

# New digital research methods

---

**Peter M. Dahlgren**

HAS, April 19, 2017

I will show some:

- Interesting data sources
- Experimentation with methods (lots of R)
- Open science and open data

## DMI-TCAT

R

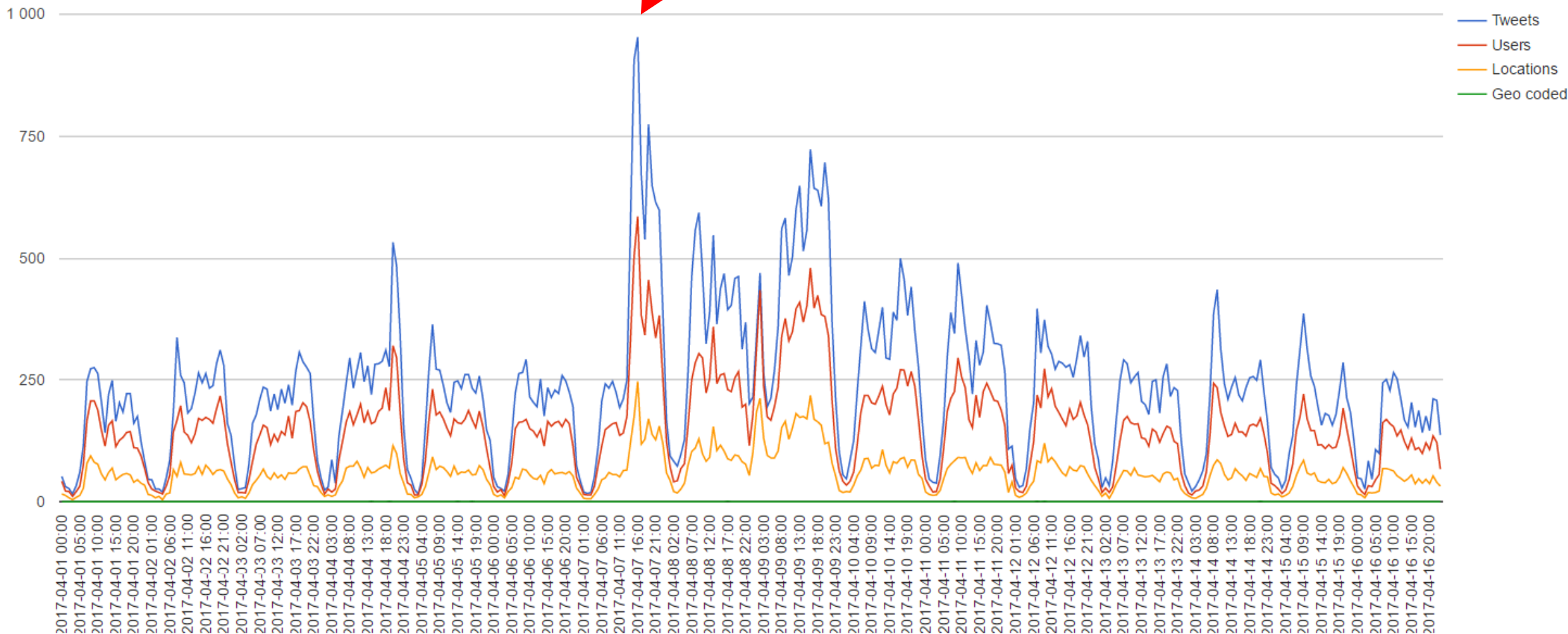
1. Trump tweets sentiment
2. ~~Drottninggatan tweets~~
3. Wikipedia breakout detection
4. GDELT (Television Explorer)
5. Microsoft Cognitive Services
6. statcheck

Open science

- Open Science Framework
- Notebooks
- Open data
- Shiny

Startdate: 2017-04-01  
Enddate: 2017-04-16  
Number of tweets: 91.462  
Number of distinct users: 14.853

#svpol



[Generate SVG](#)

Date and time are in BST (Europe/London).

