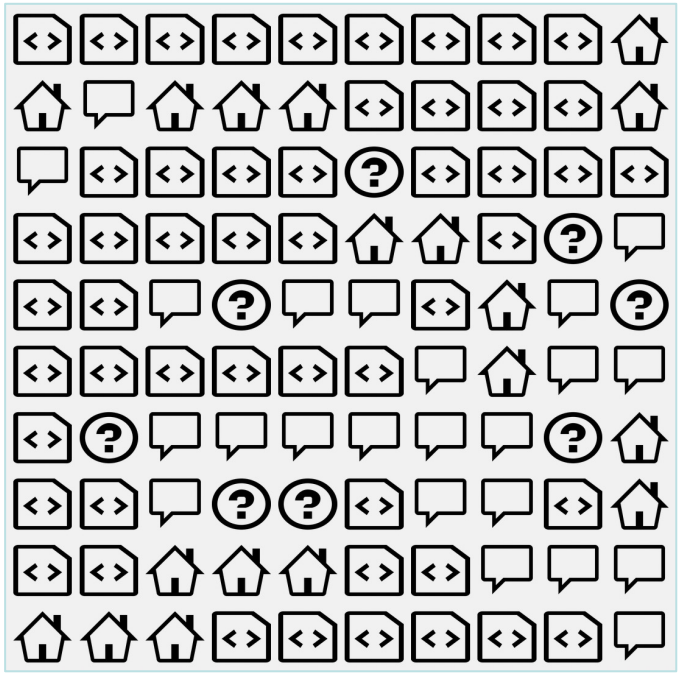
Former Farmer Chases Traces



Herbert Van de Sompel
@hvdsomp@w3c.social - https://hvdsomp.info

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

In 2009, I was invited to present a keynote at the Bielefeld Conference in, well, Bielefeld Germany. It was a special occasion because, as a child, I had lived in the German town of Arnsberg, not far from Bielefeld. And the keynote gave me the opportunity to go revisit the stomping grounds of my childhood. I felt like sharing that excitement with the audience and did so by putting up an iconic image of those days: me as a boy scout, intent on doing something meaningful with my life!

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

All materials from that 2009 Bielefeld conference are archived by the Internet Archive: program, presentation slides, audio recordings of presentations.

Exhibit 1

9th International Bielefeld Conference: Upgrading the eLibrary: Enhanced Information Services Driven by Technology and Economics, Bielefeld, Germany, 3-5 February 2009
http://web.archive.org/web/20230330152443/http://conference.ub.uni-bielefeld.de/2009/programme/

Ghent, of course, is even more special to me. I was born here, I lived here for about 40 years of my life, I got my higher education here, and I started my career here at the Ghent University Library. So, I felt that it would be appropriate to also show an iconic image that is representative of those days. Which turned out to be significantly harder because of the wide range of experiences, events, emotions in such a long period of time.

Eventually, I settled on this image: me as an experimental musician, intent on doing something creative with my life! It's a scan of a poster that announced a concert in 1987 at Ghent's cultural temple De Vooruit (meanwhile renamed to 404, what are the chances) that is available in the image database of the Ghent University Library and preserved for the long term by Meemoo, the Flemish Institute for Archives.

Our instrument was a scaffold construction prepared with contact microphones, piano strings, guitar pickups, water tubes. In the concert advertised by the poster we performed with a drumband that had won the Belgian championship the year before.

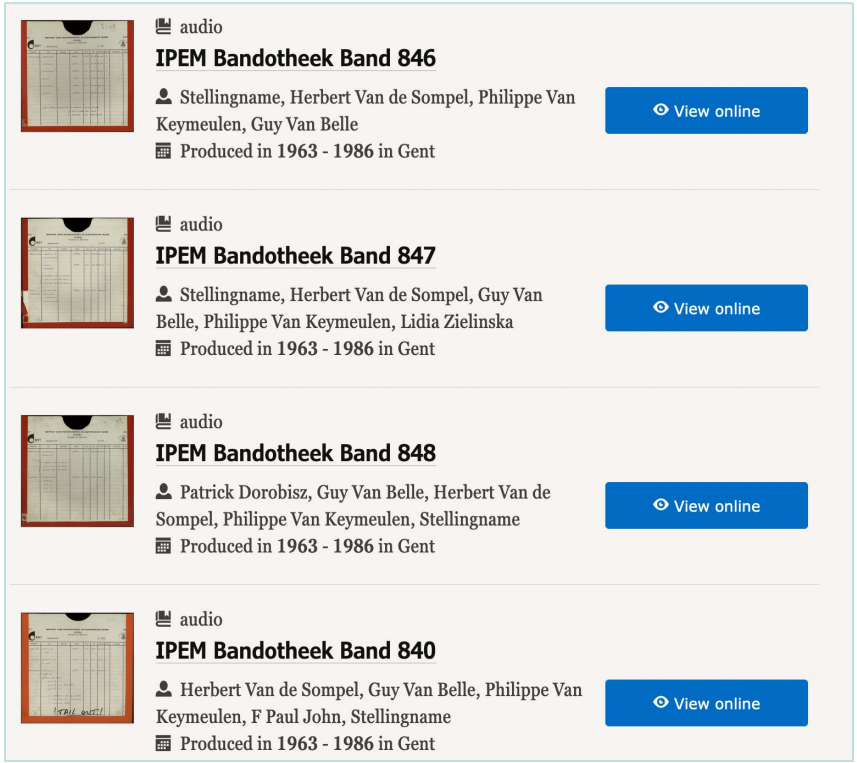
Exhibit 2

Stellingname Vooruit 4 dec.1987 : een Vooruit muziekproduktie - The Archer's paradox / foto's Michiel Hendryckx (L) & Peter Wijnendaele (R).
https://lib.ugent.be/catalog/rug01:001648386

Just in case you might be interested in that kind of music: The tapes that we recorded with Stellingname at IPEM, the Institute for Psychoacoustics and Electronic Music of Ghent University, were digitized and are also available in the Ghent Library database and are preserved at Meemoo.

Listen at your own risk.

Exhibit 3



audio

IPEM Bandotheek Band 846

Stellingname, Herbert Van de Sompel, Philippe Van Keymeulen, Guy Van Belle

Produced in **1963 - 1986** in Gent

View online

audio

IPEM Bandotheek Band 847

Stellingname, Herbert Van de Sompel, Guy Van Belle, Philippe Van Keymeulen, Lidia Zielinska

Produced in **1963 - 1986** in Gent

View online

audio

IPEM Bandotheek Band 848

Patrick Dorobisz, Guy Van Belle, Herbert Van de Sompel, Philippe Van Keymeulen, Stellingname

Produced in **1963 - 1986** in Gent

View online

audio

IPEM Bandotheek Band 840

Herbert Van de Sompel, Guy Van Belle, Philippe Van Keymeulen, F Paul John, Stellingname

Produced in **1963 - 1986** in Gent

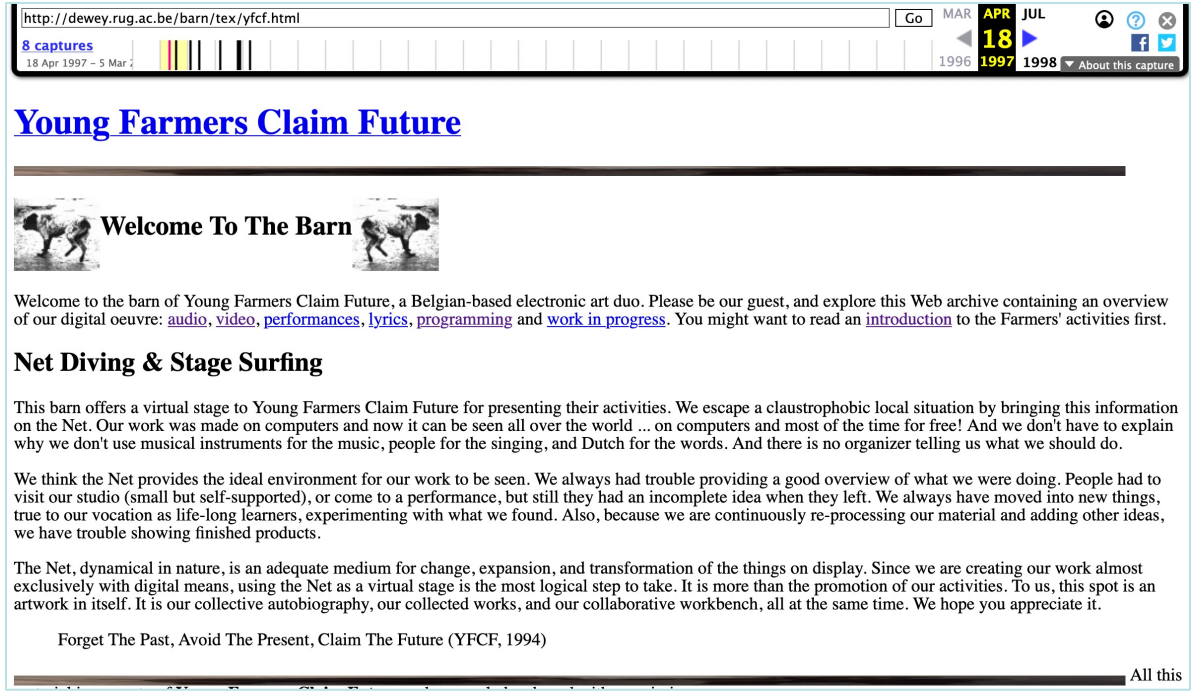
View online

https://lib.ugent.be/en/catalog?q=stellingname+ipem

After Stellingname, I got involved in a computer music band named Young Farmers Claim Future: 2 guys and 4 networked computers. We recorded two CDs each with an extended media track, played concerts in Belgium, The Netherlands, Germany, Spain. We even had a website that is archived at the Internet Archive, be it only partly.

Our slogan, early on, was ”Forget the Past, Avoid the Present, Claim the Future”. A lot has changed in my perspective since then; I wonder whether age has anything to do with that.

Exhibit 4



Forget The Past, Avoid The Present, Claim The Future (YFCF, 1994)

Young Farmers Claim Future (1996) Welcome to the Barn
http://web.archive.org/web/19970418042254/http://dewey.rug.ac.be/barn/tex/yfcf.html

In 1996, we made a contribution to a sound installation festival in Berlin, Sonambiente: 2 computers playing rock-paper-scissors with each other using voice commands output via their built-in speaker and recognized via their built-in microphones.

We have no remnants of that installation; it's all gone. Or so I thought until a Google search led me to a miniDV recording held by the Österreichische Mediathek in Vienna, where I currently happen to live.

Exhibit 5

Sonambiente 1996 MiniDV (13)
https://www.mediathek.at/katalogsuche/suche/detail/?pool=BIBL&uid=781471&cHash=21e348c89f1252cf5c9583284b09be9a

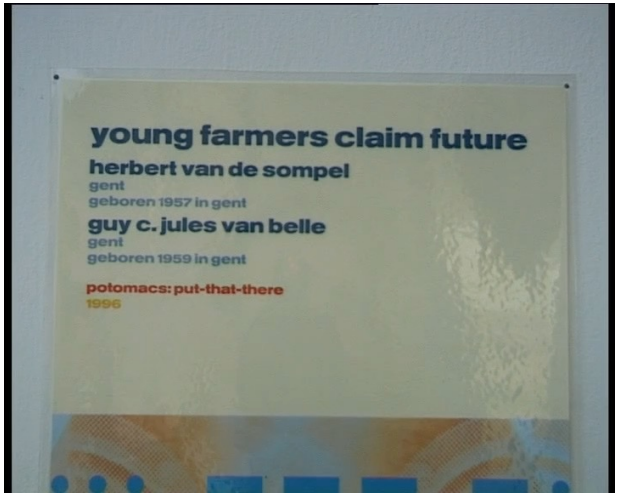
I got in touch with the people at the Österreichische Mediathek. They told me that I could get an MP4 file copy of the miniDV but that, unfortunately, the recording was defective from minute 40 onwards.

But, it turned out that YFCF's segment already started at minute 5. These are two still images taken from it.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

Exhibit 6



Sonambiente 1996 MiniDV (13) VX-10726 Österreichische Mediathek

While making music with Stellingname and YFCF, my real job was library automation, here in Ghent, at the University Library. I literally started automation by myself and from scratch. My first task was administrative automation of interlibrary loan on a PDP 11/34. To cut a long story short, over time I was able to assemble an amazing team and we became pioneers in the use of CD-ROMs to bring scientific information to our users, subscribing to databases, publishing some ourselves, and making them accessible across the campus network even when it was still running X25.

Then the Web happened. And, as was the case with most libraries, our first use of it (1996) was to advertise the availability of those databases and explaining how they could be accessed. Not via the web but using a range of access mechanisms that involved downloading software or consulting at networked PCs in Faculty Libraries.

Exhibit 7

1. Beschikbare databanken?

*Klik op de **naam van de databank** voor informatie omtrent de inhoud. Klik op de **toegangsmethode** om te weten waar en hoe de databank geraadpleegd kan worden. Klik op het **type** om te weten over welk soort databank het gaat.*

Databank	Toegangs Methode	Databank Type	Medium
ABI/Inform	Management	secundair	CD-ROM
Agecon	ERL	secundair	CD-ROM
Agricola	ERL	secundair	CD-ROM
Agris	ERL	secundair	CD-ROM
Antilope op CD-ROM	Bibinfo	catalogus	CD-ROM
Antilope via WWW	WWW	catalogus	on-line
Art Index	ERL	secundair	CD-ROM
Arts & Humanities Citation Index	Bibinfo	secundair	CD-ROM
Beast CD	ERL	secundair	CD-ROM
Biological Abstracts	ERL	secundair	CD-ROM
Boston Spa Books	Bibinfo	catalogus	CD-ROM
Boston Spa Conferences	Bibinfo	catalogus	CD-ROM
Boston Spa Serials	Bibinfo	catalogus	CD-ROM
CAB	ERL	secundair	CD-ROM
Catalogus Universiteit Gent	LIBISnet	catalogus	on-line
CCB op CD-ROM	Bibinfo	catalogus	CD-ROM
Corpus Christianorum	Bibinfo	primair	CD-ROM
Compendex Plus	Bibinfo	secundair	CD-ROM
Current Contents Agriculture	Bibinfo	secundair	CD-ROM
Current Contents Arts & Humanities	ERL	secundair	CD-ROM
Current Contents Clinical Medicine	Bibinfo	secundair	CD-ROM
Current Contents Engineering,Technology & Applied Sciences	Bibinfo	secundair	CD-ROM
Current Contents Life Sciences	Bibinfo	secundair	CD-ROM
Current Contents Physical, Chemical & Earth Sciences	Bibinfo	secundair	CD-ROM
Current Contents Sociological & Behavioural Sciences	Bibinfo	secundair	CD-ROM
Econlit	ERL	secundair	CD-ROM
Economist	Bibinfo	primair	CD-ROM
Elsevier Science Table of Contents	WWW	secundair	on-line
Eric	ERL	secundair	CD-ROM
Financial Times	Bibinfo	primair	CD-ROM
Financieel Ekonomische Tijd	Bibinfo	primair	CD-ROM
Food Sciences & Technology Abstracts	ERL	secundair	CD-ROM

Universiteitsbibliotheek Gent (1996)
https://web.archive.org/web/19961109120724/http://cbdata.rug.ac.be/e-dienst/tex/e-dbase-nl.html

That all changed in 1997, when we made a big splash launching the Executive Lounge that actually provided access to most of those databases and then some via the web, using the web as a unifying access technology. A lot of European academic libraries were impressed and we even received popular culture awards for the site.

Unfortunately, very few representative remnants of that site have been preserved. Fortunately, I wrote a paper about the effort in preparation of a conference in South Africa. Well, not a paper, really. More like a blog post on PDF. But it is available in the repository of the Ghent University Library. I do not know how it got there because it was never actually published. Anyhow, this is a screenshot taken from that paper showing part of the Executive Lounge menu system.

Exhibit 8

Current Contents: all editions

Database: CURRENT CONTENTS
Language: English
Vendor: Institute for Scientific Information (ISI)
Producer: Institute for Scientific Information (ISI)
DataType: Bibliographic

Agriculture, Biology, & Environmental Sciences

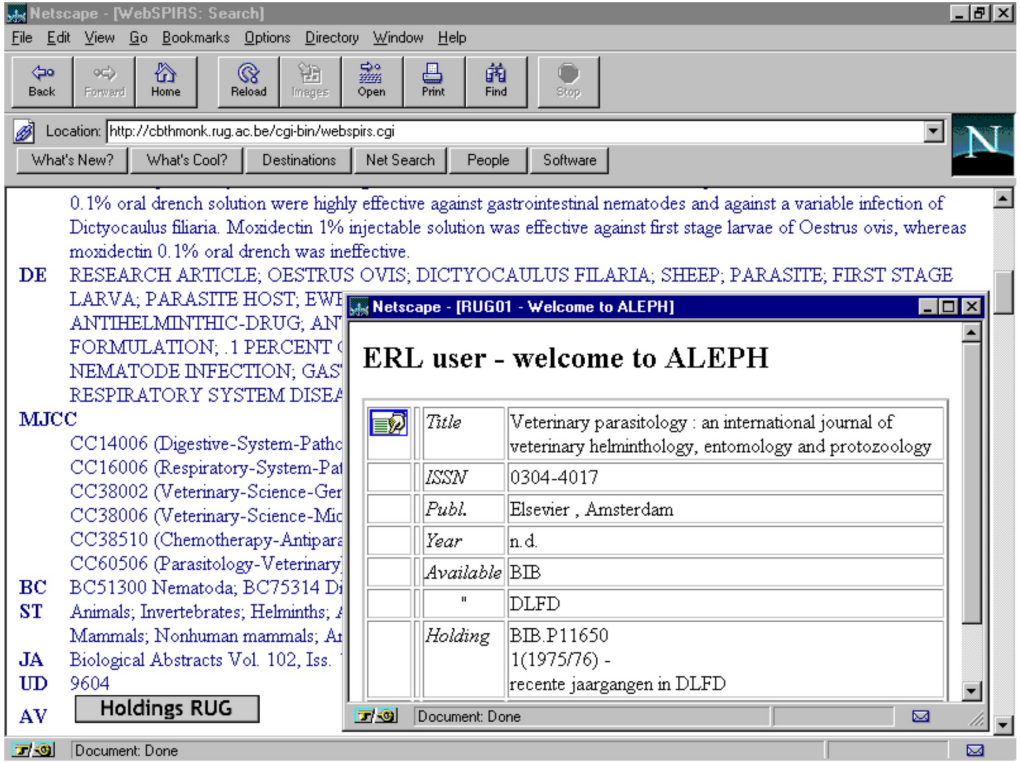
Van de Sompel, H., Hochstenbach, P., and De Pessemier, T. (1997) The hybrid information environment and our Intranet solution to access it (1997) Presented at Fourth South Africa Online Conference - https://hdl.handle.net/1854/LU-1056689

It is around that time that I also became obsessed with interlinking scholarly information. The Web had hyperlinks. I wanted to see those interlinking scholarly information too.

The early Executive Lounge already showed such linkages at work, the result of a collaboration with Ex Libris and SilverPlatter: dynamic links from Abstracting and Indexing databases to the library catalogue available via a “Holding at Ghent University” button.

CC BY SA

Executive Lounge, Linking from A&I Database to Catalogue

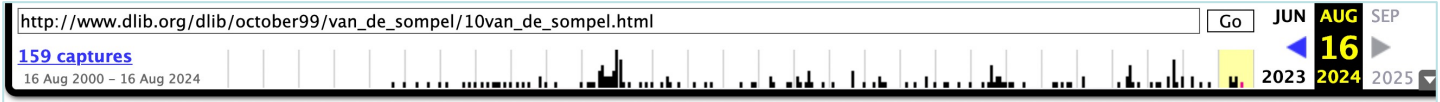


Van de Sompel, H., Hochstenbach, P., and De Pessemier, T. (1997) The hybrid information environment and our Intranet solution to access it (1997) Presented at Fourth South Africa Online Conference - https://hdl.handle.net/1854/LU-1056689

Eventually, this obsession with interlinking scholarly information led to the SFX/OpenURL library linking server work I did with Patrick Hochstenbach. It became the topic of my PhD thesis. The thesis is a bundling of papers – like this one - published in D-Lib Magazine. D-Lib Magazine has meanwhile has ceased publication but remains online. But this paper is also archived at the Internet Archive, and accessible through fatcat.

This SFX paper contained screencams that illustrated user interactions with a linking server. The screencams were provided as exe files for WinTel platforms and became obsolete quite a while ago. Thanks to advise of someone from this community – Andrew Weaver – I was recently able to migrate those screencams to a current format (Wine, CDROM image, VLC). The people at D-Lib were very helpful to make the migrated screencams accessible. And I pushed the updated version of the paper and the new screencams into the Internet Archive.

