

Using large language models for assessing the socioeconomic impacts of droughts

Mariana Madruga de Brito

Department of Urban and Environmental Sociology



Online interaction generates big unstructured text data

Volume

2.45 billion posts shared a day in Facebook

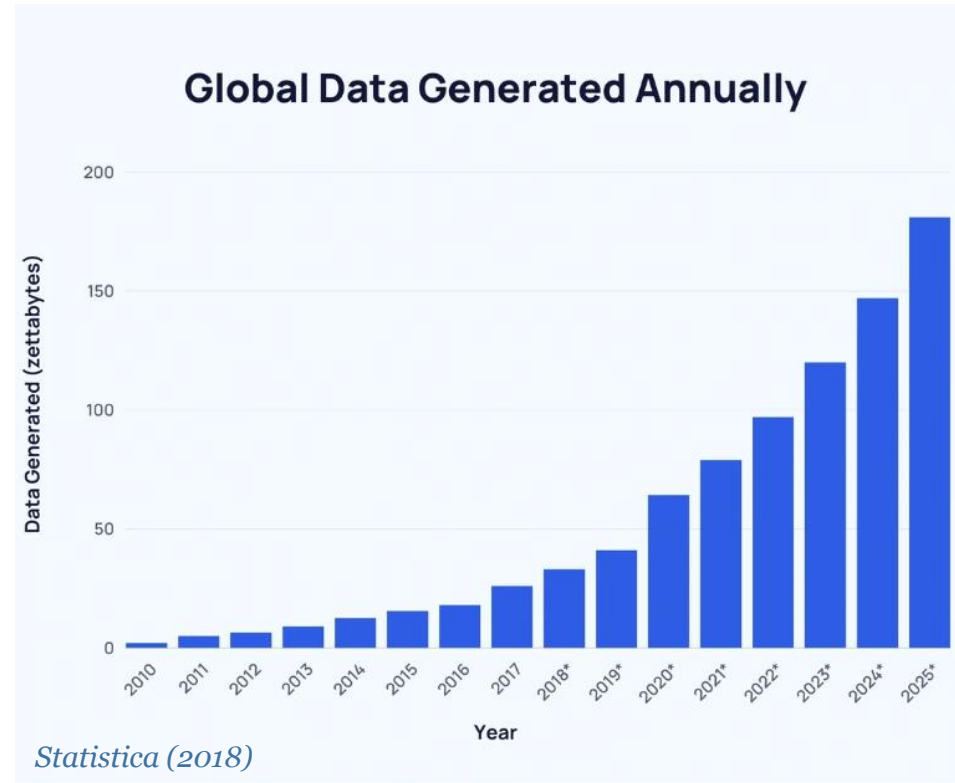
Velocity

2 Wikipedia revisions per second

Variety

X, scientific literature, news

Making sense of „messy“ data



Data collection

Newspaper aggregator database: *Wiso.de*

43,000 articles
2000 to 2022



Canva.com

Annotated data

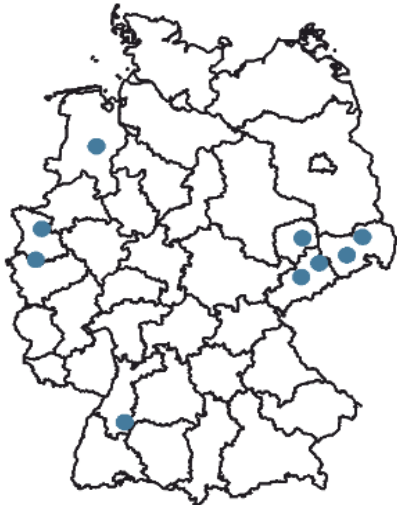
	D	E	F
1	Text	Impact 1	Impact 2
2	Die VG Altenahr sucht einen externen Dienstleister, der für die auf „fünf bis sieben Jahre“ geschätzte Dauer des Wiederaufbaus Unterstüt	0	1
3	“Wir sind in vielen Punkten mit unserem Latein am Ende,” sagte der Bürgermeister von Bad Neuenahr-Ahrweiler, Guido Orthen, bei der jü	0	1
4	Die Verbandsgemeinde Altenahr (VG) sucht europaweit nach einem „umfassend tätigen Dienstleister“, der die Verwaltung „bei der Realis	0	1
5	Die Diskussion um die Rolle der Fluthelfer Markus Wipperfürth und Wilhelm Hartmann geht weiter. Horst Gies, Erster Beigeordneter im Kr	0	1
6	Die Polizeiwache Ahrweiler hat ihren alten Standort in Ahrweiler wieder in Betrieb genommen. Nach der Flut hatte die Polizei zur Bundesal	0	1
7	Lösungen für den eklatanten Personalmangel in den Verwaltungen nach der Flutkatastrophe im Ahrtal sieht Jochen Nagel, Personalratsvor	0	1
8	Der für den Ahrtal-Einsatz in Rech freigestellte Berufssoldat Benjamin Vrijdaghs (43 Jahre, Stabsfeldwebel, parteilos) wird neuer ehrenam	1	1
9	Jürgen Saeß (SPD) ist der neue Ortsvorsteher von Heimersheim. Der 66-jährige Soldat a.D. war der einzige Bewerber bei der Direktwahl un	1	1
10	Der Ortsbürgermeister von Rech, Dominik Gieler, kandidiert für das Amt des Verbandsgemeindebürgermeisters der Verbandsgemeinde Alt	1	1
11	Der Stadtrat hatte in seiner Sitzung am 31. Januar beschlossen, dass Betroffene, deren Wohneigentum durch die Flut bis zum 31. Dezembe	1	1
12	Der Kreistag des Kreis Ahrweiler will aufgrund der Mehrbelastung durch die Flut einen hauptamtlichen Beigeordneten der Kreisverwaltung	0	1