Graph resolution ☐ days ☒ hours ☐ minutes [update graph](#)





[\[Home\]](#)

## Download

[CRAN](#)

## R Project

[About R](#)

[Logo](#)

[Contributors](#)

[What's New?](#)

[Reporting Bugs](#)

[Development Site](#)

[Conferences](#)

[Search](#)

## R Foundation

[Foundation](#)

[Board](#)

[Members](#)

[Donors](#)

[Donate](#)

## Help With R

[Getting Help](#)

# The R Project for Statistical Computing

## Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred [CRAN mirror](#).

If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

## News

- [R version 3.4.0 \(You Stupid Darkness\) prerelease versions](#) will appear starting Tuesday 2017-03-21. Final release is scheduled for Friday 2017-04-21.
- [R version 3.3.3 \(Another Canoe\)](#) has been released on Monday 2017-03-06.
- [useR! 2017](#) (July 4 - 7 in Brussels) has opened registration and more at <http://user2017.brussels/>
- Tomas Kalibera has joined the R core team.
- The R Foundation welcomes five new ordinary members: Jennifer Bryan, Dianne Cook, Julie Josse, Tomas Kalibera, and Balasubramanian Narasimhan.
- [The R Journal Volume 8/1](#) is available.
- The [useR! 2017](#) conference will take place in Brussels, July 4 - 7, 2017.
- [R version 3.2.5 \(Very, Very Secure Dishes\)](#) has been released on 2016-04-14. This is a rebadging of the quick-fix release 3.2.4-revised.
- [Notice XQuartz users \(Mac OS X\)](#) A security issue has been detected with the Sparkle update mechanism used by XQuartz. Avoid updating over insecure channels.
- The [R Logo](#) is available for download in high-resolution PNG or SVG formats.
- [useR! 2016](#), has taken place at Stanford University, CA, USA, June 27 - June 30, 2016.
- [The R Journal Volume 7/2](#) is available

[r-project.org](http://r-project.org)



[Home]

## Download

[CRAN](#)

## R Project

[About R](#)

[Logo](#)

[Contributors](#)

[What's New?](#)

[Reporting Bugs](#)

[Development Site](#)

[Conferences](#)

[Search](#)

## R Foundation

[Foundation](#)

[Board](#)

[Members](#)

[Donors](#)

[Donate](#)

## Help With R

[Getting Help](#)

# The R Project for Statistical Computing

The screenshot shows the RStudio website homepage. At the top is the RStudio logo and a navigation menu with links: [rstudio::conf](#), [Products](#), [Resources](#), [Pricing](#), [About Us](#), [Blogs](#), and a search icon. The main banner features the text "RStudio" in large white letters, followed by "Open source and enterprise-ready professional software for R". To the right of the banner are links: "Download RStudio", "Discover Shiny", "shinyapps.io Login", and "Discover RStudio Connect". Below the banner, there are three main visual elements: a screenshot of the RStudio IDE interface, a map of the United States with a "ZIP explorer" overlay, and a collection of logos for R packages: [markdown](#), [Shiny](#), [tidyr](#), [knitr](#), and [ggplot2](#).

• Notice XQuartz users (mac OS X) A security issue has been detected with the Sparkle update mechanism used by XQuartz. Avoid updating over insecure channels.

[r-project.org](https://r-project.org) + [rstudio.com](https://rstudio.com)

File for download in high-resolution PNG or SVG formats.

then place at Stanford University, CA, USA, June 27 - June 30, 2016.

• The R Journal Volume 7/2 is available





## Watching Our World Unfold

Visualization credit GDELT Project.

[gdeltproject.org](http://gdeltproject.org)

### A Global Database of Society

Supported by [Google Jigsaw](#), the GDELT Project monitors the world's broadcast, print, and web news



Watching

Computing

Downloading

Blogging

Star

Solutions

About



# Watching The Entire World

**GDELT monitors the world's news media from nearly every corner of every country in print, broadcast, and web formats, in over 100 languages, every moment of every day.**

massive computing power to see the world through others' eyes  
or local populations with the information and insights they

[gdeltproject.org](http://gdeltproject.org)

logging

Starting

# Watching Our World Unfold

## Using GDELT

The GDELT Project is an open platform for research and analysis of global society and thus all datasets released by the GDELT Project are available for unlimited and unrestricted use for any academic, commercial, or governmental use of any kind without fee.