Exhibit 9



D-Lib Magazine
October 1999

Volume 5 Number 10

ISSN 1082-9873

Reference Linking in a Hybrid Library Environment

Part 3: Generalizing the SFX solution in the "SFX@Ghent & SFX@LANL" experiment

Herbert Van de Sompel
Los Alamos National Laboratory - Research Library
herbert.vandesompel@rug.ac.be

Patrick Hochstenbach
Automation Department of the Central Library
University of Ghent, Belgium
patrick.hochstenbach@rug.ac.be

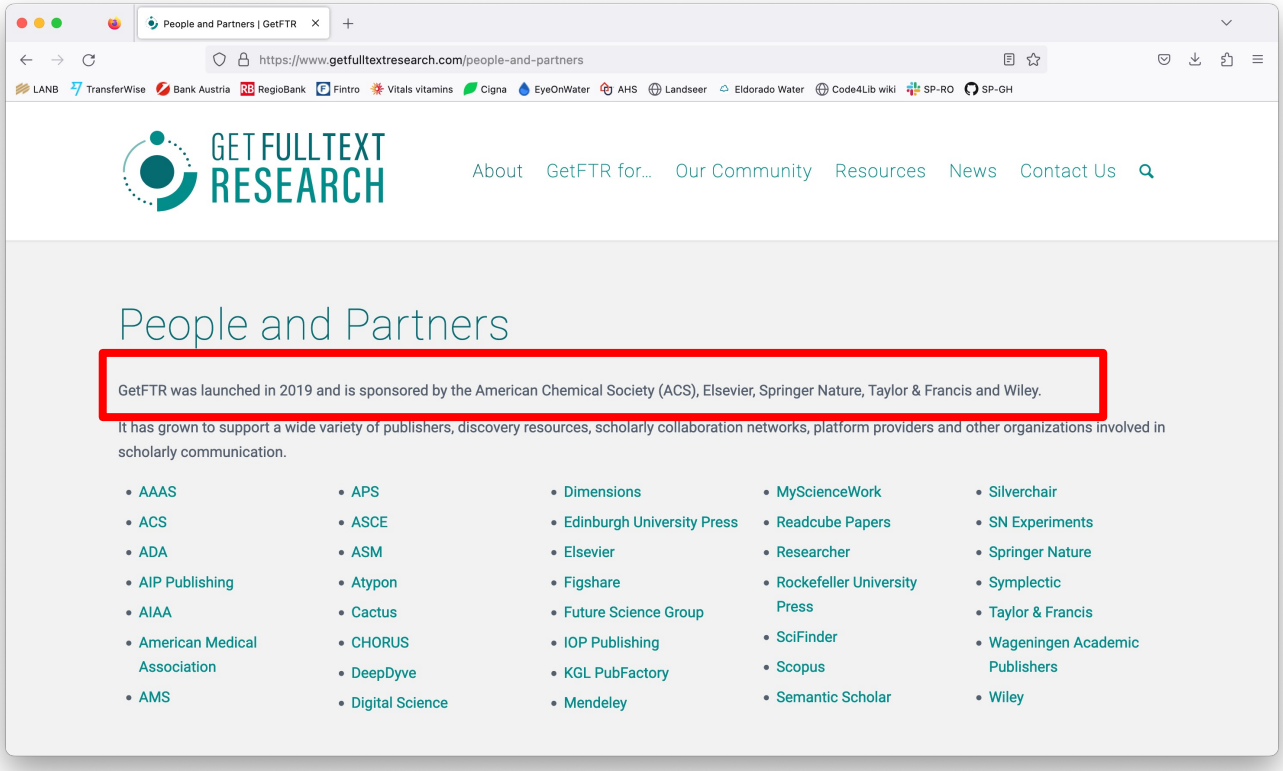
Editorial Note: This article includes 7 Screencams created as Lotus executable files for Windows that no longer play on any platform. In August 2024 the authors converted the screencams to MP4 files. Links to the MP4 versions have been added.

Van de Sompel, H., and Hochstenbach, P. (1999) Reference Linking in a Hybrid Library Environment. Part 3: Generalizing the SFX solution in the "SFX@Ghent & SFX@LANL" experiment.
https://web.archive.org/web/20240816072424/http://www.dlib.org/dlib/october99/van_de_sompel/10van_de_sompel.html

OpenURL became an ANSI/NISO standard, Patrick's SFX software got taken on board by Ex Libris and was commercialized. Over time, several competing library linking server products emerged.

Linking servers are still used by academic libraries, worldwide, but they are meanwhile facing heads-on competition from a centralized solution sponsored by the major commercial publishers, which technically consists of a combination of SeamlessAccess and GetFTR.

GetFTR: Easy Access to Content Researchers are Entitled to Read

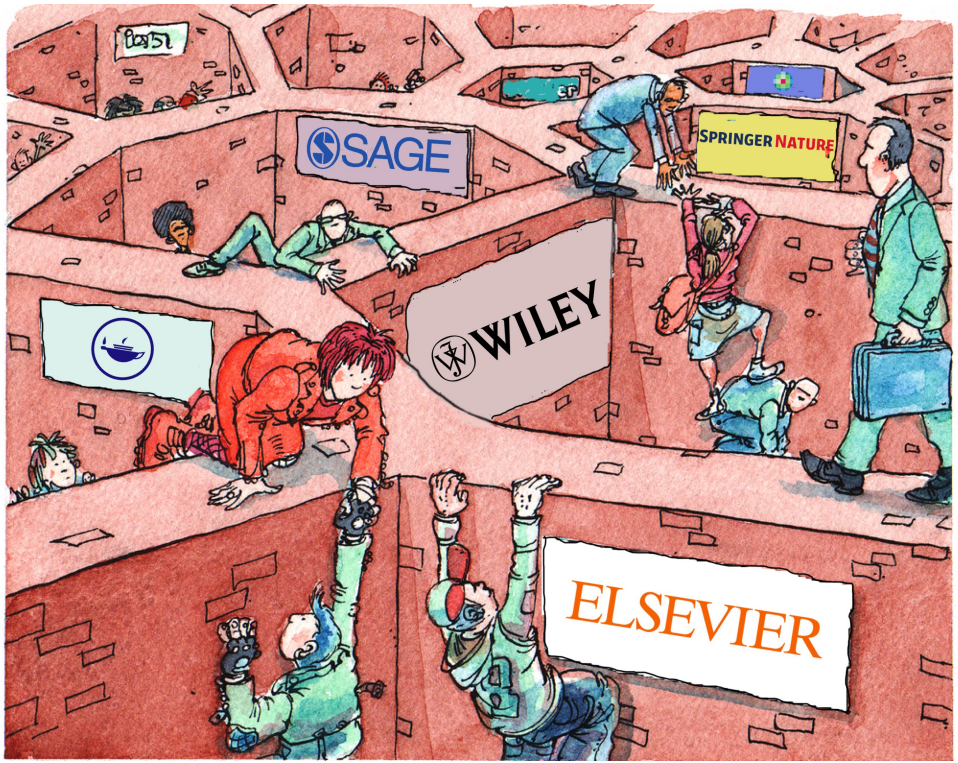


GetFTR (2019)
https://www.getfulltextresearch.com/people-and-partners

This is where I had planned to go off on a rant about the sorry state of the research communication system, the ongoing consolidation of publishing venues, the technical centralization that comes with it, the surveillance capability that results, the excessive profit margins, ... and the dire need for an alternative system.

But, since I consider this talk a celebration, “the Return of the Homeboy” or a belated retirement party if you will, I eventually decided not to go there. Been there, done that. If you want to hear me reminisce about that, I can highly recommend my 2017 CNI keynote.

Scholarly Communication: Deconstruct & Decentralize?



Van de Sompel, H. (2017) CNI Paul Evan Peters Award lecture - Scholarly Communication: Deconstruct & Decentralize? https://vimeo.com/248056214

So, I’m going to head in another direction here. A direction that came to me while compiling the slides so far. What we’ve seen thus far, using myself as an example:

1. Unsurprisingly, as we go about our lives, we’re leaving traces/artifacts all over the place. Some get captured as a result of the collection development policy of archives. Others, are never captured; artifacts slip between the maze although at some point they might have been important to us.
2. How does one know in which archives to look for our artifacts? In my examples, the Ghent Library archive is probably rather predictable. But the Österreichische Mediathek in Vienna less so.
3. If these artifacts are archived, how does one know that they are related, that they are traces left by the same individual? Where’s the aggregate view?

Archived Personal Traces

- **Some traces left by individuals get captured as a result of the collection development policy of archives, others slip between the maze** although they might have been important to the individual.

- **Discovery of archived personal traces** – No obvious way to determine in which archives an individual's traces can be found.

- **Inventory of archived personal traces** – No aggregate resource that connects these archived traces, that makes it obvious that the archived traces available in various archives were left by a same individual.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

I am going to take you on a journey in search of archival approaches that have a focus on traces left by individuals that try to address the challenges I just mentioned.

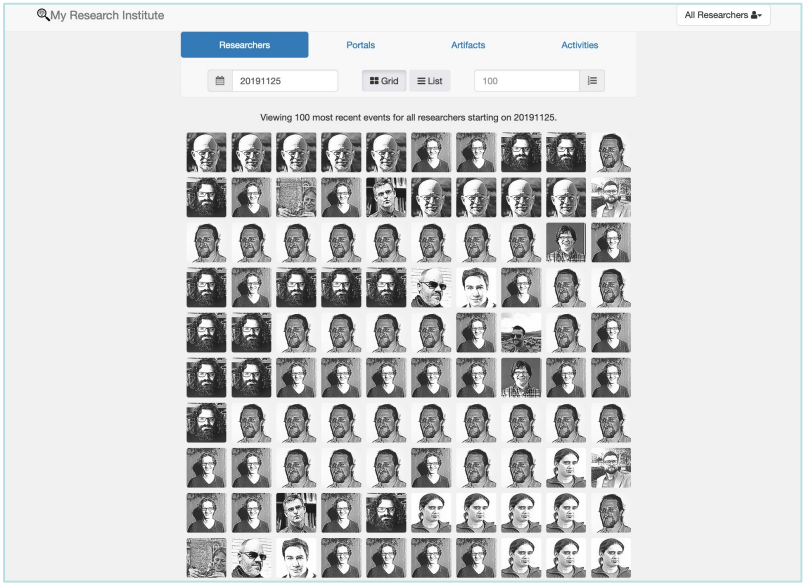
I will explore this in the narrower context of **research communication on the web**. Existing archival frameworks in research communication understandably take a systemic perspective, not a personalized perspective on archiving. **The attention is on the medium not on the messenger.**

I will talk about two experiments that look at archiving from another perspective: the myresearch.institute experiment that takes an institutional perspective and the mycontributions.info experiment that takes an individual perspective.

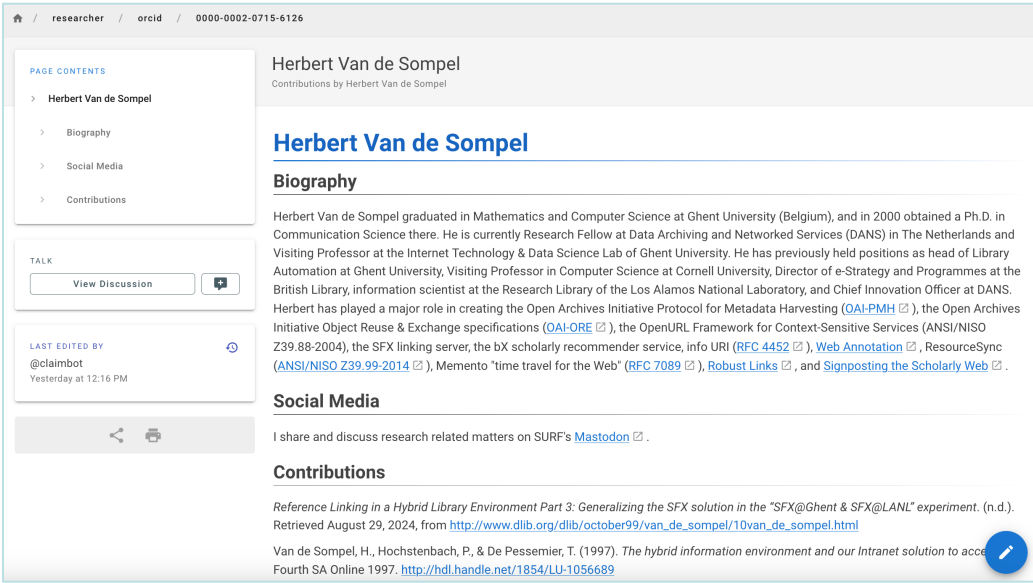
CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

myresearch.institute



mycontributions.info



CC BY SA

UNIVERSITEIT GENT

Data Archiving and Networked Services
DANS

But before digging into the experiments, I want to disclose the basic inspiration for both: Personal Web Observatories. What you see here is a screenshot of Amy Guy's personal web observatory. Well, a snapshot of a that observatory dating back to 2017.

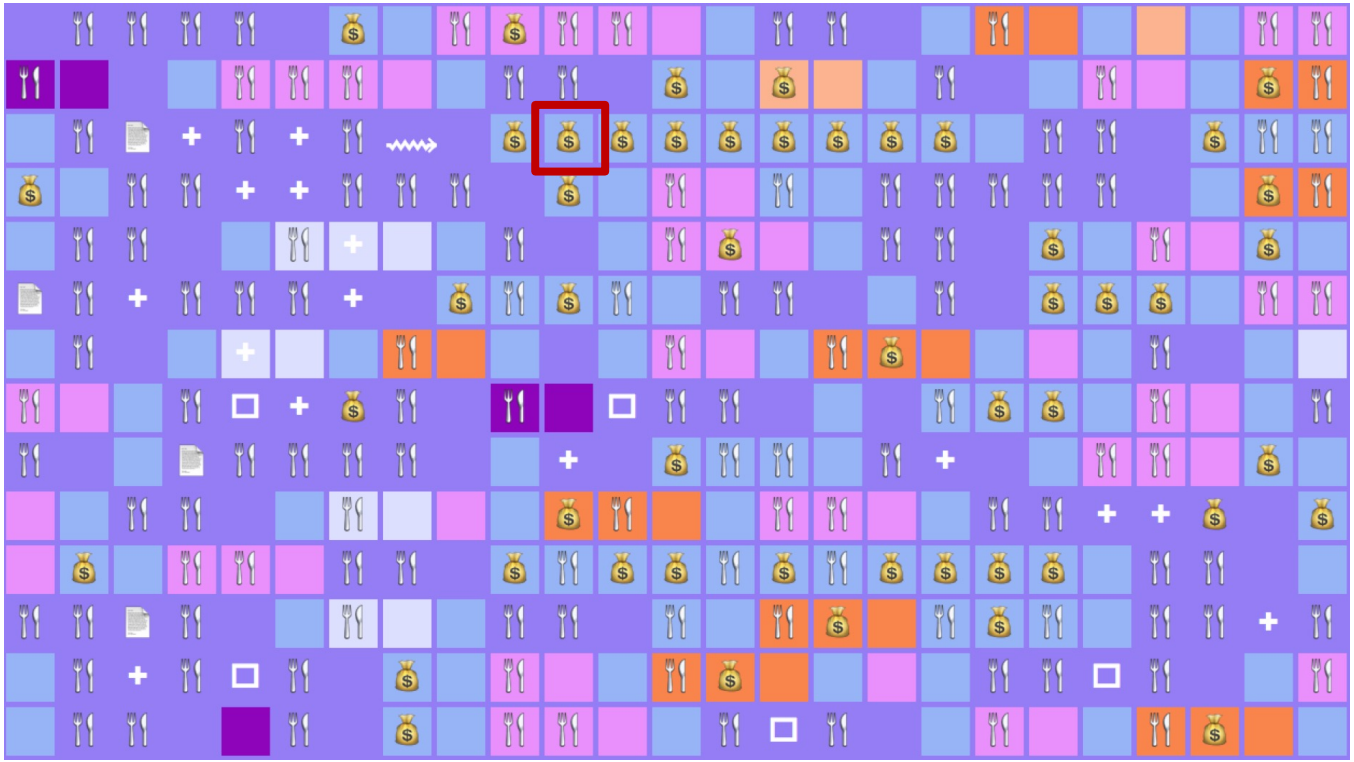
Amy Guy is a computer scientist with an education from the University of Edinburgh who has been involved in a range of W3C standardization efforts and is also on the W3C TAG.

The web observatory provides a visualization of the activities that Amy conducts in a variety of portals on the web. While leaving traces in those portals, Amy's personal site provides a summarization of it all. So, for example, when I click that money bag there …

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

Amy Guy's Personal Web Observatory



Amy Guy (2017) sloph
https://rhiaro.co.uk/

I can see that Amy bought some vegan food at Luton Airport and how much was spent on that. Amy recorded this using some app or portal. Other activities are recorded using other apps and portals and eventually they all get listed in Amy's personal web observatory. Each little square in the screenshot I showed is such an activity.

A Personal Trace

Tuesday the 5th of September at 7:25pm (+01:00)

Food (M&S, Luton Airport)

5.50gbp

---

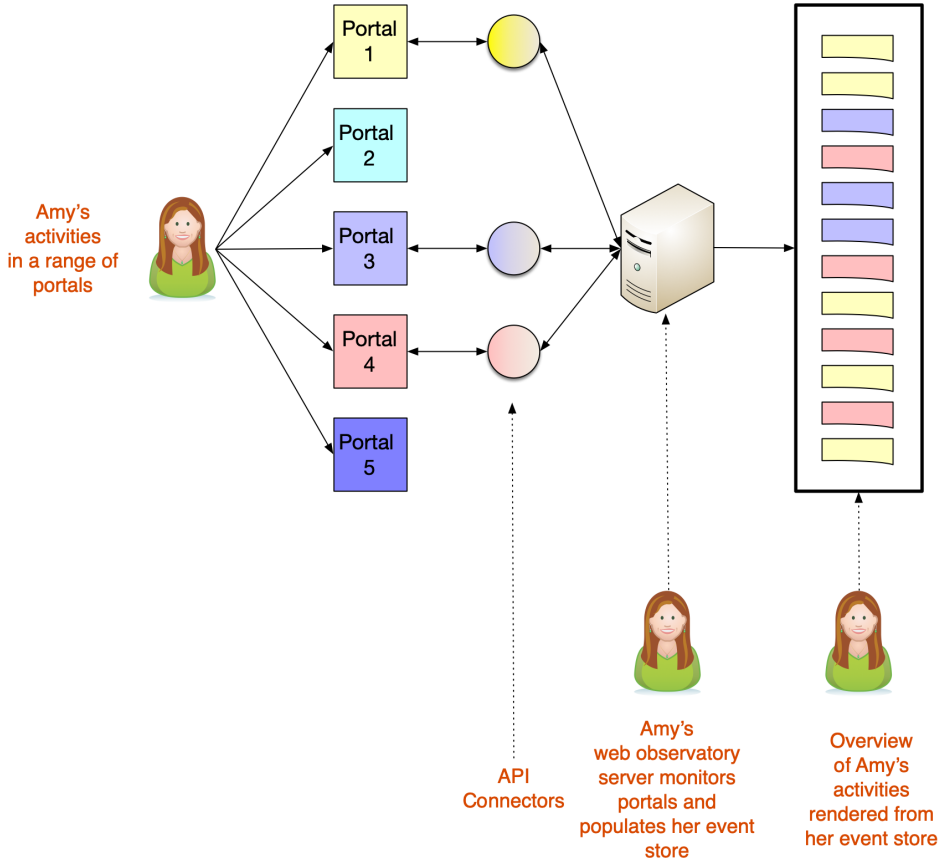
https://rhiaro.co.uk/2017/09/59aef516c7041

So, in essence, the premise of a personal web observatory is as follows:

- personal activities and conducted/recorded in variety of portals
- a personal server automatically monitors a select number of portals looking for new activities by the individual; the server uses connectors for the APIs of various portals to achieve this
- the activities recorded in those portals are collected and stored in a personal event store
- a user interface overlaying the personal store provides a cross-portal overview of personal activities

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS



based on: Amy Guy (2017) Observing the decentralised social web (one telescope at a time)
https://rhiaro.co.uk/2017/personal-web-observatory

The Personal Web Observatory work was a direct inspiration for the myresearch.institute experiment that was part of the Mellon-funded Scholarly Orphans project, a collaboration between my team at Los Alamos and Michael Nelson's team at Old Dominion University.

