

Developing a rule-based method for identifying researchers on Twitter: The case of vaccine discussions

SND-ID: snd1117-1. **Version:** 1.0. **DOI:** <https://doi.org/10.5878/akmc-va16>

Download data

class.py (7.62 KB)

Citation

Ekström, B. (2019) Developing a rule-based method for identifying researchers on Twitter: The case of vaccine discussions (Version 1.0) [Data set]. University of Borås. Available at: <https://doi.org/10.5878/akmc-va16>

Creator/Principal investigator(s)

Björn Ekström

Research principal

[University of Borås](#) - Akademin för bibliotek, information, pedagogik och IT

Principal's reference number

FO2017/23

Description

This study seeks to develop a method for identifying the occurrences and proportions of researchers, media and other professionals active in Twitter discussions. As a case example, an anonymised dataset from Twitter vaccine discussions is used. The study proposes a method of using keywords as strings within lists to identify classes from user biographies. This provides a way to apply multiple classification principles to a set of Twitter biographies using semantic rules through the Python programming language. The script used for the study is here deposited.

Method development for Twitter biography classification concerning occurrences of academics, academically related groups and individuals, media, other groups and members of the general public. Written in the Python programming language.

Language

[English](#)

Unit of analysis

[Group](#)

[Individual](#)

[Organization/Institution](#)

[Other](#)

Population

Twitter users

Time Method

[Other](#)

Sampling procedure

[Other](#)

Time period(s) investigated

2018-06-01 – 2019-10-31

Data format / data structure

[Software](#)

Responsible department/unit

Akademien för bibliotek, information, pedagogik och IT

Funding

- Funding agency: Horizon 2020
- Funding agency's reference number: 770531

Research area

[Language technology \(computational linguistics\)](#) (Standard för svensk indelning av forskningsämnen 2011)

[Social sciences](#) (Standard för svensk indelning av forskningsämnen 2011)

[Information studies](#) (Standard för svensk indelning av forskningsämnen 2011)

[Higher and further education](#) (CESSDA Topic Classification)

[Information society](#) (CESSDA Topic Classification)

[Language and linguistics](#) (CESSDA Topic Classification)

Keywords

[Classification](#), [Social media](#)

Publications

Ekström, B. (2019). Developing a rule-based method for identifying researchers on Twitter: The case of vaccine discussions. Poster abstract accepted to ISSI, 17th International Society of Scientometrics and Informetrics Conference, Rome, 2-5 September.

Accessibility level

Access to data through SND

Data are freely accessible

Use of data

[Things to consider when using data shared through SND](#)

License

[CC BY 4.0](#)

Versions

Version 1.0. 2019-08-23

CLARIN Virtual Collection Registry

[Add to collection](#)

A virtual collection is connected to a specific research purpose and contains links to data resources from various digital archives. It is easy to create, access, and cite the collection.

Read more about virtual collections on the [CLARIN website](#).

Download metadata

[DataCite](#)

[DDI 2.5](#)

[DDI 3.3](#)

[DCAT-AP-SE 2.0](#)

[JSON-LD](#)

[PDF](#)

[Citation \(CSL\)](#)

[File overview \(CSV\)](#)

Published: 2019-08-23

Last updated: 2019-09-20