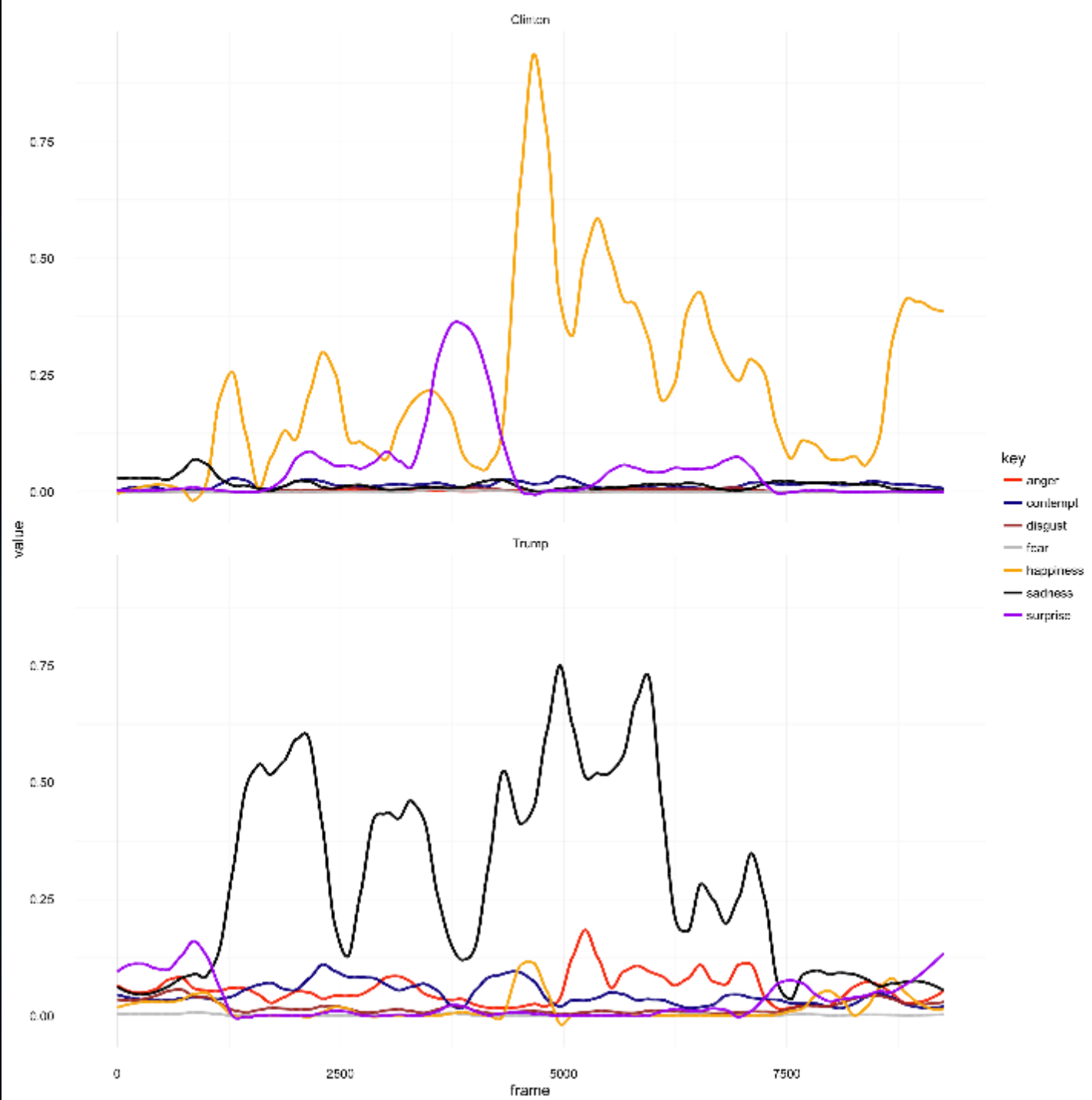
monitors the world's broadcast, print, and web news

# Discover the APIs

It's building time. Integrate these vision, speech, language, knowledge, and search APIs into your app, on the platform of your choice.