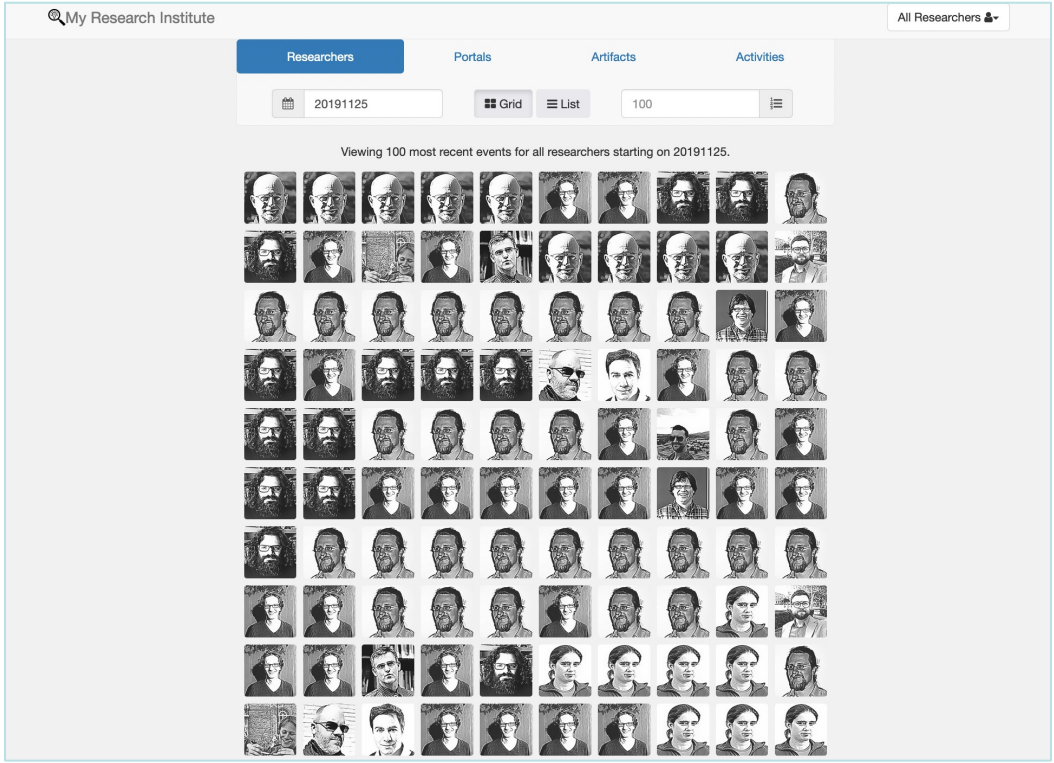
It was the last experiment I set up with my team at Los Alamos prior to leaving for DANS in The Netherlands. Martin Klein, who has meanwhile left Los Alamos himself, took over the experiment from me.

The results of the experiment are still online. The question is for how much longer, now that both Martin and I have left Los Alamos.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

myresearch.institute



CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

A core motivation for the Scholarly Orphans project was an insight that resulted from the Hiberlink project, a collaboration between the University of Edinburgh and Los Alamos led by Peter Burnhill, that focused on the archival status of web resources referenced in scholarly papers. It found that such resources, in many cases non-traditional scholarly resources, were very poorly archived.

The results of the Hiberlink research were published in PLOS One and I am happy to be able to publicly thank Andy Jackson for his approach to depict link rot and content drift over time that we used in this paper.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

Hiberlink – Poorly Archived References to Web Resources

Scholarly Context Adrift: Three out of Four URI References Lead to Changed Content

Shawn M. Jones, Herbert Van de Sompel, Harihar Shankar, Martin Klein, Richard Tobin, Claire Grover

Published: December 2, 2016 • https://doi.org/10.1371/journal.pone.0167475

Article | Authors | Metrics | Comments | Media Coverage

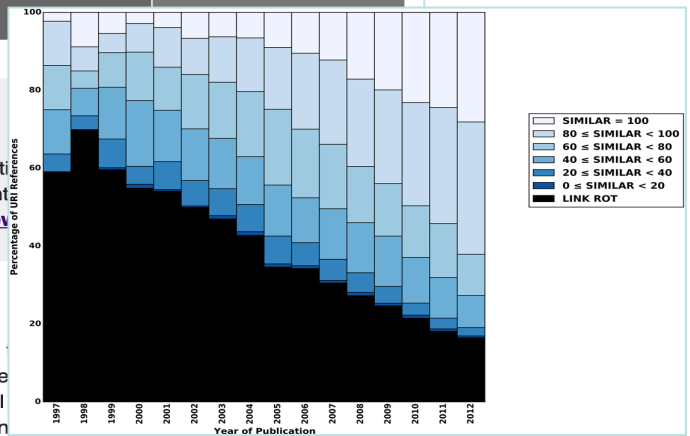
- Correction
- Abstract
- Introduction
- Related Work
- Methods
- Results
- Discussion
- Supporting Information
- Acknowledgments
- Author Contributions
- References

Correction

25 Jan 2017: The PLOS ONE Staff (2017) Correcti
out of Four URI References Lead to Changed Cont
https://doi.org/10.1371/journal.pone.0171057 | Viev

Abstract

Increasingly, scholarly articles contain URI references
project web sites, scholarly wikis, ontologies, online de
Authors reference such resources to provide essential
A reader who visits a web at large resource by followin
time after its publication, is led to believe that the resource's content is representative of what the author originally referenced. However, due to the dynamic nature of the web, that may very



Jones, S. M,, Van de Sompel, H. Shankar, H., Klein, M., Tobin, R., and Grover, C. (2016) Scholarly Context Adrift https://doi.org/10.1371/journal.pone.0171057

Overall the considerations at the basis of the project were that:

- Researchers use a variety of web platforms for collaboration and communication
- They use platforms that are dedicated to scholarship as well as general purpose
- The researchers' institutions are not aware of contributions made in general purpose portals and many times also not of those in scholarly portals that deal with less traditional contributions.
- Although contributions are not systematically archived, institutions have an incentive to obtain broad coverage of the contributions of their researchers. After all, institutions have set up institutional repositories and/or CRIS systems to cover contributions by their researchers to the research literature. Institutions have started setting up repositories to cover research data. It seems like a logical next step for institutions to start looking into systems that can cover a broader range of their researchers' outputs.

myresearch.institute - Considerations

- Researchers use a variety of web platforms for collaboration and communication.
- They use platforms that are dedicated to scholarship as well as general purpose.
- The researchers' institutions are not aware of contributions made in general purpose portals.
- Contributions are not systematically archived.
- Institutions have an incentive to obtain broad coverage of the contributions of their researchers.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

We decided to explore the potential of the Personal Web Observatory approach in an institutional setting, i.e. institutions operate infrastructure to automatically track contributions of researchers in web portals.

We created a fictitious research institution and populated it with 16 researchers who had given their consent to track them across a range of portals.

In our institutional observatory, the researchers were uniquely identified by means of their ORCID and in each portal they had a web identity that our observatory knew about. That is essential to be able to track them.

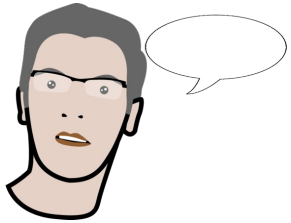
CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

myresearch.institute - Researchers

- 16 researchers
- Uniquely identified by ORCIDs
- Web identities in multiple portals
- Make a variety of contributions

CC BY SA



We selected 11 portals in which to track our researchers. Some are general purpose portals and some are focused on research: hypothes.is, blogger, slideshare, wordpress, github, figshare, wikipedia, home page, publons, stack overflow, ...

10.187 unique contributions were tracked between August 2018 and March 2020.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

myresearch.institute - Portals

- 11 portals
- Tracking started August 2018
- Tracking ended March 2020
- 10.187 unique contributions tracked

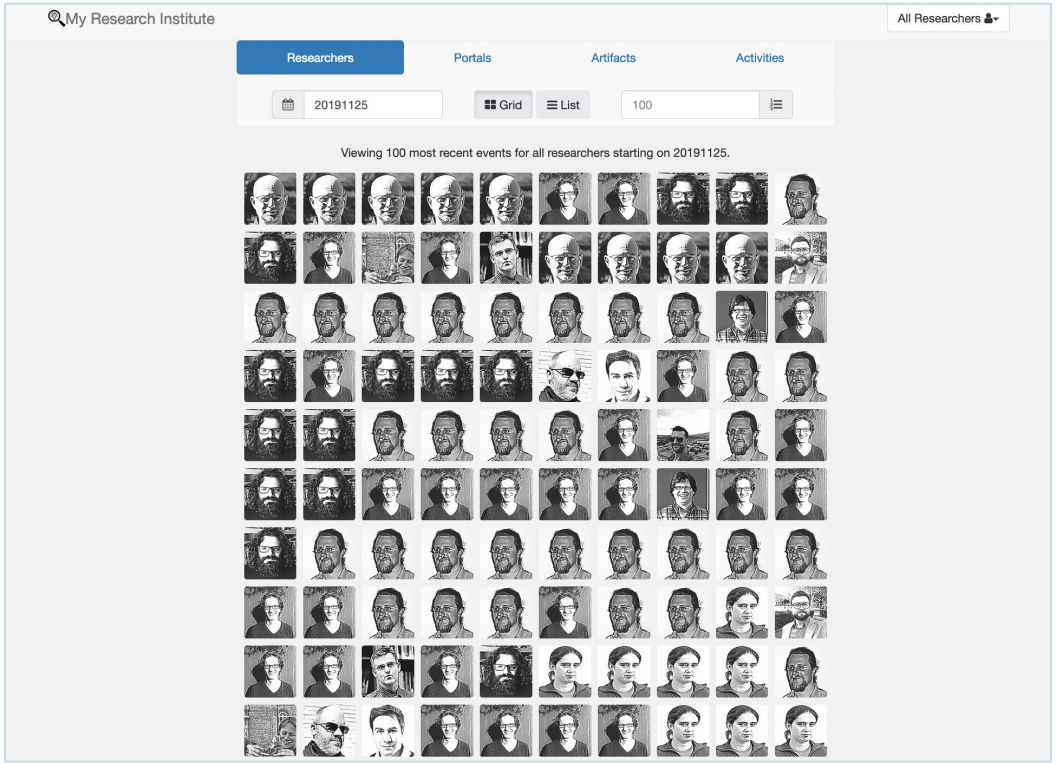
CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

The event store was gradually populated over time and a user interface, rather inspired by Amy Guy's personal web observatory, allowed getting an overview of our researcher's contributions.

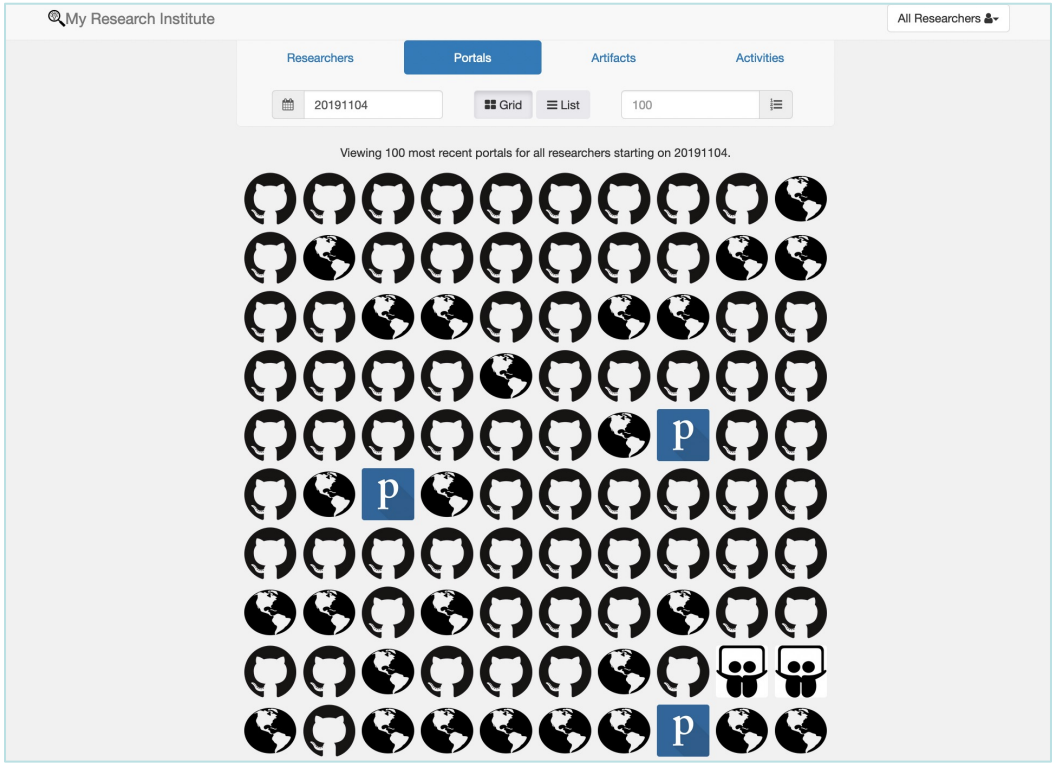
It provided ways to get a chronological overview of contributions by all researchers, organized by the researcher who made the contribution, the portal, the contribution type, and the activity type. It also allowed to zoom in on a single researcher.

myresearch.institute – Contributions - Researcher View



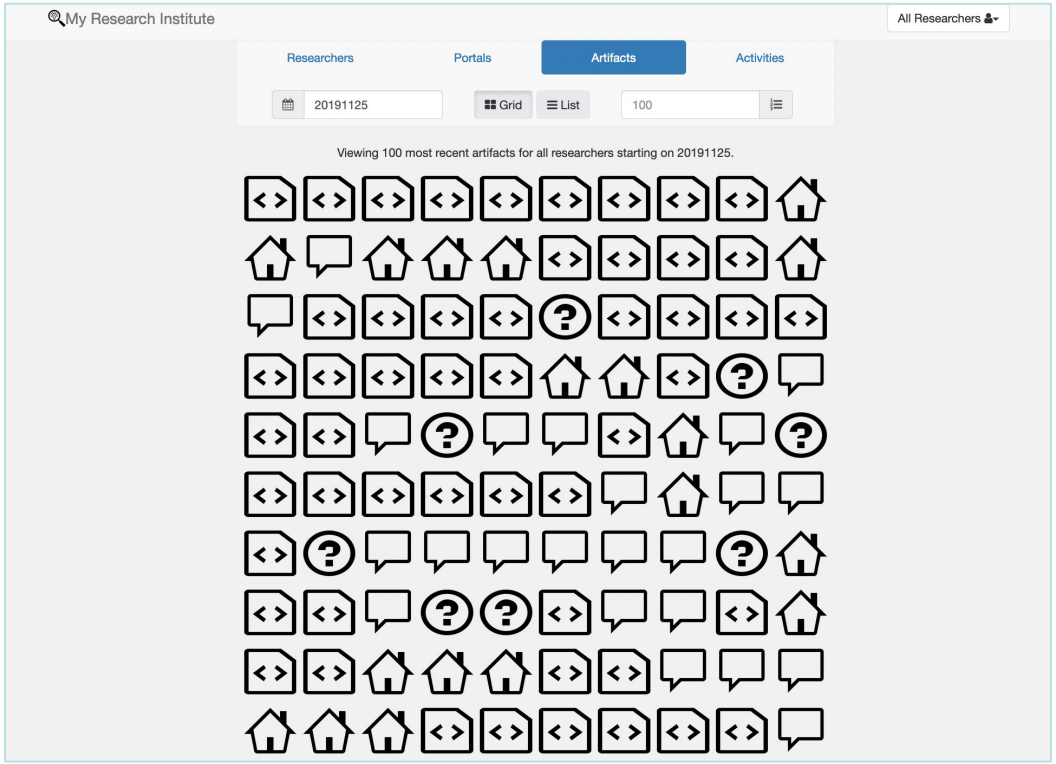
https://myresearch.institute/

myresearch.institute – Contributions – Portal View



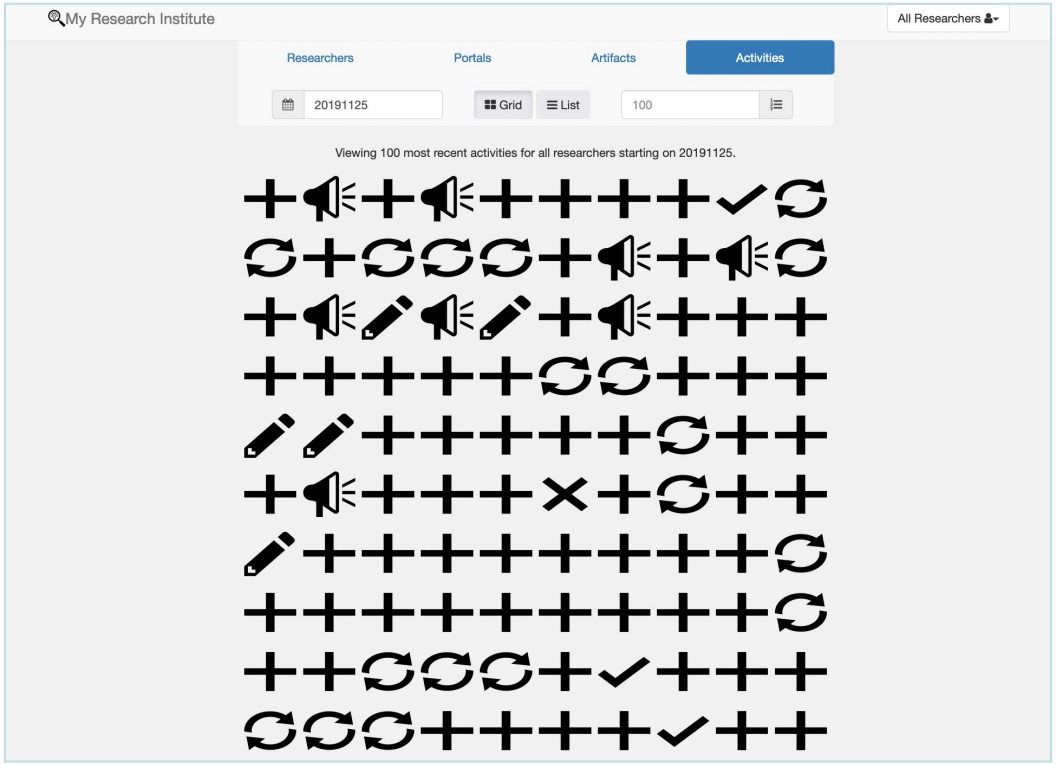
https://myresearch.institute/all/portals/

myresearch.institute – Contributions – Artifact Type View



https://myresearch.institute/all/artifacts/

myresearch.institute – Contributions – Activity Type View



https://myresearch.institute/all/activities/

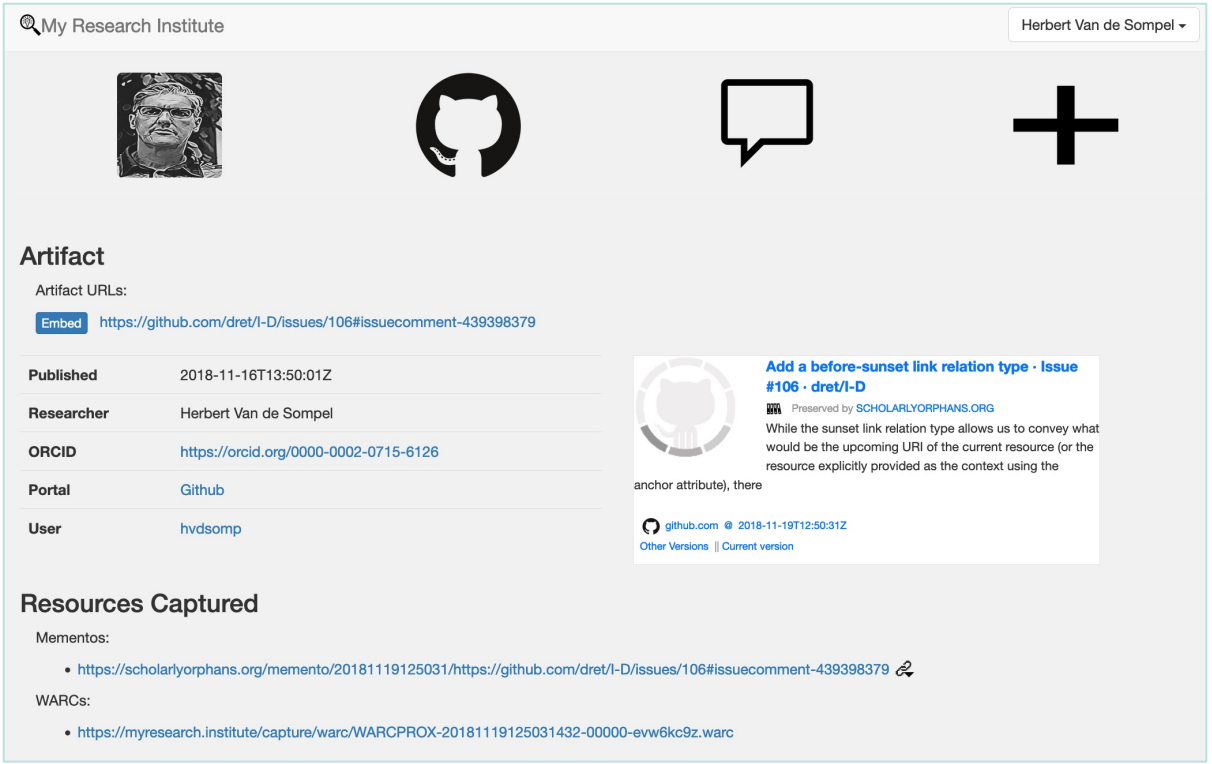
Details about a specific contribution were presented on a sort of landing page. This page is about a contribution I made to Github.

