Beyond twitter

Exploring `bluesky.social` for digital disease detection and prototyping a data extraction pipeline for ILI surveillance

Heiner Atze, MSc, PhD

Hasselt University

2024-04-24

Outline I

- Introduction
  - short history of bluesky [1]

---

[1]Signorini (2011)

Introduction

short history of bluesky

- What is special ? decentralized Balduf et al. (2024)
- Where does it come from … twitter

some figures and stats

Brief literature review of articles using `bluesky`

- Who is on there ?
- Polarisation and stuff
- any epidemiological research ?

Exploration of bluesky data

bluesky API

- accessibility
- terminology
- metadata
- availabilty of data
  - for now messages are returned exhaustively and deterministically

(Un)official statistics

Some research

Methodology

- focusing on French (time zone problem)
- extraction using list of keywords
  - grippe
  - rhume
  - fievre
  - courbature
-

Balduf, Leonhard, Saidu Sokoto, Onur Ascigil, Gareth Tyson, Björn Scheuermann, Maciej Korczyński, Ignacio Castro, and Michał Król. 2024. "Looking at the Blue Skies of Bluesky." In *Proceedings of the 2024 ACM on Internet Measurement Conference*, 76–91.

Signorini, Alberto Maria AND Polgreen, Alessio AND Segre. 2011. "The Use of Twitter to Track Levels of Disease Activity and Public Concern in the u.s. During the Influenza a H1N1 Pandemic." *PLOS ONE* 6 (5): 1–10. https://doi.org/10.1371/journal.pone.0019467.