[Get started for free](#)

## Vision





# HFST-SweNER – A New NER Resource for Swedish

Dimitrios Kokkinakis<sup>§</sup>, Jyrki Niemi<sup>±</sup>, Sam Hardwick<sup>±</sup>,  
Krister Lindén<sup>±</sup>, Lars Borin<sup>§</sup>

<sup>§</sup>Språkbanken, Department of Swedish  
University of Gothenburg, Sweden  
E-mail: first.last@svenska.gu.se

<sup>±</sup>Department of Modern Languages  
University of Helsinki, Finland  
E-mail: first.last@helsinki.fi

## Abstract

Named entity recognition (NER) is a knowledge-intensive information extraction task that is used for recognizing textual mentions of entities that belong to a predefined set of categories such as locations, organizations and time expressions. NER is a challenging

difficult,  
language  
of social  
environn  
(HFST)  
generic  
(gazettee

Keywor

Named  
language

(a) `<s id=jc04-041> Han kom till <ENAMEX TYPE="LOC" SBT="PPL">Stockholm</ENAMEX> <TIMEX TYPE="TME" SBT="DAT">1885</TIMEX> , fick en organisatorisk bas i <ENAMEX TYPE="ORG" SBT="PLT">Socialdemokratiska klubben</ENAMEX> ( senare <ENAMEX TYPE="ORG" SBT="PLT">Socialdemokratiska förbundet</ENAMEX> ) , kunde starta <ENAMEX TYPE="PRS" SBT="HUM">Social-Demokraten</ENAMEX> och fick medarbetare som <ENAMEX TYPE="PRS" SBT="HUM">Axel Danielsson</ENAMEX> , <ENAMEX TYPE="PRS" SBT="HUM">Fredrik Sterky</ENAMEX> och , <TIMEX TYPE="TME" SBT="DAT">året därpå</TIMEX> , <ENAMEX TYPE="PRS" SBT="HUM">Hjalmar Branting</ENAMEX> . <s>`

(b) `<s id=jc04-041> Han kom till <name type=place> Stockholm </name> 1885 , fick en organisatorisk bas i <name type=inst> Socialdemokratiska klubben </name> ( senare <name type=inst> Socialdemokratiska förbundet </name> ) , kunde starta <name type=inst> Social-Demokraten </name> och fick medarbetare som <name type=person> Axel Danielsson </name> , <name type=person> Fredrik Sterky </name> och , året därpå , <name type=person> Hjalmar Branting </name> . <s>`

Figure 1. Example sentence from the gold standard used for evaluation (see Section 7.3), annotated by the HFST-SweNER (a) and the original manual annotation (b). Note that in (a) the entity *Social-Demokraten*, is erroneously annotated as a person (PRS), which according to the manual annotation in the gold standard it should have been an *inst* (ORGanization).

# HFST-SweNER – A New NER Resource for Swedish

Dimitrios Kokkinakis<sup>§</sup>, Jyrki Niemi<sup>±</sup>, Sam Hardwick<sup>±</sup>,  
Krister Lindén<sup>±</sup>, Lars Borin<sup>§</sup>

<sup>§</sup>Språkbanken, Department of Swedish  
University of Gothenburg, Sweden  
E-mail: first.last@svenska.gu.se

<sup>±</sup>Department of Modern Languages  
University of Helsinki, Finland  
E-mail: first.last@helsinki.fi

## Abstract

Named entity recognition (NER) is a knowledge-intensive information extraction task that is used for recognizing textual mentions of entities that belong to a predefined set of categories, such as locations, organizations and time expressions. NER is a challenging

Collect

Clean

Extract

Analyze

Keyword: TYPE="PRS" SBT="HUM">Hjalmar Branting</ENAMEX> . <s>

Named language: (b) <s id=jc04-041> Han kom till <name type=place> Stockholm </name> 1885 , fick en organisatorisk bas i <name type=inst> Socialdemokratiska klubben </name> ( senare <name type=inst> Socialdemokratiska förbundet </name> ) , kunde starta <name type=inst> Social-Demokraten </name> och fick medarbetare som <name type=person> Axel Danielsson </name> , <name type=person> Fredrik Sterky </name> och , året därpå , <name type=person> Hjalmar Branting </name> . <s>

Figure 1. Example sentence from the gold standard used for evaluation (see Section 7.3), annotated by the HFST-SweNER (a) and the original manual annotation (b). Note that in (a) the entity *Social-Demokraten*, is erroneously annotated as a person (PRS), which according to the manual annotation in the gold standard it should have been an *inst* (ORGanization).