I did mention that a basic consideration of the Scholarly Orphans project was that these contributions were poorly archived. So, the event store, while useful in its own right as a means to get a broad perspective on researchers' contributions, also acted as a seed for web archiving those contributions. Contributions were captured and deposited both in the institutional observatory (as a WARC file) and in a cross-institutional scholarly web archive – scholarlyorphans.org - where it could be replayed.

The landing page provides access to these captures.

CC BY SA

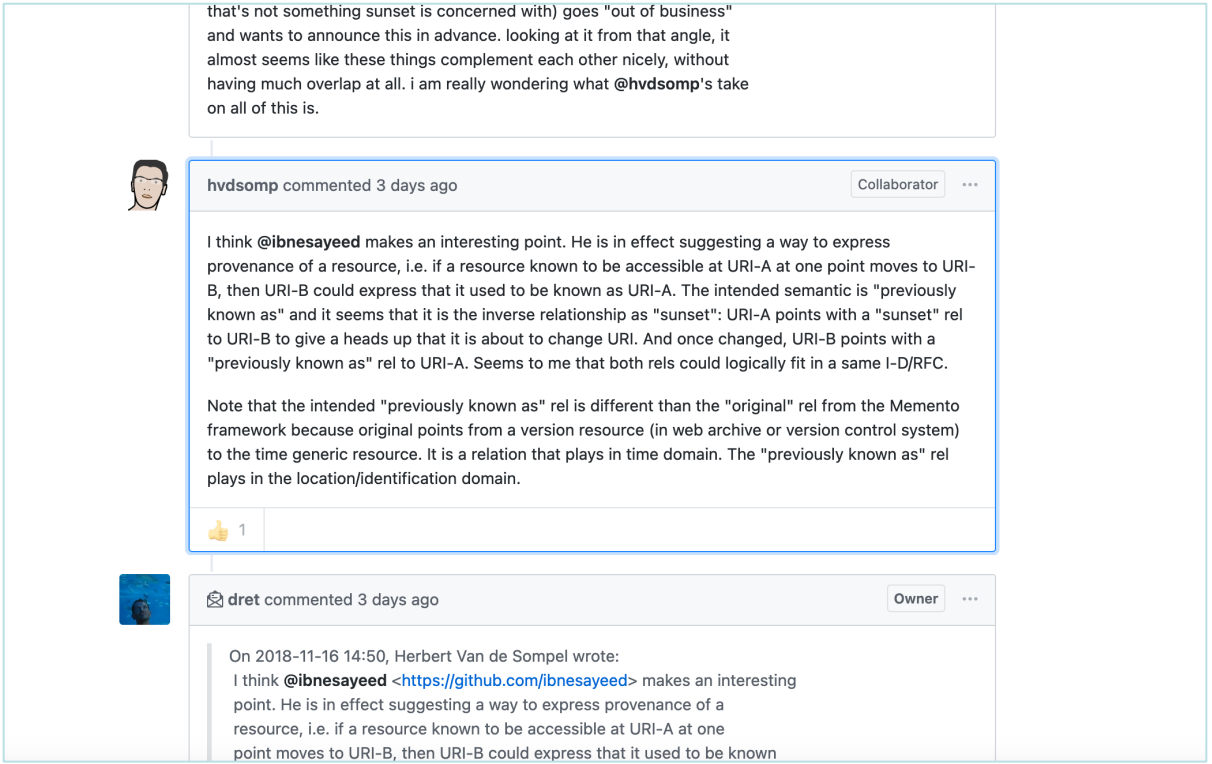
myresearch.institute – Landing Page for Contribution



https://myresearch.institute/event/7644e5d859e84360895d34a31573ee4f/

And this is a Memento of my Github contribution, archived at scholarlyorphans.org where it can also be replayed.

Memento of Contribution in scholarlyorphans.org



that's not something sunset is concerned with) goes "out of business" and wants to announce this in advance. looking at it from that angle, it almost seems like these things complement each other nicely, without having much overlap at all. i am really wondering what **@hvdsomp**'s take on all of this is.

hvdsomp commented 3 days ago — Collaborator

I think **@ibnesayeed** makes an interesting point. He is in effect suggesting a way to express provenance of a resource, i.e. if a resource known to be accessible at URI-A at one point moves to URI-B, then URI-B could express that it used to be known as URI-A. The intended semantic is "previously known as" and it seems that it is the inverse relationship as "sunset": URI-A points with a "sunset" rel to URI-B to give a heads up that it is about to change URI. And once changed, URI-B points with a "previously known as" rel to URI-A. Seems to me that both rels could logically fit in a same I-D/RFC.

Note that the intended "previously known as" rel is different than the "original" rel from the Memento framework because original points from a version resource (in web archive or version control system) to the time generic resource. It is a relation that plays in time domain. The "previously known as" rel plays in the location/identification domain.

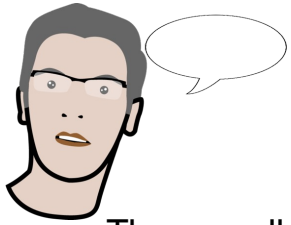
1

dret commented 3 days ago — Owner

> On 2018-11-16 14:50, Herbert Van de Sompel wrote:
> I think **@ibnesayeed** <https://github.com/ibnesayeed> makes an interesting point. He is in effect suggesting a way to express provenance of a resource, i.e. if a resource known to be accessible at URI-A at one point moves to URI-B, then URI-B could express that it used to be known

---

https://scholarlyorphans.org/memento/20181119125031/https://github.com/dret/I-D/issues/106#issuecomment-439398379



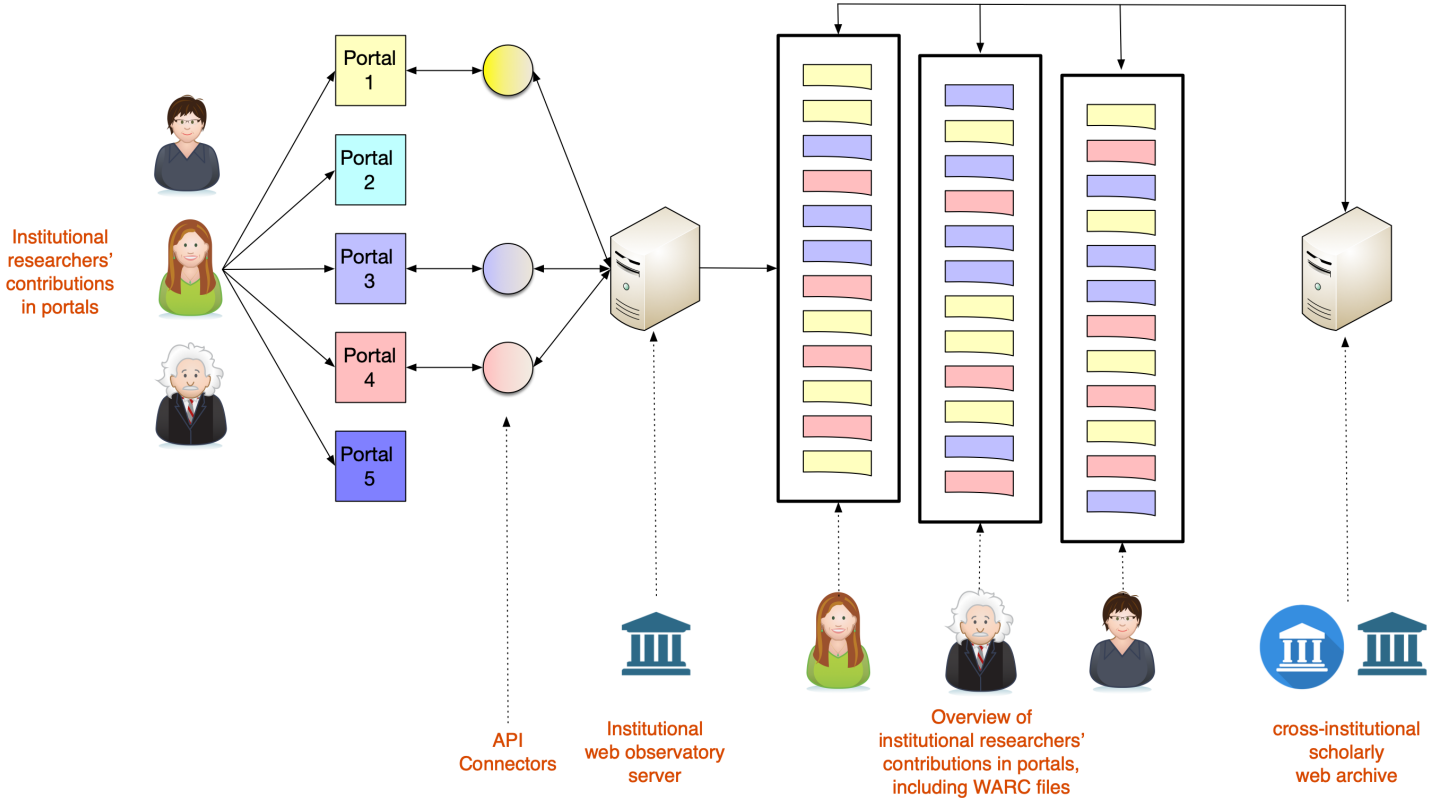
The overall architecture indeed looks very much like Amy Guy's Personal Web Observatory. But, here, multiple observatories are operated at the same time by an institution instead of by an individual.

Each observatory tracks a single researcher but populates a shared event database. The tracking techniques used are the same as in the Personal Web Observatory, i.e. API connectors continuously monitor remote portals; the web identities of researchers in each portal are essential to be able to do so.

The Institutional Web Observatory provides an overview of institutional and personal contributions.

Additionally, there is the archival component. Whenever a new contribution is detected, it gets captured and deposited as a WARC at the institution and as a replay-able capture in a cross-institutional scholarly web archive.

CC BY SA



CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

I want to share a few insights gained from this effort.

First, I mentioned that web identities of researchers are essential in order to be able to efficiently monitor platforms for new contributions.

We found that all general purpose web portals support such functionality: given the unique identifier of a user in a portal, retrieve a list of contributions, or even recent contributions.

We found that, despite the significant advocacy around unique contributor identifiers in scholarly communication scholarly portals typically did not have such API features. That was, of course, quite a few years ago but a recent survey conducted by colleagues at DANS, coordinated by Ingrid Dillo, found that not much has changed with that regard.

myresearch.institute - Insights

- Poor support in scholarly portals for retrieving contributions by means of a contributor identifier.

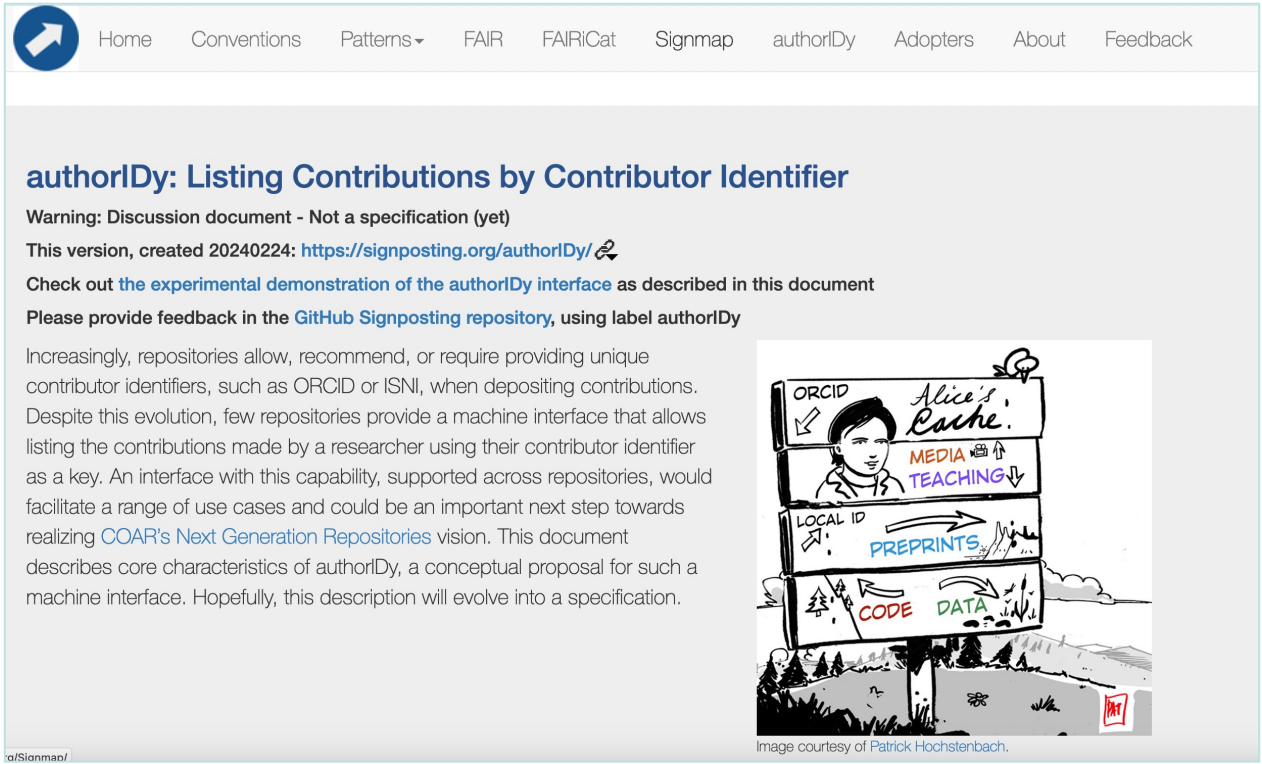
CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

This insight led to an effort to propose an API that could be supported by repositories and that can be used to list contributions by providing a contributor identifier.

It’s called authorIDy, it’s published under the Signposting umbrella, and is awaiting feedback. There was significant interest in it when I did a presentation at the recent Open Repositories conference in Götenburg. But we’ll have to see whether anything concrete happens with it.

authorIDy: Listing Contributions by Contributor Identifier



And there are also a few insights gained from this effort that informed the overall approach of the mycontributions.info experiment that I will talk about in a bit:

- It's not just institutions that do not know about researcher contributions in those web portals. The researchers know they make them but loose track of it all. As such, they much appreciated the overview provided in the institutional web observatory. They even asked whether our experimental setup had an API so they could get their stuff out.

- One can track only so many portals and obviously researchers will make contributions to portals that were not tracked

- Not all tracked contributions seem to be scholarly; there's definitely a mix of private and scholarly on general purpose platforms

CC BY SA

myresearch.institute - Insights

- Poor support in scholarly portals for retrieving contributions by means of a contributor identifier.
- Researchers don’t remember all the contributions they make and as such much appreciated the overview provided in the institutional web observatory.
- Researchers make contributions to portals that were not tracked; can’t track them all.
- Not all tracked contributions seem to be scholarly; there’s definitely a mix of private and scholarly on general purpose platforms.

CC BY SA

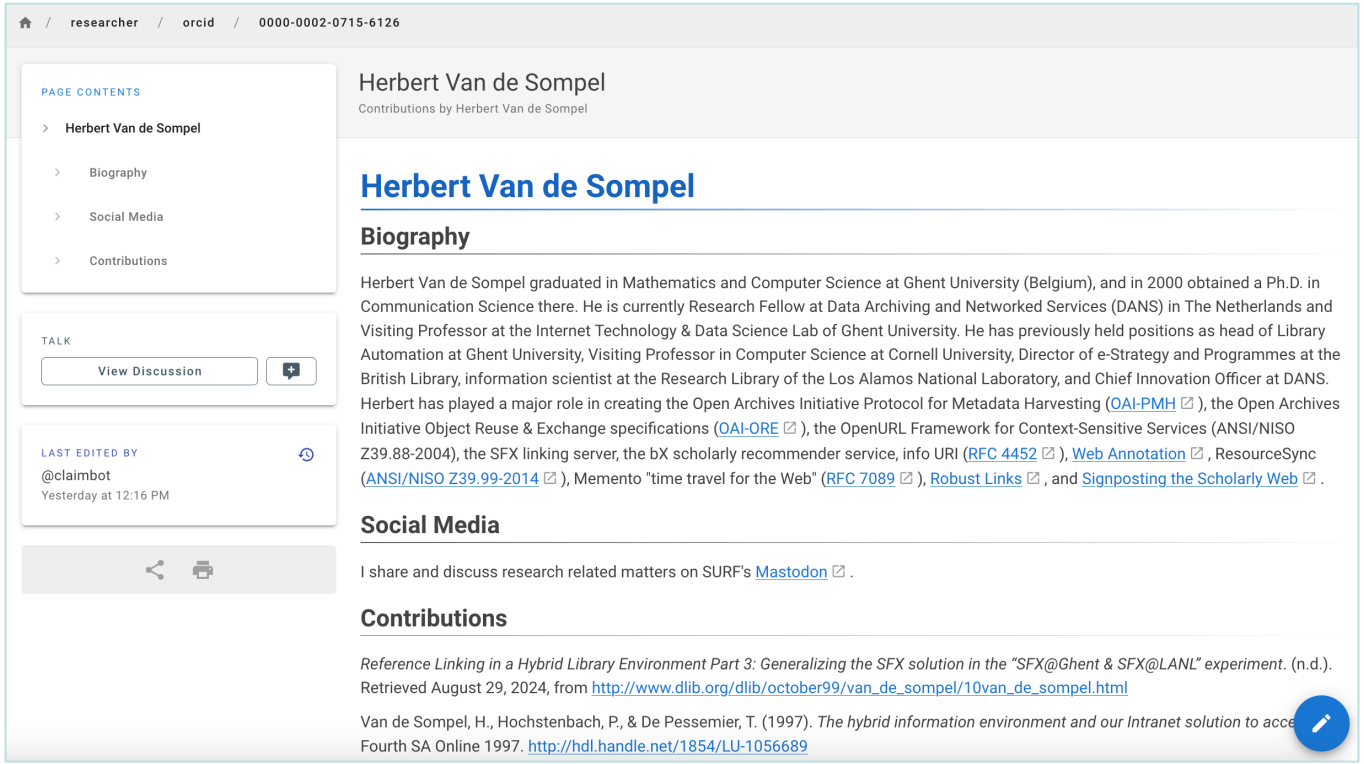
UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

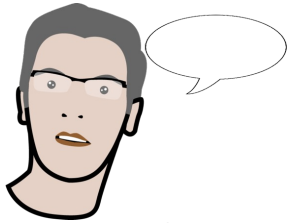
Which brings us to the mycontributions.info experiment that is currently being conducted by my long-term collaborator Patrick Hochstenbach.

That is, indeed, the same Patrick Hochstenbach who developed the first SFX linking server software 25 years ago. The mycontributions.info experiment is a collaboration with SURF and is part of Patrick's PhD research; Ruben Verborgh and I are his mentors on his PhD journey.

