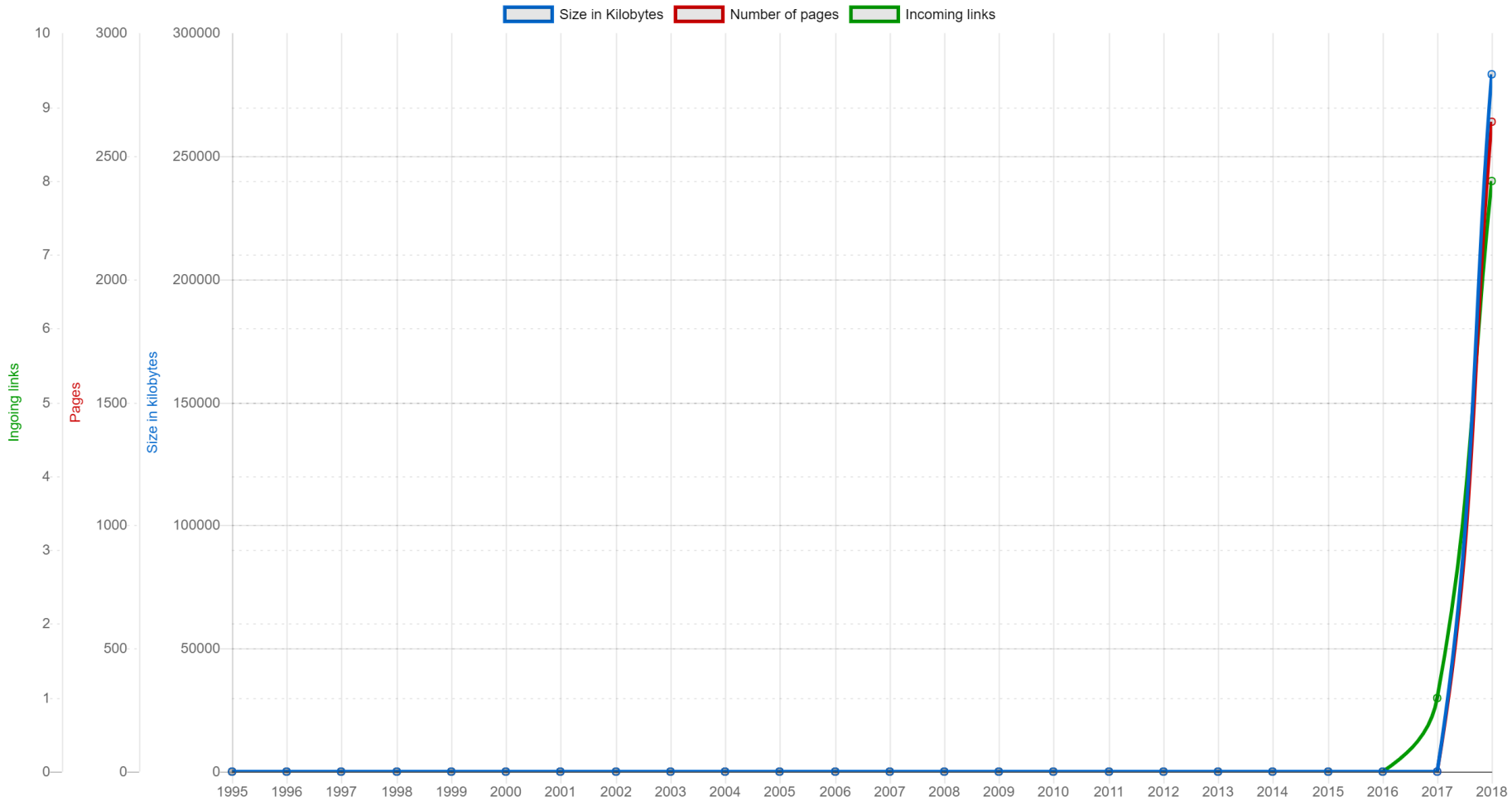
Latitude: 56.17
Longitude: 10.2

Number of hits: 50

Download result as CSV ▼

DET KGL. BIBLIOTEK
Royal Danish Library

Search by gps location for images having exif location information

Domain developement for: mnm.hu

mnm.hu Check domain