## GROUP DETAILS

[Charter](#)[Group Participants](#)[Wiki](#)[Issue Tracker](#)[Patent Policy Status](#)

## RECOMMENDATIONS AND NOTES

## Recommendations

[Web Annotation Data](#)[Model](#)[Web Annotation](#)[Vocabulary](#)[Web Annotation](#)[Protocol](#)

## Working Group Notes

[Embedding Web](#)[Annotations in HTML](#)[Selectors and States](#)

*Warning: this Working Group is now closed. It has published its [Recommendations and Working Group Notes](#) on the 23rd of February 2017. For further discussions on the evolution of Web Annotation technologies at W3C, possible errata, etc, please join the [W3C Open Annotation Community Group](#).*

## WEB ANNOTATION WORKING GROUP

The W3C Web Annotation Working Group is part of the [Publishing@W3C](#), but the scope is more than just publishing; Web annotation is for all content on all devices.

A detailed list of this group's [publications](#) is available in the working group's specification repository.

The W3C Web Annotation Working Group is [chartered](#) to develop a set of specifications for an interoperable, sharable, distributed Web Annotation architecture.

Over the course of its charter, this group has created the following Recommendations:

[Web Annotation Data Model](#)

Describes the underlying **Annotation Abstract Data Model** as well as a JSON-LD serialization

[Web Annotation Vocabulary](#)

The **Vocabulary** which underpins the Web Annotation Data Model

[Web Annotation Protocol](#)

The **HTTP API** for publishing, syndicating, and distributing Web Annotations

## What are Web Annotations?

Traditional annotations are marginalia, errata, and highlights in printed books, maps, picture, and other physical media. Web annotations are an attempt to recreate and extend that functionality as a new layer of interactivity and linking on top of the Web. It will allow anyone to annotate anything anywhere, be it a web page, an ebook, a video, an image, an audio stream, or data in raw or visualized form. Web annotations can be linked, shared between services, tracked back to their origins, searched and discovered, and stored wherever the author wishes: the vision is for



Medieforskarna.se samlar bloggar från svenska forskare om medier, journalistik och kommunikation.

Prenumerera på RSS-flödet

Följ @medieforskarna

### Ämnen

- Allmänt
- Debatt
- Internationellt
- Internet och sociala medier
- Journalistik
- Konferenser
- Kriser
- Skogsbranden utanför Sala

[medieforskarna.se](https://medieforskarna.se)

Pelle Snickars 1 dag sedan

## MIT Press manuscript slowly emerging

Within our research project around Spotify we are currently together writing and building a book manuscript to be delivered to MIT Press before summer. Things are proceeding, albeit somewhat slowly. I am presently in charge of Chapter 2, which [...]

Bengt Johansson 5 dagar sedan

## Baad PR

United Airlines är i blåsväder. Att släpa blodiga resenärer nerför passagerargången som inte vill gå av ett överbokat plan är inte att rekommendera. I alla fall inte om man vill ha bra PR och [...]

Margareta Rönnberg 1 vecka sedan

## Är skräckhistorier av typen vandringsägner som VARUHuset verkligen något för 6-åringar?

I torsdags var jag dygnet-runt-föräldra-stand-in för mina barnbarn och skulle i vanlig ordning inför nattningen läsa för 6-åringen och 8-åringen. Jag hade tagit med Dahls och Johanna [...]

Margareta Rönnberg 2 dagar sedan

## Beror verkligen ökningen av barns närsynthet på det myckna "skärmtittandet"? På bokläsning?

Till listan av decenniets alla falskheter sällar sig nu också "falsk närsynthet"... Det senaste halvåret har i pressen och i tv:s nyhetsprogram ännu ett medielarm dykt upp: barns ögon sägs [...]