CC BY SA

mycontributions.info





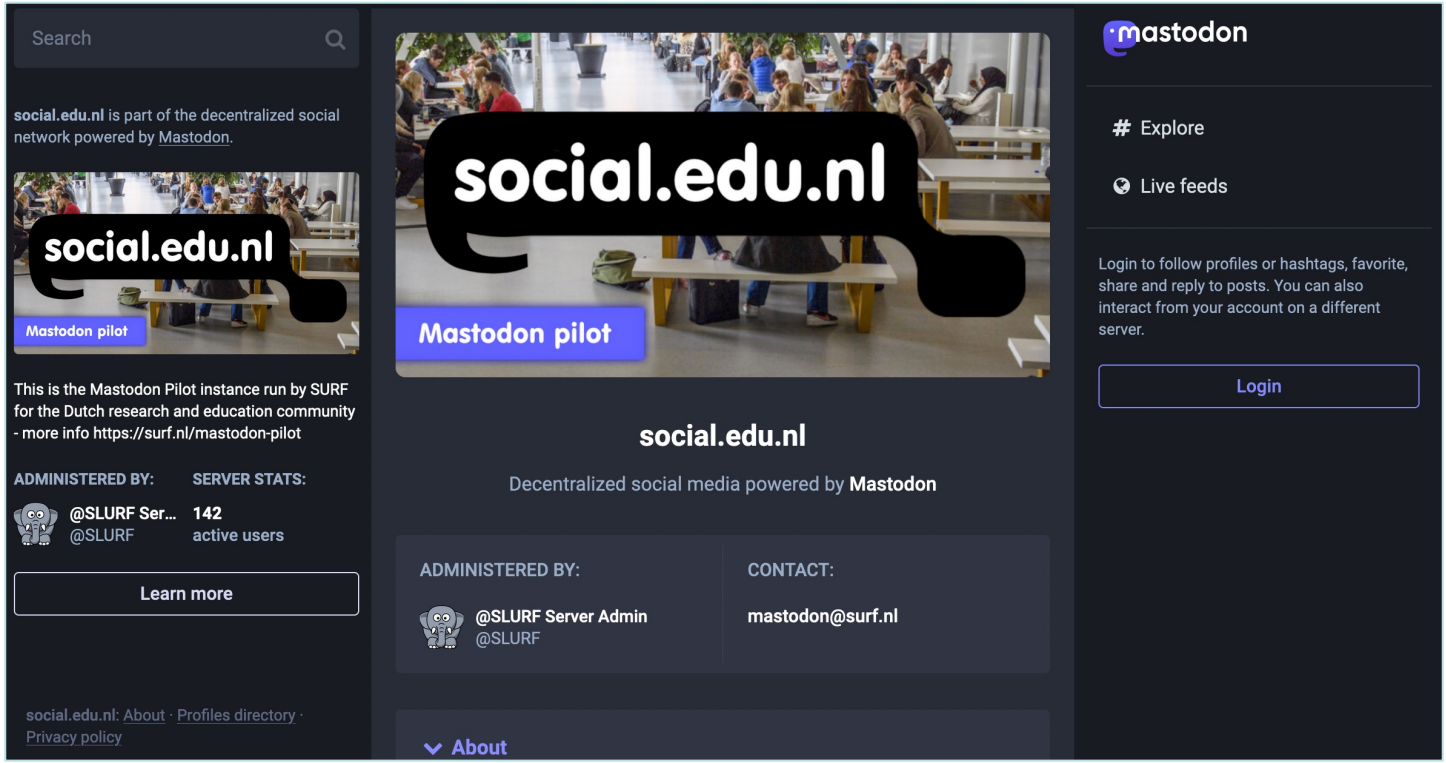
Before addressing the design of the experiment, I want to provide some context. SURF is the ICT cooperative of education and research institutions in the Netherlands. SURF allocates resources to explorative innovation activities in a range of areas. Societal relevance is an important criterion to decide which innovation activities to pursue.

Last year, as part of its Innovation activities, SURF set up a Mastodon instance. It has not been successful to this date, going by the “142 active users” stat.

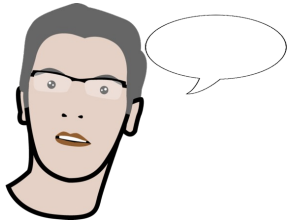
CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

SURF Mastodon Pilot



https://social.edu.nl/about



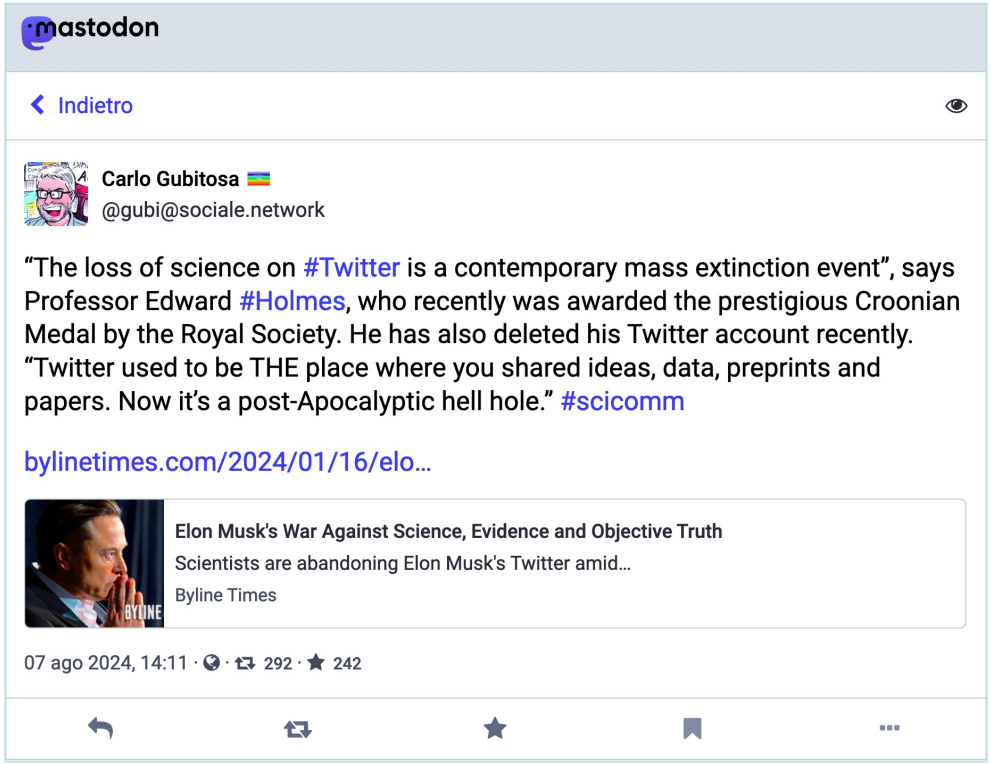
Earlier this year, SURF organized a brainstorm to discuss potential innovative activities in the realm of scholarly communication and the distributed web. Two main and seemingly unrelated themes/challenges emerged from the meeting.

The first theme was the disappearance of informal research communication on twitter resulting from Musk's take-over. It has been characterized as a contemporary mass extinction event. The question hanging around the room was whether the Fediverse with its distributed Mastodon instances could provide an alternative.

CC BY SA

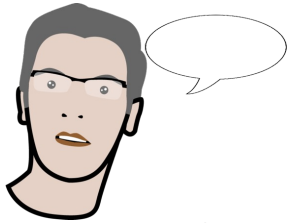
UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

Theme 1: Xodus, Loss of Science Communication on Twitter



https://sociale.network/@gubi/112920672314200247

Markolin, P. (16 January 2024) Elon Musk’s War Against Science, Evidence and Objective Truth https://bylinetimes.com/2024/01/16/elon-musks-war-against-science-evidence-and-objective-truth/



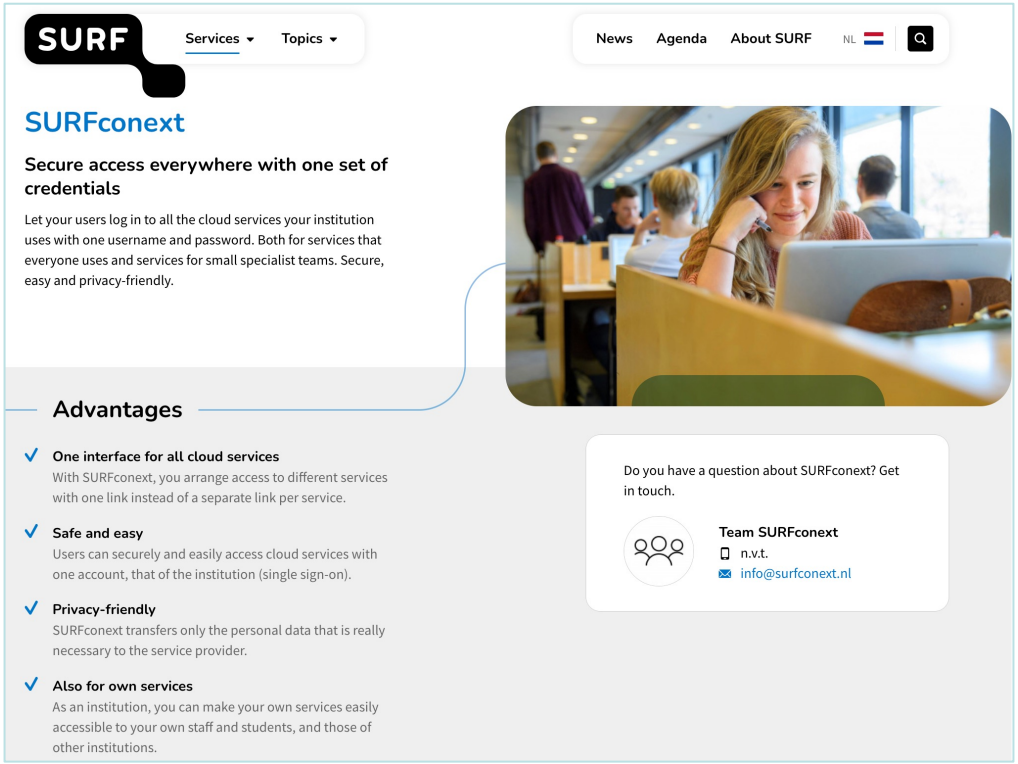
The SURF Mastodon instance has a very interesting feature with this regard: Users are registered and authenticated by means of SURFConext, the single sign-on service for high education in The Netherlands that leverages institutional identity providers. Meaning posts by users of the SURF Mastodon instance can be unambiguously attributed, traced back to an actual person who works or studies at a Dutch higher education institution. Which is very attractive from a research communication perspective.

But how to get them to join? What is the carrot?

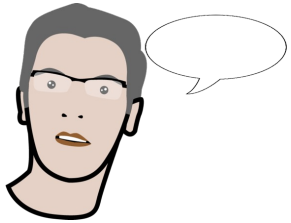
CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

SURFconext Used for Mastodon Registration/Authentication



https://www.surf.nl/en/services/surfconext



The second theme related to the efforts in the context of Open Science to assess researchers in novel ways, to take a broad range of activities/contributions into account for assessments.

- There is global activity in this realm, for example, under the umbrella of COARA, the Coalition for Advancing Research Assessment.

- And recently I came across a UNESCO document “Principles of Open Science Monitoring” that is in request for feedback status. It actually lists Roberto Di Cosmo as one on the contributing experts.

Theme 2 - Holistic Assessment of Researcher Contributions

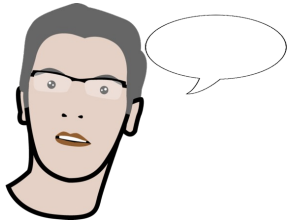
Coalition for Advancing Research Assessment

Our vision is that the assessment of research, researchers and research organisations recognises the diverse outputs, practices and activities that maximise the quality and impact of research. This requires basing assessment primarily on qualitative judgement, for which peer review is central, supported by responsible use of quantitative indicators.

Call for inputs: Global Consultation on the Draft Principles of Open Science Monitoring

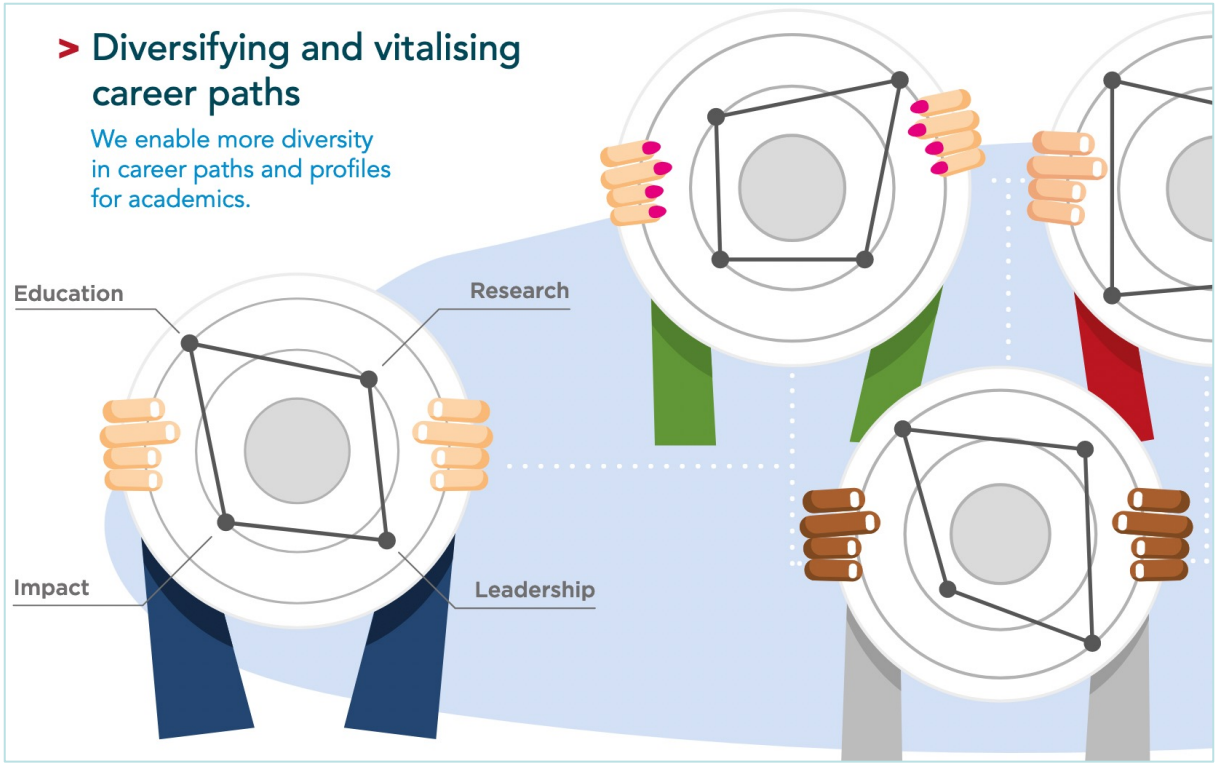
As part of the ongoing efforts to advance open science monitoring in alignment with the 2021 UNESCO Recommendation on Open Science, UNESCO is calling for inputs and comments from all the regions and interested stakeholders on the Draft Principles for Open Science Monitoring.
Deadline for inputs: 30 November 2024

https://coara.eu/
https://www.unesco.org/en/articles/call-inputs-global-consultation-draft-principles-open-science-monitoring



This desire to assess researchers in novel ways also very much exists in the Netherlands. And there is clear interest in areas that go beyond the creation of scholarly artifacts, including, under the ”impact” umbrella penetration of research results into society, the media, i.e. a newspaper article about a researcher’s work, an interview on the radio, an appearance on a TV program, etc.

Theme 2 - Holistic Assessment of Researcher Contributions



VSNU, NFU, KNAW, NWO and ZonMw (2019) Room for everyone's talent: towards a new balance in the recognition and rewards of academics
https://recognitionrewards.nl/wp-content/uploads/2020/12/position-paper-room-for-everyones-talent.pdf

A challenge with this regard, though, is how to get an overview of all pertinent activities/contributions in a manner that doesn't put all the burden on the researchers.

The EC-funded GRASPOS project operates in this realm and aims at generating holistic researcher profiles, for example, even including things like conference attendance. I heard OpenAIRE's Nathalia Manola talk about the project at the recent CRIS2024 conference in Vienna. She confirmed the challenge.

As we learned in the myresearch.institute perspective, one can't monitor it all. The monitoring approach can work when a select number of environments are deemed relevant but it breaks down when the environment that is to be monitored is the entire web.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

graspos – Open Research Assessment Dataspace

28

graspos Researcher Profile

Narrative that goes beyond research outputs

- Supported by evidence
- Using a variety of open data sources

How

- Different templates for different uses
- Novel ideas for presentation

Challenges

- How to capture information beyond research outputs
- ORCID extensions a valuable tool, but far from becoming extensively used

OpenAIRE

EuroCRIS 2024 | Vienna | May 15-17, 2024

CC BY SA

Manola, N. (2024) Transforming the Research Landscape: The Intersection of Open Science, Open Infrastructure and CRIS. Keynote at EuroCRIS2024, May 15-17, Vienna, Austria
http://hdl.handle.net/11366/2599

chiving and Networked Services
ANS

So, one needs to turn the whole thing around. And that was a major insight of the brainstorm at SURF: instead of detecting researchers' traces by monitoring portal APIs, have researchers themselves announce them, in the case of the SURF experiment by posting about them on Mastodon.

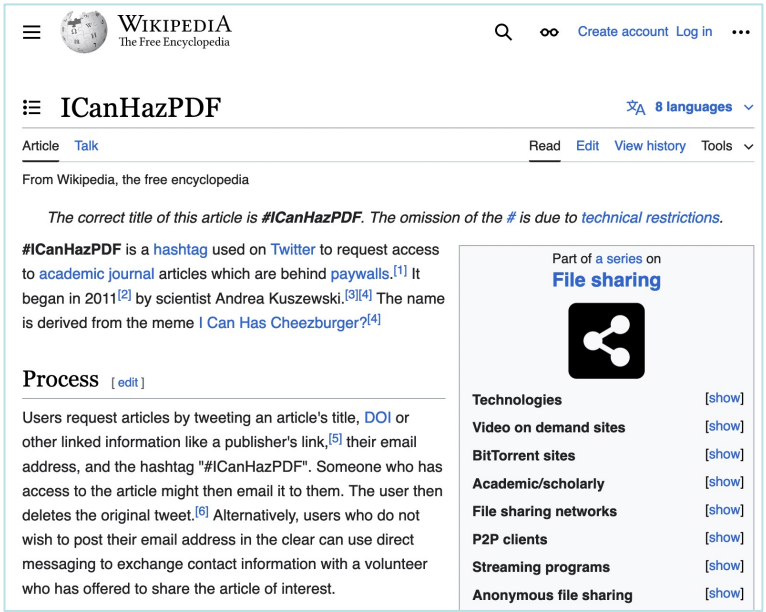
The specific approach chosen draws inspiration from tools that had been created on Twitter, such as #ICanHazPDF to obtain a copy of a research article that sits behind a paywall and @icanhazmemento to request the creation of a Memento for a certain URL. In this case, researchers will post about resources that they consider their contributions.

The incentive for the researcher to post is to obtain holistic data that can be used for assessments but, honestly, also to just keep track of what they've been up to. And the hope is that the carrot of the contribution list will be sufficient to attract Dutch researchers to Mastodon and start informal research communication there. We will learn whether that hope is justified once SURF decides on a community in which to run a pilot.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

#ICanHazPDF



WIKIPEDIA
The Free Encyclopedia

Create account Log in

ICanHazPDF

8 languages

Article Talk Read Edit View history Tools

From Wikipedia, the free encyclopedia

*The correct title of this article is **#ICanHazPDF**. The omission of the # is due to technical restrictions.*

#ICanHazPDF is a hashtag used on Twitter to request access to academic journal articles which are behind paywalls.[1] It began in 2011[2] by scientist Andrea Kuszewski.[3][4] The name is derived from the meme I Can Has Cheezburger?[4]

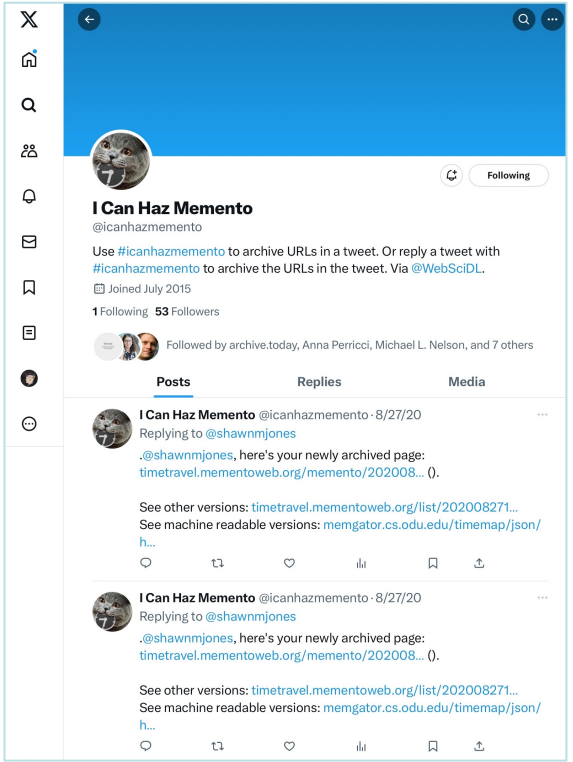
Process [edit]

Users request articles by tweeting an article's title, DOI or other linked information like a publisher's link,[5] their email address, and the hashtag "#ICanHazPDF". Someone who has access to the article might then email it to them. The user then deletes the original tweet.[6] Alternatively, users who do not wish to post their email address in the clear can use direct messaging to exchange contact information with a volunteer who has offered to share the article of interest.

Part of a series on
File sharing

Technologies	[show]
Video on demand sites	[show]
BitTorrent sites	[show]
Academic/scholarly	[show]
File sharing networks	[show]
P2P clients	[show]
Streaming programs	[show]
Anonymous file sharing	[show]

@icanhazmemento



Following

I Can Haz Memento
@icanhazmemento

Use #icanhazmemento to archive URLs in a tweet. Or reply a tweet with #icanhazmemento to archive the URLs in the tweet. Via @WebSciDL.

Joined July 2015

1 Following **53** Followers

Followed by archive.today, Anna Perricci, Michael L. Nelson, and 7 others

Posts Replies Media

I Can Haz Memento @icanhazmemento · 8/27/20
Replying to @shawnmjones
.@shawnmjones, here's your newly archived page: timetravel.mementoweb.org/memento/202008... ().