Peter M. Dahlgren 6 dagar sedan

## "Det är algoritmerna" – en förklaring som betyder "jag har ingen förklaring"

Stiftelsen för internetinfrastruktur har släppt en ny rapport vid namn Algoritmer: Så påverkar de din vardag, skriven av Jutta Haider och Olof Sundin. Här tänkte jag redogöra för min kritik mot [...]

Ulrika Hedman 1 vecka sedan

## #drottninggatan #openstockholm

Det delades en hel del rykten, spekulationer och överdrifter på Twitter igår. Det var skottlossning här, utpekade gärningsmän där, falska citat från ministrar på ett håll, ett malande om ett [...]

Bengt Johansson 3 dagar sedan

## Raserar kriser samhällsförtroendet? KRISAMS drar igång

Kerlin Björn 15.00 Normal 0 21 false false false SV JA X-NONE [...]

Mediespanarna poddradio 5 dagar sedan

## 276. Terrorismens gräns och journalistens mall

Efter attentatet på Drottninggatan i Stockholm diskuterar Mediespanarna hur journalistiska genrer försvårar rapportering och vad efterföljande debatter egentligen handlar om.

Bengt Johansson 1 vecka sedan

## Where have all the flowers gone?

Blommorna som partisymboler är på väg ut. I förra veckan lanserade kristdemokraterna sin nya logotype och därmed var vitsippan borttagen från den svenska partirymden. Istället innehåller loggan bara [...]

Ulrika Hedman 1 vecka sedan

## Några tankar om våra nyhetsmedier så här dagen

# Selective exposure

Private


Make Public



0

Contributors: [Peter M. Dahlgren](#)

Date created: 2016-12-16 07:59 PM | Last Updated: 2016-12-16 07:59 PM

Category: Project 

Description: No description

License: No license

## Wiki



No wiki content

## Files



Click on a storage provider or drag and drop to upload

 Filter

Name ^ ▾

Modified ^ ▾

 Selective exposure-  OSF Storage

## Citation

osf.io/mzj3f ▾

## Components

Add Component


Link Projects

No components have been added to this project.

## Tags

add a tag

## Recent Activity

 [Peter M. Dahlgren](#) created [Selective exposure](#)

2016-12-16 07:59 PM

# Liberals and Conservatives are Similarly Motivated to Avoid Exposure to One Another's Opinions



Jeremy Frimer, Linda Skitka, Matt Motyl

Added on: April 17, 2017 | Last edited: April 17, 2017

[Download preprint](#)

Downloads: 13

## Abstract

Ideologically committed people are similarly motivated to avoid ideologically crosscutting information. Although some previous research has found that political conservatives may be more prone to selective exposure than liberals are, we find similar selective exposure motives on the political left and right across a variety of issues. The majority ...

[See more](#)

## Article DOI

10.1016/j.jesp.2017.04.003

## License

CC0 1.0 Universal ▶

## Disciplines

Social and Behavioral Sciences Psychology Social Psychology

## Tags

confirmation bias ideological symmetry ideology motivated reasoning  
political psychology selective exposure social psychology

## Liberals and Conservatives are Similarly Motivated to Avoid Exposure to One Another's Opinions

Jeremy A. Frimer, University of Winnipeg

Linda J. Skitka, & Matt Motyl, University of Illinois, Chicago

May be cited as:

(2017). Liberals and conservatives are similarly motivated  
her's opinions. *Journal of Experimental Social Psychology*,

[osf.io/preprints](https://osf.io/preprints)