See other versions: timetravel.mementoweb.org/list/202008271...
See machine readable versions: memgator.cs.odu.edu/timemap/json/h...

I Can Haz Memento @icanhazmemento · 8/27/20
Replying to @shawnmjones
.@shawnmjones, here's your newly archived page: timetravel.mementoweb.org/memento/202008... ().

See other versions: timetravel.mementoweb.org/list/202008271...
See machine readable versions: memgator.cs.odu.edu/timemap/json/h...

https://en.wikipedia.org/wiki/ICanHazPDF
https://x.com/icanhazmemento

Here’s a high-level overview of the the design of the experimental setup that Patrick is currently working on.

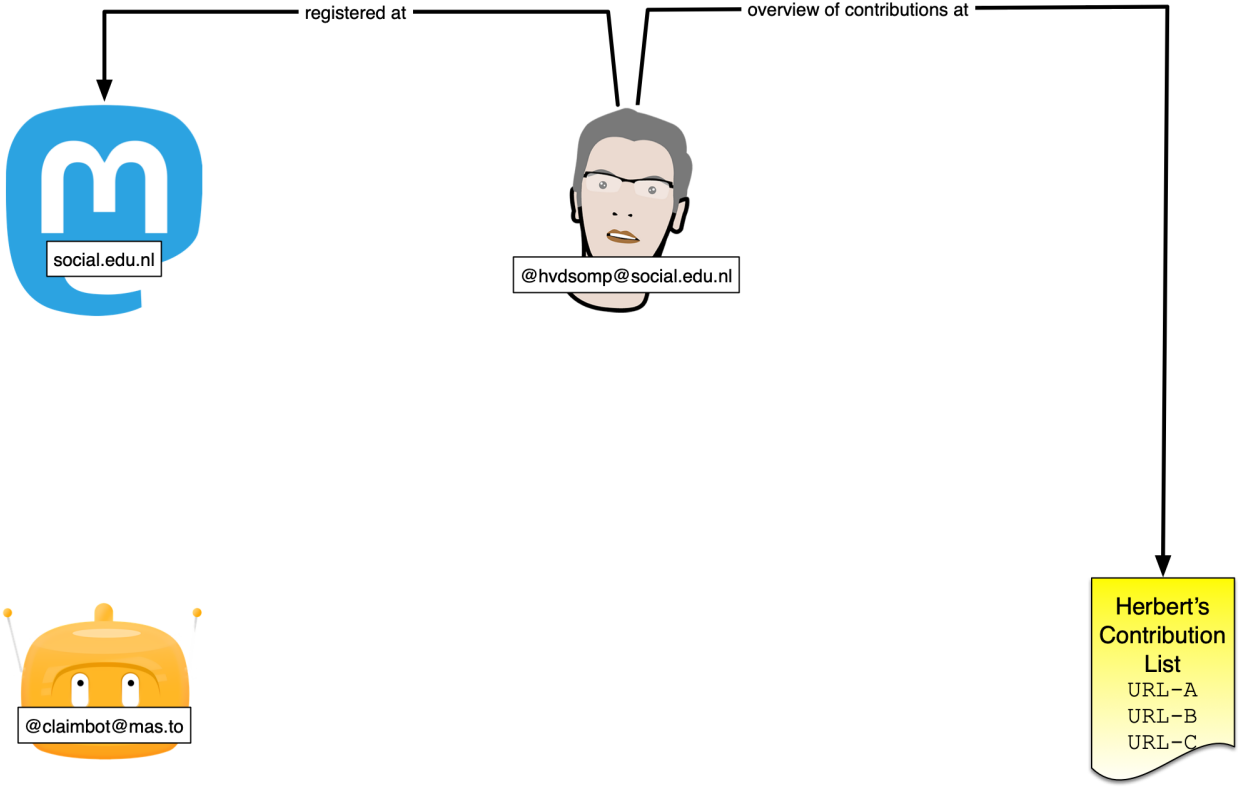
I am registered with the SURF Mastodon instance.

I have a Contribution List.

There’s also a bot in the Fediverse – the claimbot - that accepts requests from users registered at the SURF Mastodon instance.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS



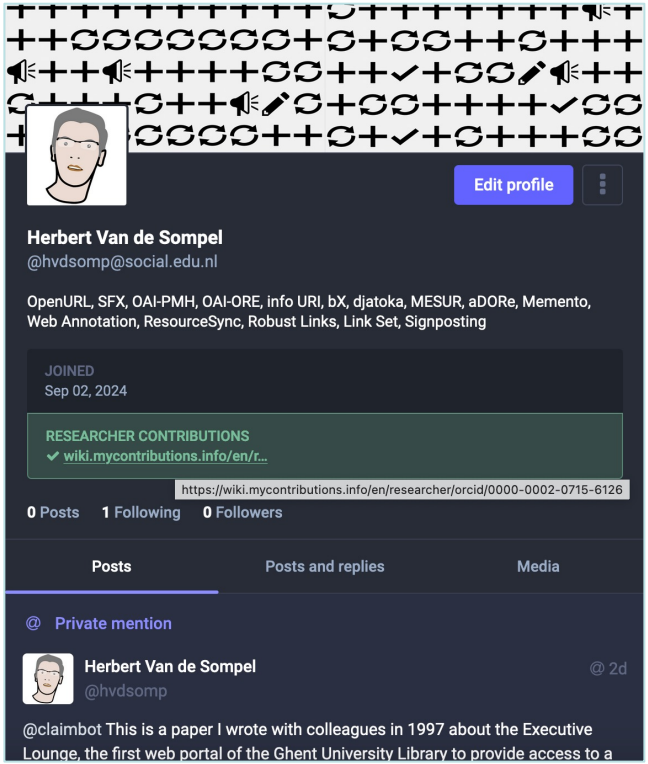
CC BY SA

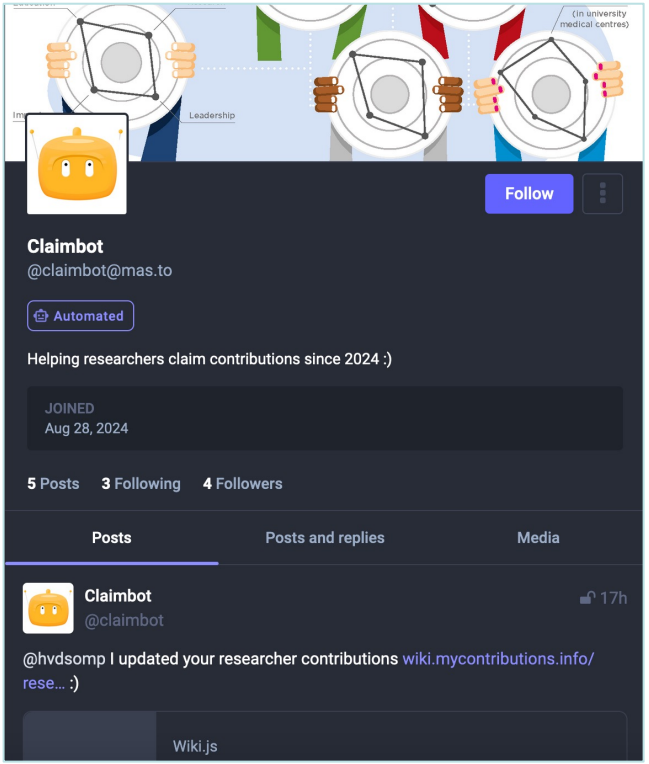
UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

To the left is my personal mastodon account on the SURF instance; you can seen the verified URL of my Contribution List there too.

To the right is the claimbot's mastodon account.

mycontributions.info – Personal and Bot Mastodon Accounts



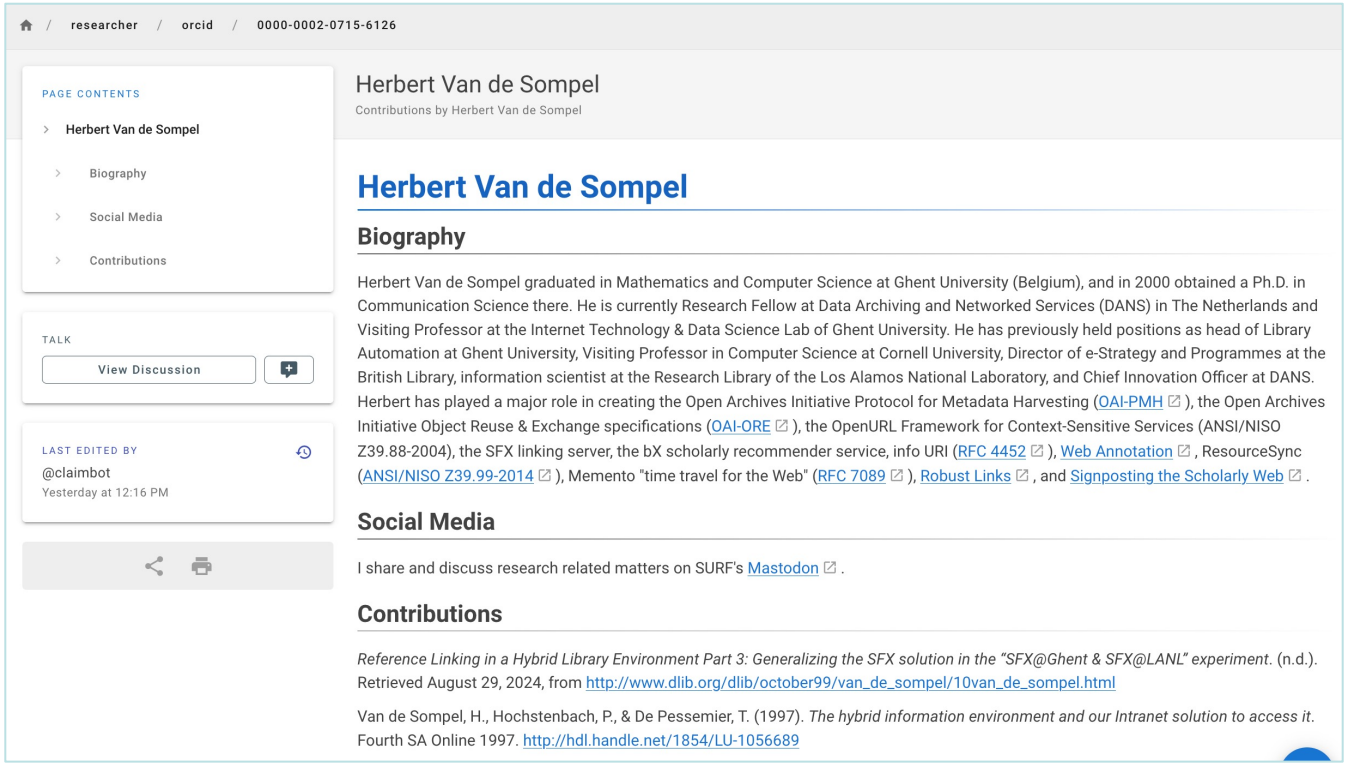


https://social.edu.nl/@hvdsomp
https://mas.to/@claimbot

This is my Contribution List that I have prepopulated with scholarly contributions I mentioned at the beginning of my talk: the D-Lib paper and the paper/blog about the Executive Lounge.

CC BY SA

mycontributions.info – Contribution List

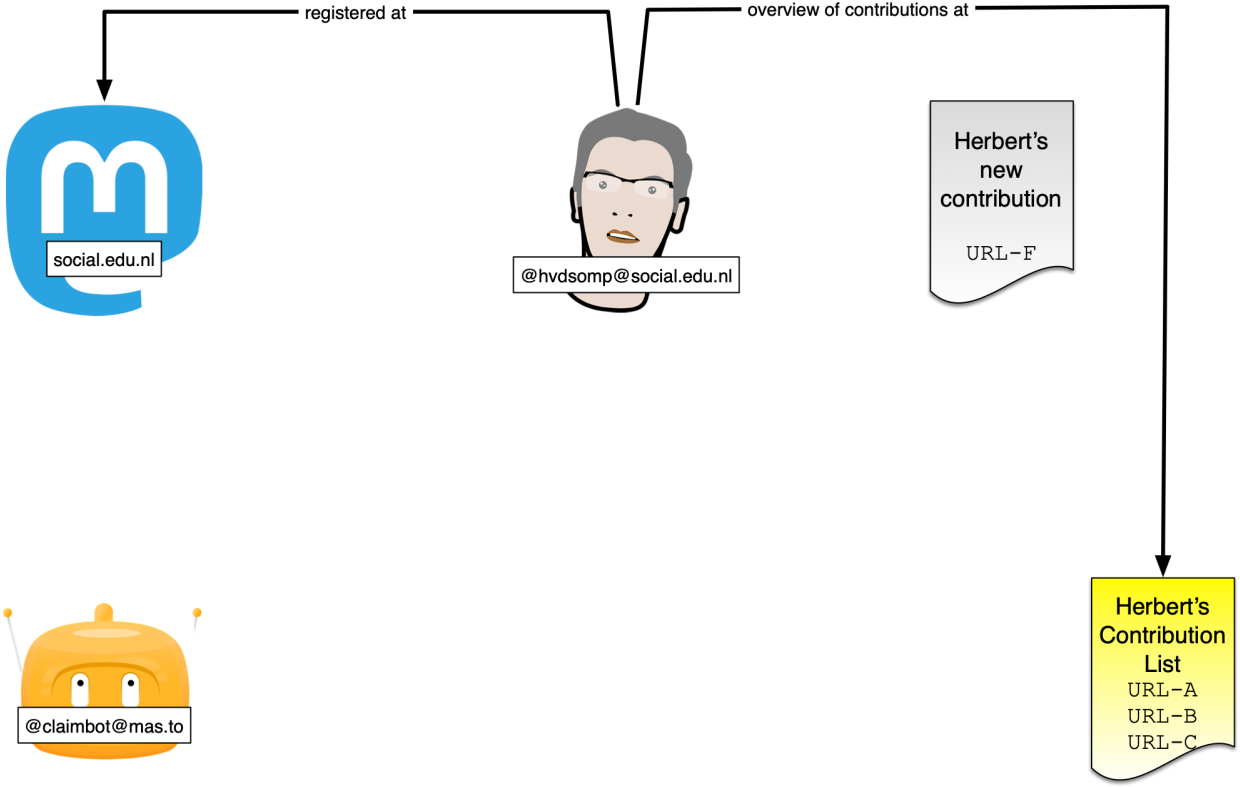


https://wiki.mycontributions.info/researcher/orcid/0000-0002-0715-6126

Let’s now introduce a new contribution I made, which I want to claim.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS



CC BY SA

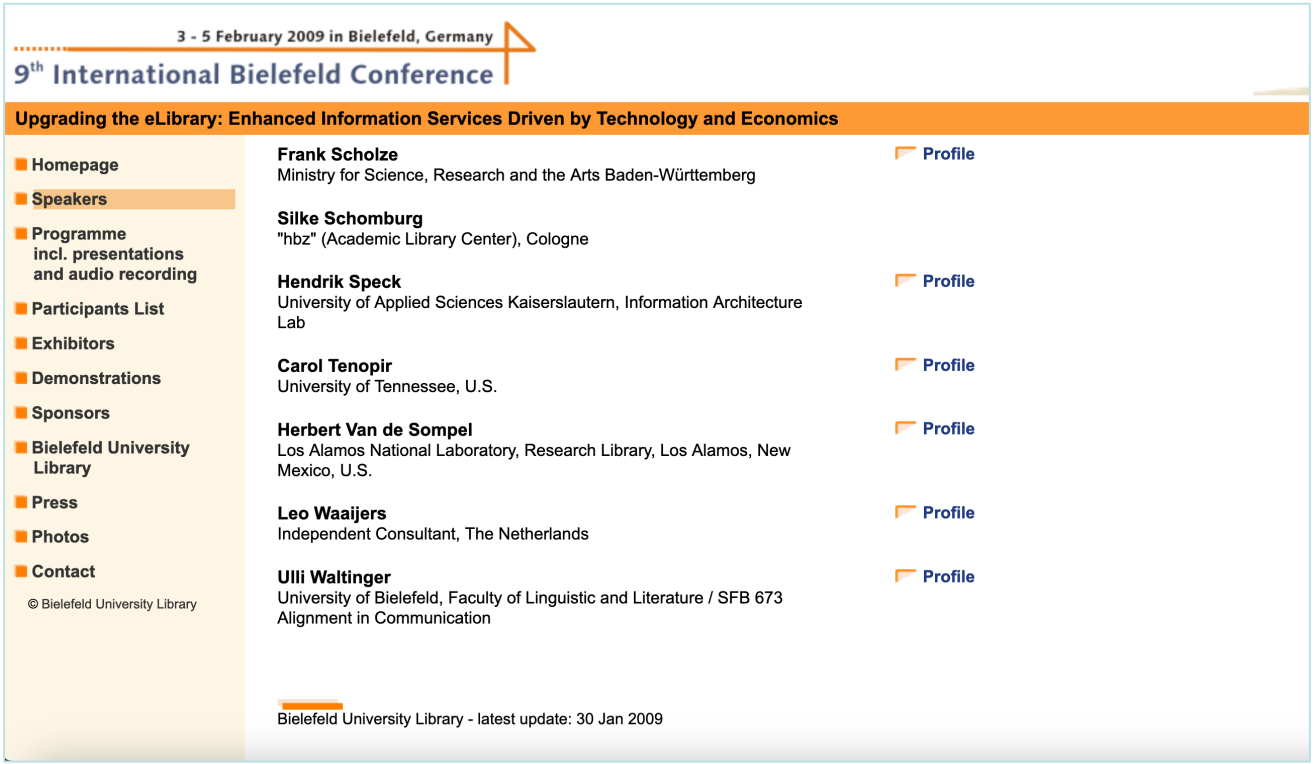
UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

In order to do so, we are going to roll back time for a moment and pretend it's 2009 again, when I was speaking at the Bielefeld Conference.