# Tomorrow's Research Today

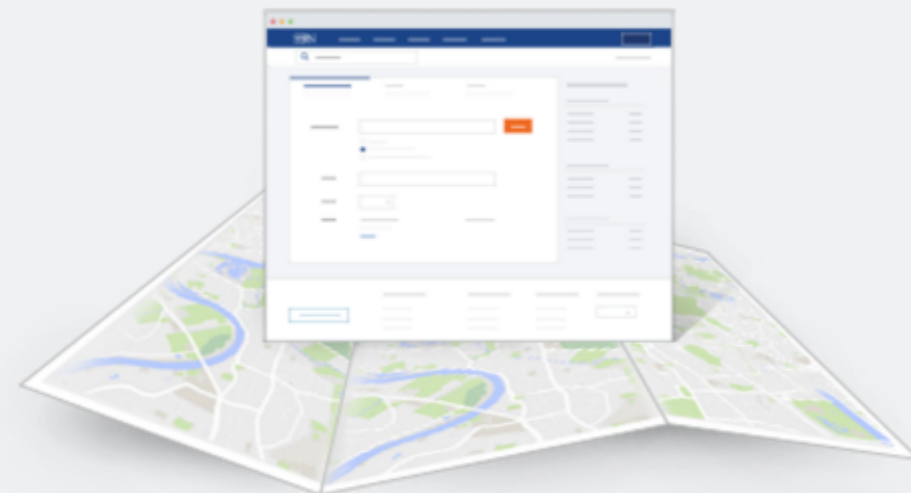
Welcome to the new SSRN Home Page. Check out the 'New Look & Feel' link above the map icon for more information.

SSRN's eLibrary provides 728,359 research papers from 336,368 researchers across 30 disciplines.



Advanced Search

[New Look & Feel of Site >](#)



## Research Networks

Accounting	<a href="#">ARN</a>	Hebrew	<a href="#">HRN</a>	Law	<a href="#">LSN</a>
Anthropology & Archaeology	<a href="#">AARN</a>	Humanities: Classics	<a href="#">CRN</a>	Leadership	<a href="#">LRN</a>
	<a href="#">BioRN</a>	Humanities: Literature	<a href="#">LIT</a>	Management	<a href="#">MRN</a>
	<a href="#">CSN</a>	Humanities: Music & Composition	<a href="#">MRCN</a>	Marketing	<a href="#">MKT</a>

## Quick Links



Research Paper Series



Conference Papers



# Open Data



Download and use open data that I have collected. I only require you to cite it if you decide to publish an analysis from it.

## 1. One million news articles with Facebook and Twitter shares

**Download data:** <http://dx.doi.org/10.5878/002853>

This dataset contains 953,069 news articles (title and URL with shares on Facebook and Twitter) from the 12 largest Swedish news sites during 20 months, from February 2014 to October 2015. Format: CSV, TSV, SQL.

### Cita in APA

Dahlgren, P. M. (2017). Swedish online news articles 2014-2015. *Swedish National Data Service*.  
<http://dx.doi.org/10.5878/002853>

[peterdahlgren.com/opendata](http://peterdahlgren.com/opendata)

## Easy download of articles from home

If you work from home, you might have noticed the difficulty downloading articles. At least if you don't use the University VPN.

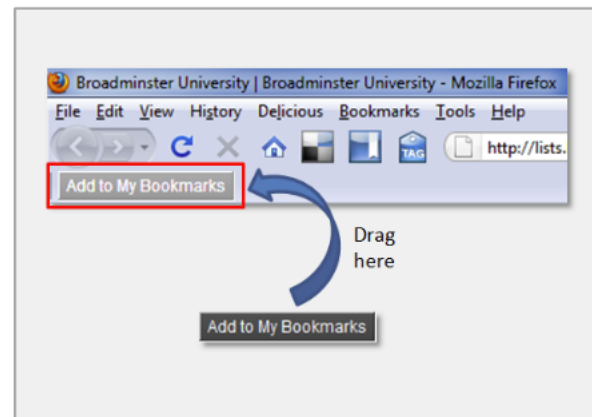
With this little bookmarklet, you can easily download articles you find on the web, without having to search for the article via the University of Gothenburg Library.

### How to Install

Just drag the link below and drop it onto your bookmarks in your web browser. You can put it anywhere.

Ezproxy

← Drag and drop to your bookmarks



[peterdahlgren.com/ezproxy](http://peterdahlgren.com/ezproxy)

For any interesting article you want to download, just press your



Download this presentation, code and data at

[github.com/peterdalle/has](https://github.com/peterdalle/has)



Myndigheten för  
samhällsskydd  
och beredskap

FORSKNING/STUDIE

## **Metoder och verktyg för att analysera opinioner och händelser i sociala medier**