I am going to claim two new contributions with that regard:

- A page that shows that I am a speaker at the Bielefeld Conference
- The slides that I use for my talk

Contribution 1: Speaker at Bielefeld Conference



3 - 5 February 2009 in Bielefeld, Germany

9th International Bielefeld Conference

Upgrading the eLibrary: Enhanced Information Services Driven by Technology and Economics

- Homepage
- Speakers
- Programme incl. presentations and audio recording
- Participants List
- Exhibitors
- Demonstrations
- Sponsors
- Bielefeld University Library
- Press
- Photos
- Contact

© Bielefeld University Library

Frank Scholze
Ministry for Science, Research and the Arts Baden-Württemberg
Profile

Silke Schomburg
"hbz" (Academic Library Center), Cologne

Hendrik Speck
University of Applied Sciences Kaiserslautern, Information Architecture Lab
Profile

Carol Tenopir
University of Tennessee, U.S.
Profile

Herbert Van de Sompel
Los Alamos National Laboratory, Research Library, Los Alamos, New Mexico, U.S.
Profile

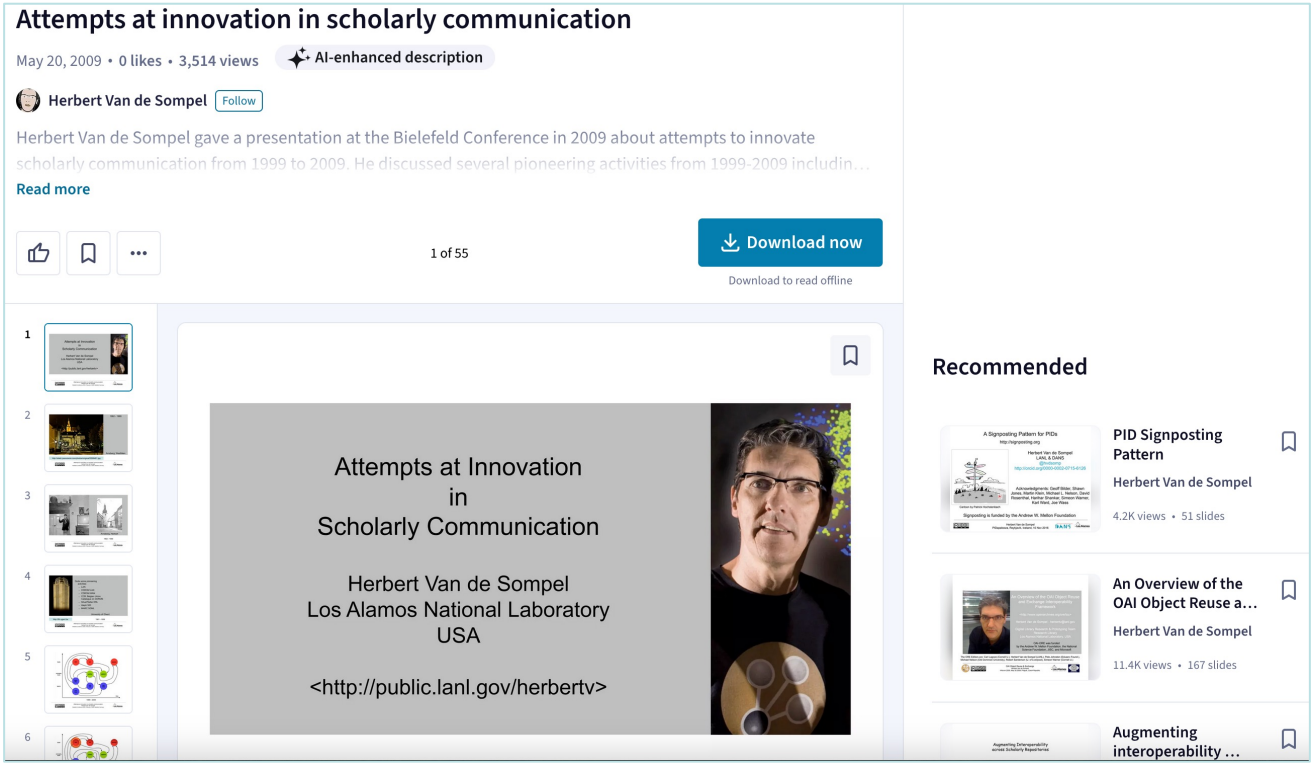
Leo Waaijers
Independent Consultant, The Netherlands
Profile

Ulli Waltinger
University of Bielefeld, Faculty of Linguistic and Literature / SFB 673 Alignment in Communication
Profile

Bielefeld University Library - latest update: 30 Jan 2009

http://conference.ub.uni-bielefeld.de/2009/speakers/

Contribution 2: Presentation Slides for Bielefeld Conference

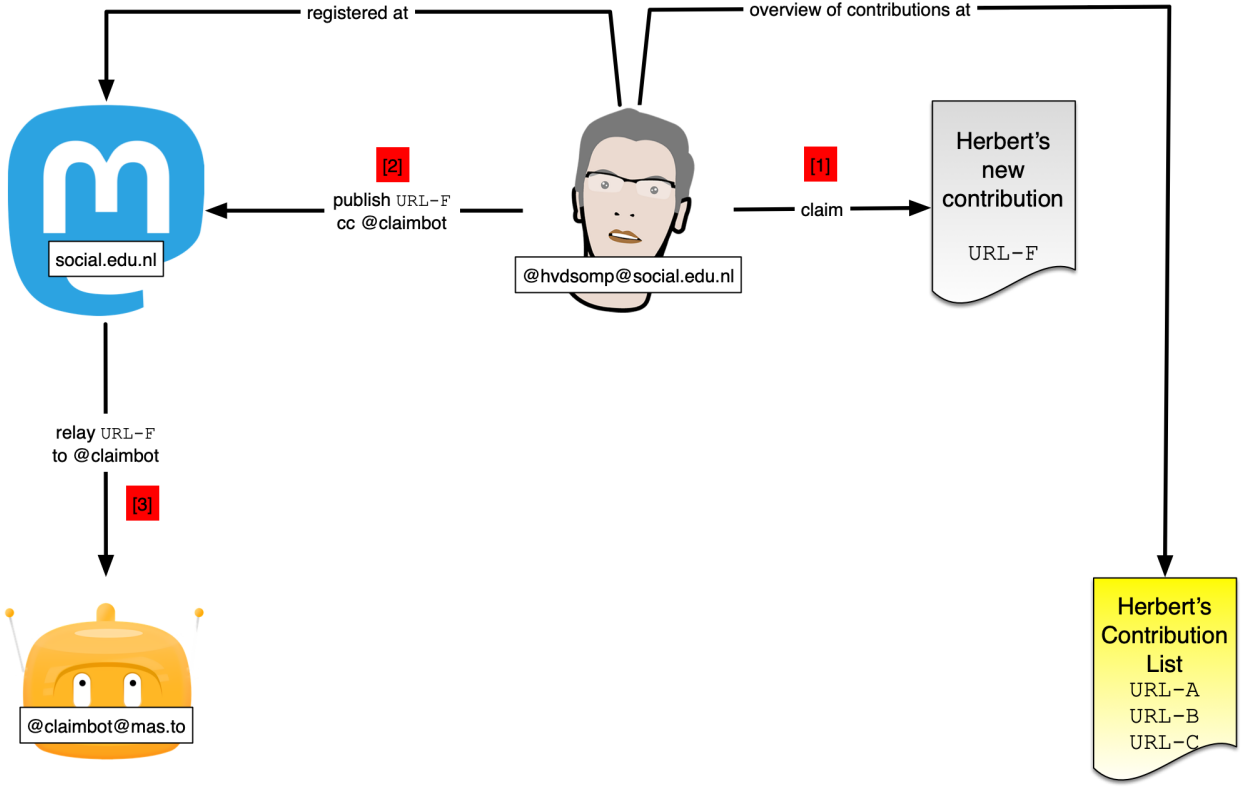


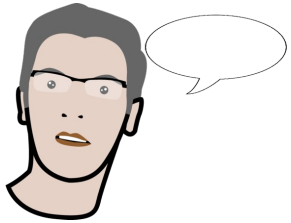
https://www.slideshare.net/slideshow/attempts-at-innovation-in-scholarly-communication/1466198

So, I am going to create Mastodon posts with the URL of these pages and I will target them at the claimbot.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS


registered at
overview of contributions at
social.edu.nl
[2]
publish URL-F
cc @claimbot
@hvdsomp@social.edu.nl
[1]
claim
Herbert's new contribution
URL-F
relay URL-F
to @claimbot
[3]
@claimbot@mas.to
Herbert's Contribution List
URL-A
URL-B
URL-C

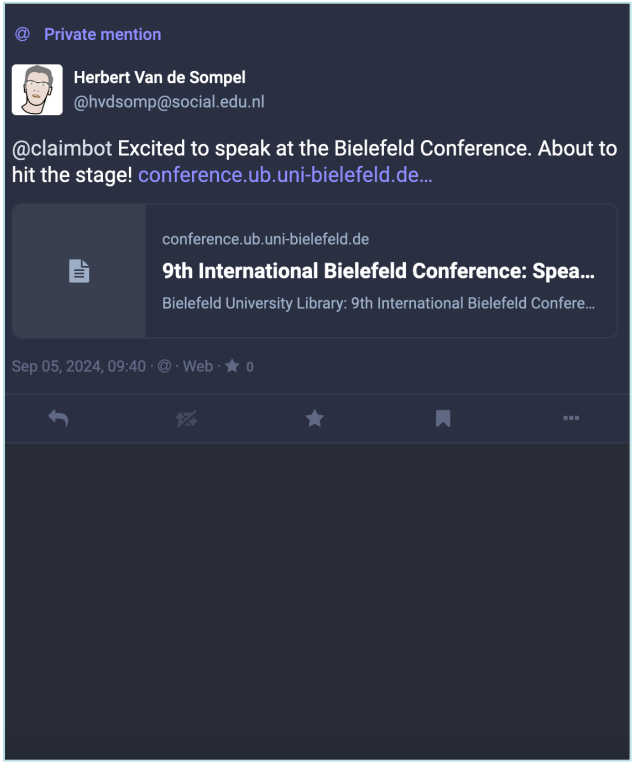


- Prior to the Conference, I express my excitement to be speaking
- After the talk, I share my slides

Note that I have sent these as private messages so as to not spam the SURF Mastodon instance. But in reality I would send these publicly so that any one can see my new contributions.

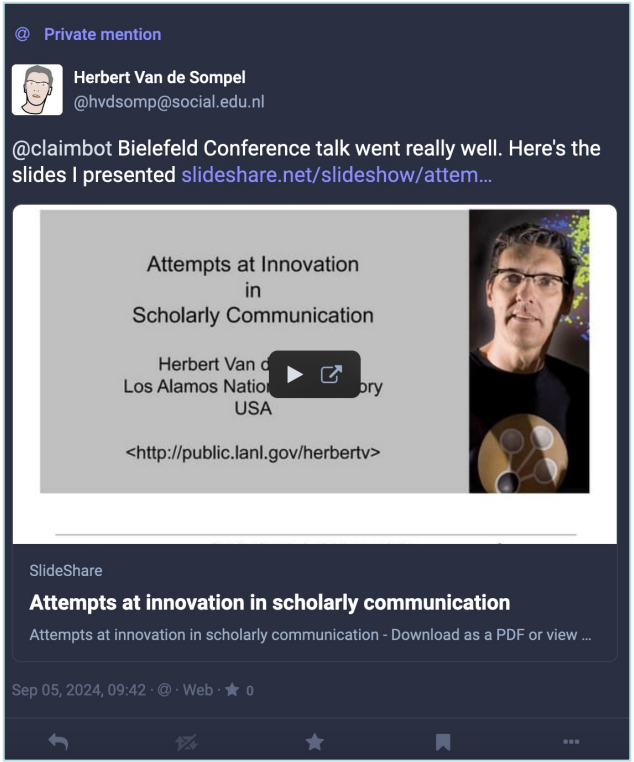
CC BY SA
UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

mycontributions.info – Announce Contribution 1, cc @claimbot



https://social.edu.nl/@hvdsomp/113083813675476591

mycontributions.info – Announce Contribution 2, cc @claimbot



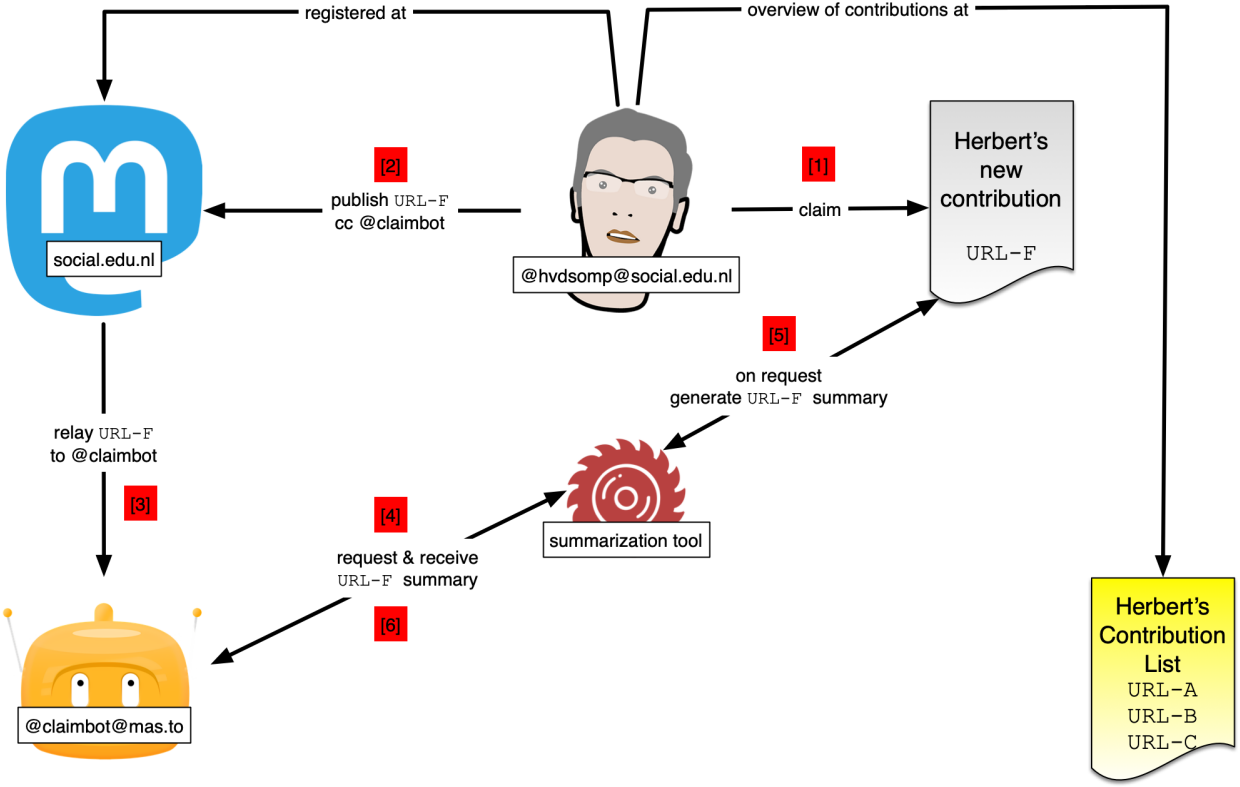
https://social.edu.nl/@hvdsomp/113083819150666111

The claimbot receives the post, checks whether it originates from someone registered with the SURF Mastodon instance, and if so, calls a service that automatically extracts pertinent metadata about the contribution.

That information extraction is not conducted by calling APIs as was the case in myresearch.institute but by obtaining information via the contribution page.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS



CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

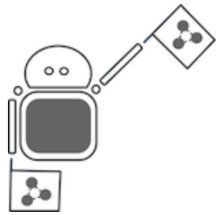
The communication between the claimbot and the extraction tool uses the notification approach described in the Event Notifications in Value-Adding Networks specification authored by Patrick Hochstenbach and colleagues.

The approach is itself based on Linked Data Notifications and ActivityStreams2 vocabulary and is, among others, also used in the COAR Notify project for communication between repositories and overlay peer-review services. But. the notification approach is increasingly attracting attention for other use cases such as requesting archiving of resources, requesting/announcing augmentation of metadata, registration of datasets in registries, etc.

Event Notifications in Value-Adding Networks

Event Notifications in Value-Adding Networks

Living Document, 6 September 2023



This version:
https://www.eventnotifications.net

Latest published version:
https://www.eventnotifications.net

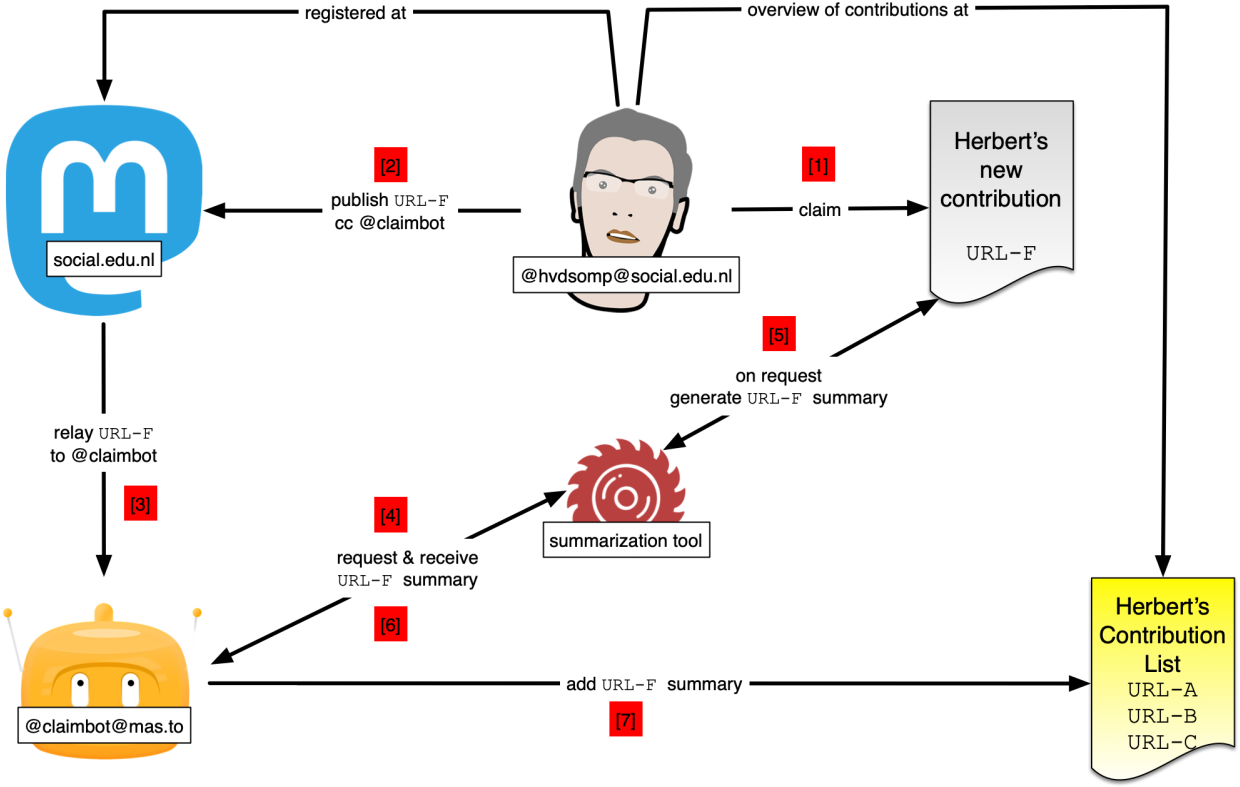
Previous Versions:
https://www.eventnotifications.net/0.1/

Feedback:
Inline In Spec

Editors:
Patrick Hochstenbach (Ghent University Library)
Miel Vander Sande (meemoo - Flemish Institute for Archives)
Ruben Dedecker (IDLab - Ghent University)
Paul Walk (Antleaf)
Martin Klein (Los Alamos National Laboratory)
Herbert Van de Sompel (IDLab - Ghent University)

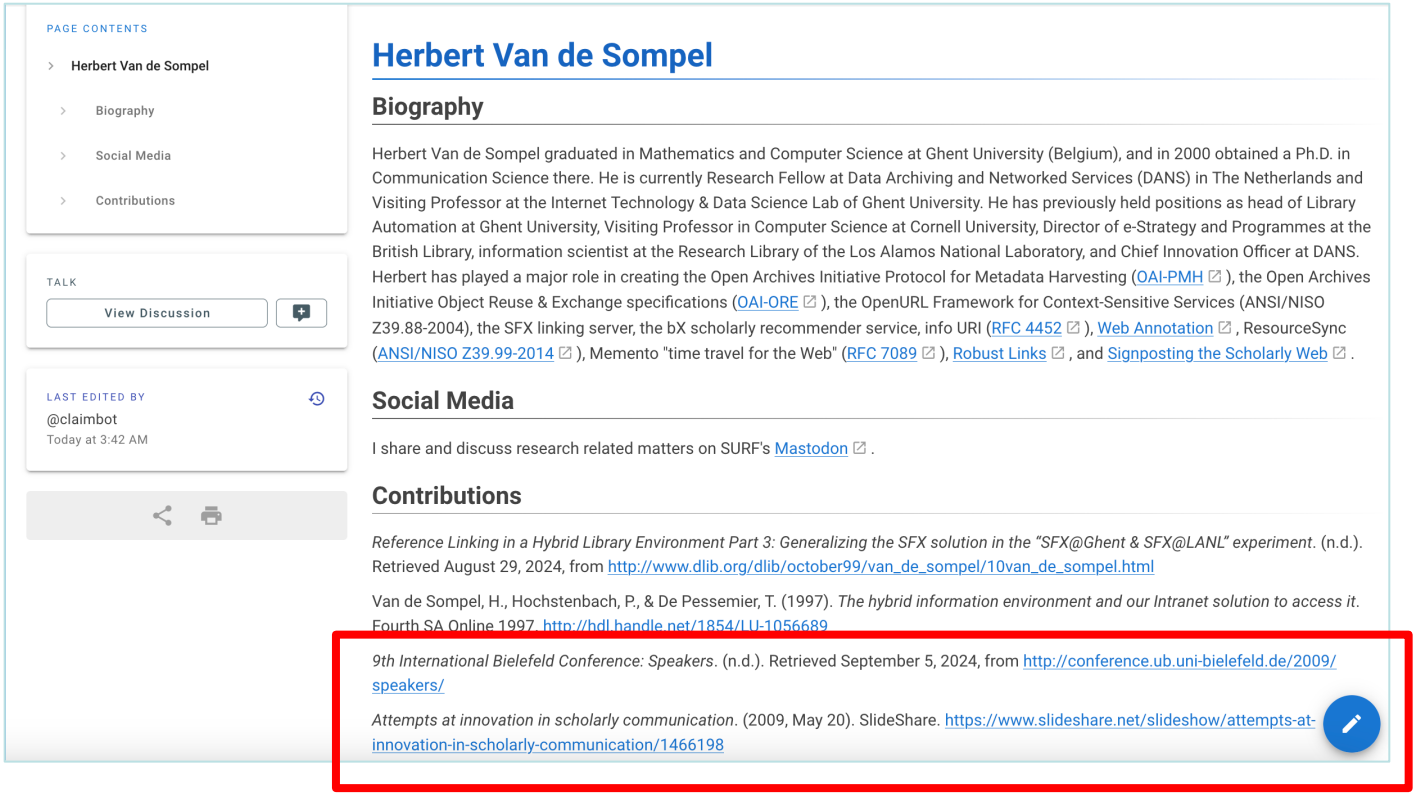
Hochstenbach, P., Vander Sande, M. Dedecker, R. et al. (2023) Event Notifications in Value-Adding Networks https://www.eventnotifications.net/

The claimbot adds an entry for these new contributions to my Contribution List.


registered at
overview of contributions at
social.edu.nl
[2]
publish URL-F
cc @claimbot
@hvdsomp@social.edu.nl
[1]
claim
Herbert's new contribution
URL-F
[5]
on request
generate URL-F summary
relay URL-F
to @claimbot
[3]
summarization tool
[4]
request & receive
URL-F summary
[6]
@claimbot@mas.to
add URL-F summary
[7]
Herbert's Contribution List
URL-A
URL-B
URL-C

As you can see the two new contributions were added to my Contribution List. The metadata extracted is pretty minimal in this case but there's at least an entry for the contributions.

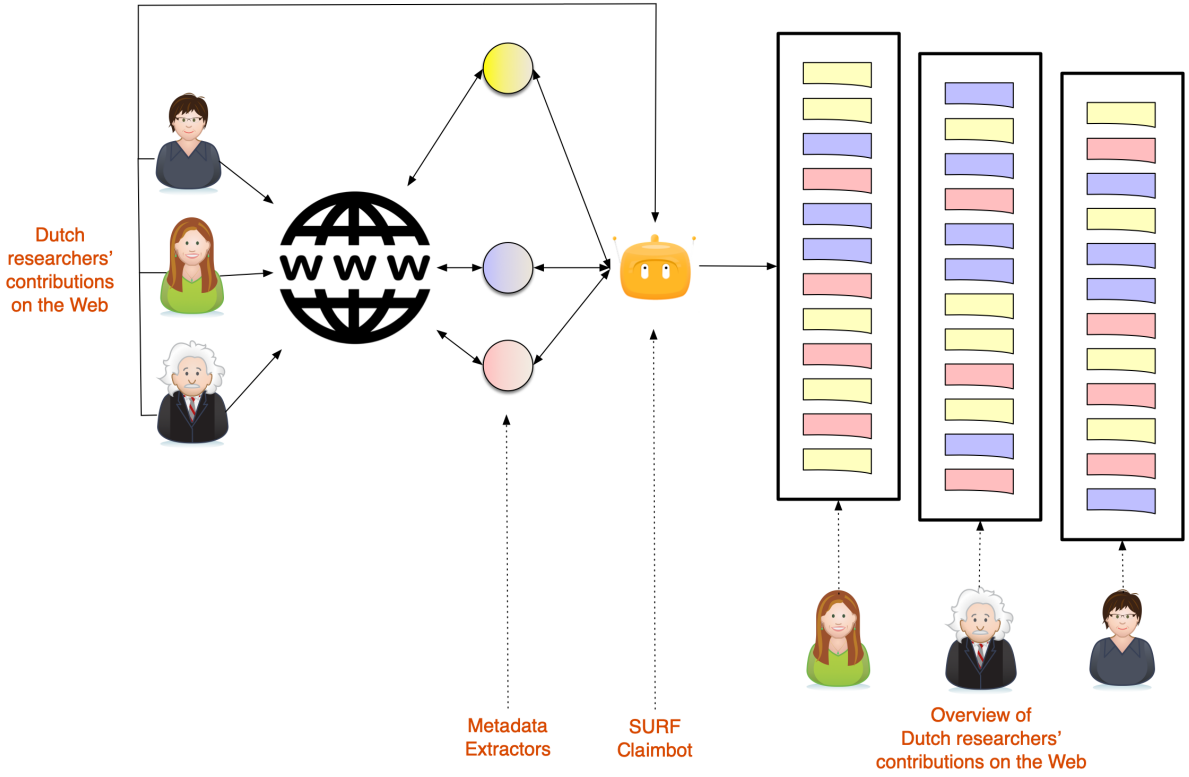
mycontributions.info – Updated Contribution List



https://wiki.mycontributions.info/researcher/orcid/0000-0002-0715-6126

The overall approach in this case still draws inspiration from the Personal Observatory work in that overviews of contributions by researchers that reside on the web are created.

But the action radius of this approach is not just a select number of portals in which researchers are active but rather the web at large. Granted, researchers can not just sit back and relax while their contributions are automatically detected. They need to be pro-active and announce them. But in return, their Contribution Lists are updated all the while they are also informally communicating about scholarship on mastodon.



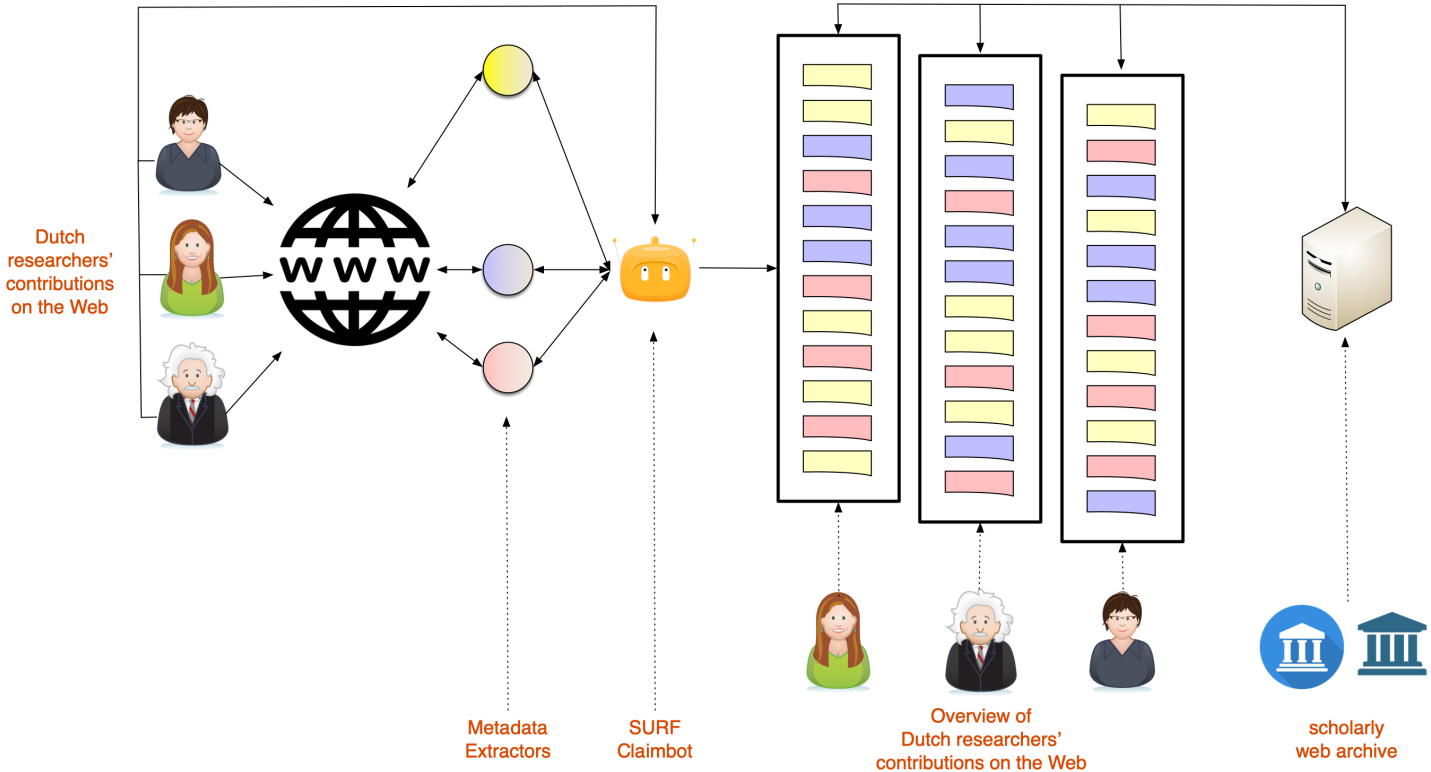
CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

The focus in this experiment is on novel research assessment approaches but it is obvious that Contribution Lists can also be used as seeds for archiving, just like the traces obtained by observing APIs in the myresearch.institute experiment.

The approach could be passive, i.e. the Contribution List just sitting there, waiting, until an archival bot comes by.

Or, much better, the approach could be pro-active, for example, by using event notifications to request resources to be archived. In which case, Contribution Lists can link to Mementos of the contributions, which is beneficial to the scholarly record as such but also augments the value of Contribution Lists as sources for evaluation, indicators, etc. Because even when the contributions are gone from the live web, their nature can be checked via archives.



CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

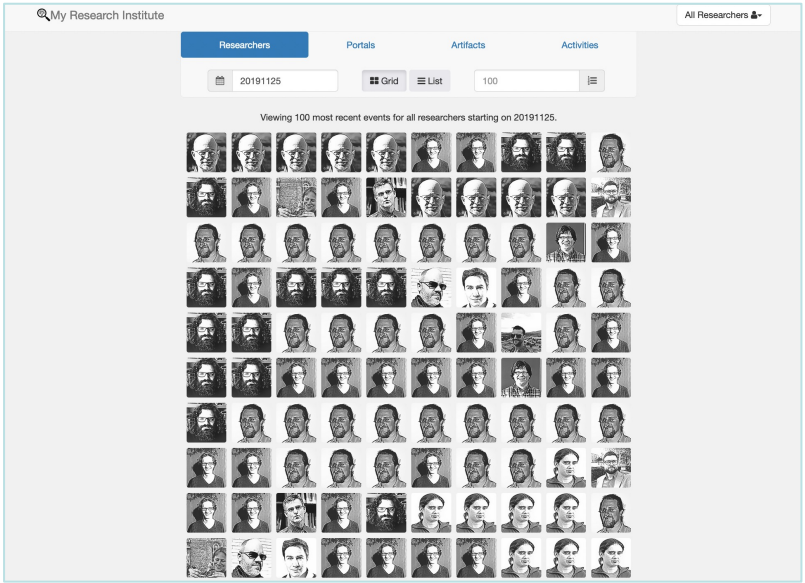
I have discussed two experiments, myresearch.institute and mycontributions.info that take a different perspective on archiving the web-based scholarly record. Not a systemic but rather an institutional and personal perspective, respectively.

Both approaches, if they would be pursued in reality instead of as a mere experiments, would lead to a better archival coverage of scholarly contributions; they would complement existing archival frameworks. And, as I've shown, in both approaches, incentives are in place for these parties to invest time and effort to optimize archival coverage of research contributions.

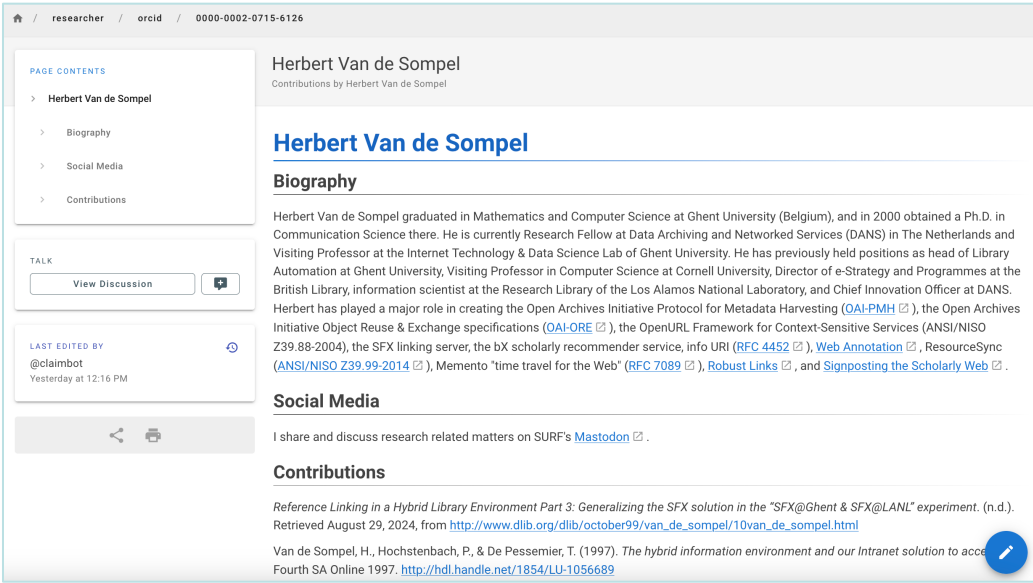
I very much understand that there are very significant challenges involved in turning these experiments into reality. Organizational, social, economical, legal, etc., etc. But, as one can when presenting a keynote, I am going to leave those to the reader.

CC BY SA

myresearch.institute



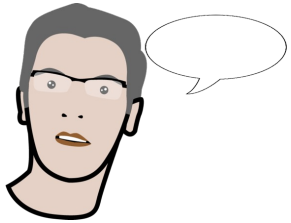
mycontributions.info



CC BY SA

UNIVERSITEIT GENT

Data Archiving and Networked Services
DANS



I want to loop back to the start of the presentation where I discussed challenges with archiving personal traces.

Remember: artifacts slipping through the maze, the lack of a resource that ties all archived artifacts of an individual together, problems discovering personal traces because they reside in various archives.

There's a lot, especially in the mycontributions.info approach that I find inspiring when thinking more generally about archiving the personal traces that individuals leave on the web.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

Archived Personal Traces

- **Some traces left by individuals get captured as a result of the collection development policy of archives, others slip between the maze** although they might have been important to the individual.

- **Discovery of archived personal traces** – No obvious way to determine in which archives an individual's traces can be found.

- **Inventory of archived personal traces** – No aggregate resource that connects these archived traces, that makes it obvious that the archived traces available in various archives were left by a same individual.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

I am going to characterize the approaches I discussed in terms of the various phases as identified by these Tips for Personal Web Archiving that are very much inspired by those for Personal Digital Archiving. And I am going to look into the Who does What in each phase.

These Tips are about getting your stuff off the web, onto your own machine, and onto redundant storage. That's not what I am after here because:

- I want the personal archive to be usable and I want it to complement other archives, which it can not do when it sits on my laptop;
- I don't want to and am not able to run my own archival infrastructure;
- I don't think I can dedicate the necessary attention to this because I am busy living my life and trying to protect myself from cyberscams.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

Personal Web Archiving Tips

	How	Who
Identify	locate your content on the web	individual
Select	select what's important to you	individual
Capture	save as ; export	individual
Organize	directories ; file names	individual
Persist	redundancy ; migration	individual

based on : Personal Archiving: Preserving your Digital Memories
https://www.digitalpreservation.gov/personalarchiving/documents/PA_websites.pdf

Let’s look at the mycontributions.info approach in this way too:

- Individuals identify and select the resources they deem relevant by tooting about them and creating Contribution Lists; the lists provides seeds for archival processes.

- Memory organizations take care of the last three phases and as such individuals do not need web archiving nor long-term preservation expertise.

- Using an appropriate interface between a Contribution List and a web archive, the List can be made to reflect in which archive contributions are archived. And, obviously, not all contributions would need to end up in the same archive.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

mycontributions.info

	How	Who
Identify	toot ; contribution list	individual
Select	toot ; contribution list	individual
Capture	web crawl/capture	organization
Organize	contribution list ; web archive replay	indiv ; org
Persist	redundancy ; migration	organization

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

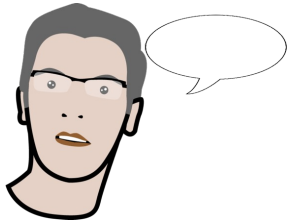
Contrast this with the myresearch.institute approach in which the researcher’s institution takes care of identification by choosing a set op portals to monitor but contributions slipped through the maze because one can’t monitor them all.

myresearch.institute

	How	Who
Identify	track researchers via portal APIs	institution
Select	--	institution
Capture	web crawl/capture	organization
Organize	event database ; web archive replay	inst ; org
Persist	redundancy ; migration	organization

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS



This leads us to the following perspective on Personal Web Archiving in which I reorganized the parameters of the previous tables a bit:

- The individual is in charge of Identification/Selection because the individual is keen to get specific traces archived.

- The memory organization in charge of Capture/Organize/Persist because that's its expertise and – in a best case scenario – its time horizon is longer than that of an individual.

A partnership, basically, between individuals and memory organizations for Personal Web Archiving.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

Personal Web Archiving as a Partnership?

	Individual	Memory Organization
Identify	nomination	
Select	nomination	
Capture		web crawl/capture
Organize		web archive replay
Persist		redundancy ; migration

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

I want to make some observations here, some technical, some organizational. First technical:

- There's something very very nice about the overviews of personal traces along with their corresponding Mementos as discussed in both myresearch.institute and mycontributions.info. They are an inventory documenting personal traces and their archival copies.

- Capture challenges (complex web pages, bots increasingly being blocked e.g. because of aggressive AI bots, authorization) may require desktop capture even though that's more demanding for the individual. Desktop capture undoubtedly has its place in web archiving but maybe the work involved may be prohibitive in the context of personal web archiving.

- There’s something really attractive about the nomination approach in mycontributions.info, in which the individuals are just in their normal flow on the web, and nomination totally integrates into that flow. It did so with toots but seamlessness can be achieved in other ways, e.g. Zotero personal collection & Robust Links add-on. It would be nice if a level of seamlessness could be achieved for both the web crawl/capture and desktop capture paradigms.

- Since we’re dealing with an environment with multiple archives interoperability regarding capture requests would be advantageous:
  - by-reference requests when dealing with web capture;
  - by-value requests when dealing with desktop capture;
  - responses to these requests in order to streamline management of Inventories.

Technical Considerations

- Inventory of resources and their corresponding Mementos.

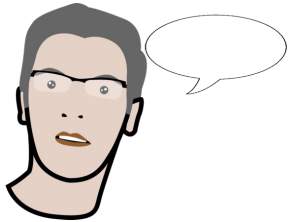
- Capture challenges (complex web resources, blocking of bots, authorization) may require a desktop capture approach rather than a web crawl/capture approach that is easier on the individual.

- Seamless, in-the-flow, nomination of resources to be archived for both capturing paradigms.

- Web archive interoperability for capture requests/responses:
  - By reference, to request a web archive to web crawl/capture resources;
  - By value, to transfer desktop-captured resources to a web archive;
  - Responses to these capture requests ~ management of Inventory.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS



This all leads us to a somehow enhanced characterization of the Personal web Archiving partnership with, as discussed, the necessary interoperability ingredients to make information flow between phases seamless.

Personal Web Archiving as a Partnership?

	Individual	Memory Organization
Identify	in-the-flow nomination	
Select	in-the-flow nomination	
Capture	desktop capture	web crawl/capture
Organize	artifact-memento inventory	web archive replay
Persist		redundancy ; migration

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services DANS

I also want to share some organizational observations:

- I would very much prefer a memory organization that has a long-term horizon to be in charge of the archive instead of a commercial company that might eventually enshittify its offering and ultimately prevent individuals from exporting their archive

- I have really been wondering why personal archiving is not a more common, integral component of the collection development policy of memory organizations. This has struck me big time when everyone was scrambling to get their stuff out of twitter, as I did, and all those unusable “archives” that now sit on hard disks. It seems to me that memory organizations can play a role in this personal web archiving realm and that there is a win/win involved:
  - Memory organizations increase their coverage and do so with explicit consent of individuals.
  - Individuals get a usable long-term archive.

Bottom Line: When it comes to personal web archiving individuals could use some help. Memory organizations come to mind as parties that could be the Rescue Heroes.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

Organizational Considerations

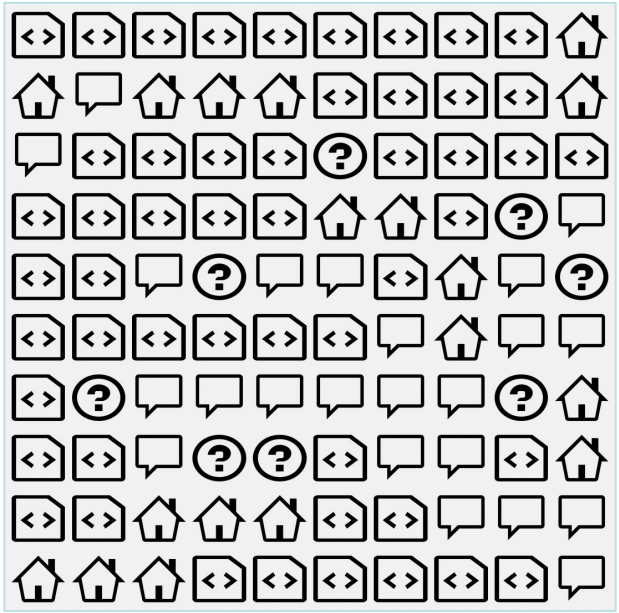
- Strong preference for memory organization with a long-term horizon; significant concerns regarding eventual enshittification of offerings by commercial players.

- Few memory organizations have Personal Web Archiving as an integral collection development approach despite what seems like a potential win/win:
  - memory organizations increase their coverage with explicit consent of individuals;
  - Individuals get a usable long-term archive.

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS

Former Farmer Chases Traces



Herbert Van de Sompel
@hvdsomp@w3c.social - https://hvdsomp.info

CC BY SA

UNIVERSITEIT GENT
Data Archiving and Networked Services
DANS