13	Für die betroffenen Winzer und Landwirte ist also auch ein Jahr nach der Flut noch vieles im Unklaren, berichten Hubert Pauly, Präsidenten	0	1
14	Bislang wurden Landwirten und Winzern Entschädigungen für den Ernteausfall und die Übernahme von Entsorgungskosten in Höhe von ru	1	1
15	Die Auszahlung der Spendengelder in Millionenhöhe an von der Flut betroffene Winzer im Ahrtal war zunächst daran gescheitert, dass gen	1	1
16	Landwirte und Winzer können noch bis Freitag, 5. August, ihr Interesse zur Teilnahme an einer Förderung von Agrarumwelt- und Klimamaß	0	1
17	Das Bodenordnungsverfahren für den Wiederaufbau von Rebflächen in Dernau, Rech und Mayschoß wurde vorzeitig freigegeben, teilte La	0	0
18	Der Hauptgeschäftsführer des Gesamtverbands der Deutschen Versicherungswirtschaft (GDV), Jörg Asmussen, fordert klare Bauverbote in	1	0
19	Der Ortsbeirat von Heimersheim hat sich in seiner Sitzung gegen einen Neubau der Kita St. Maritus am Wiesweg ausgesprochen und wic	1	0
20	Der Dernauer Bürgermeister Alfred Sebastian zieht nach fast einem Jahr Wiederaufbau eine ernüchternde Bilanz: Es hake an allen Ecken u	0	0
21	Das Planungsbüro Stadt-Land-Plus hat sein Konzept für den Wiederaufbau nach der Flutkatastrophe für Mayschoß im Rahmen einer Einw	0	0
22	Interview mit Guido Orthen, BürgermeisterDas “vorläufig gesicherte Überschwemmungsgebiet” umfasst etwa 18 Prozent der Siedlungsflä	1	0
23	Fernsehbericht von Report Mainz “Angst nach der Flut - Neue Fehler beim Wiederaufbau?“Dr. Thomas Roggenkamp Hydrologe von der Uni	1	0
24	Anke Hupperich, Ortsvorsteherin von Kreuzberg kritisiert schleppende Genehmigung eines vom Gemeinderat im März 2022 beschlossenen	0	0
25	In einer öffentlichen Sitzung des Altenahrer Wiederaufbauausschusses präsentierte Klaus Zimmermann, Inhaber des Institut für Immission	1	0

Location extraction

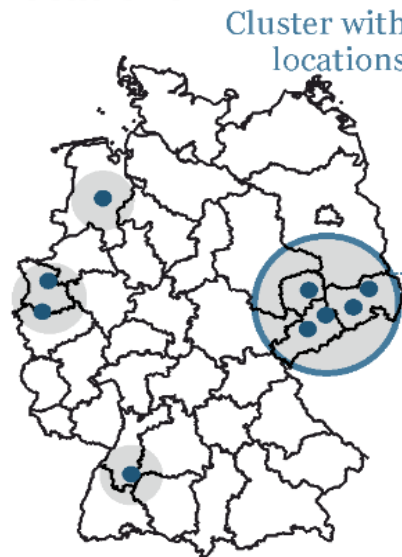
Detection of potential locations in text using NER

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et **accusam** et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy **eirmod** tempor invidunt ut labore et dolore magna aliquyam erat, sed diam **voluptua**. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren,

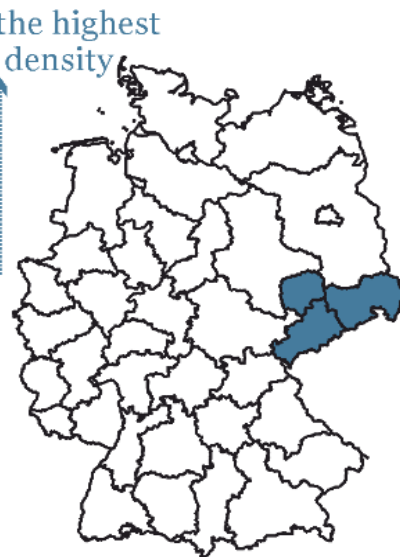
Geoparsing potential locations via gazetteer data



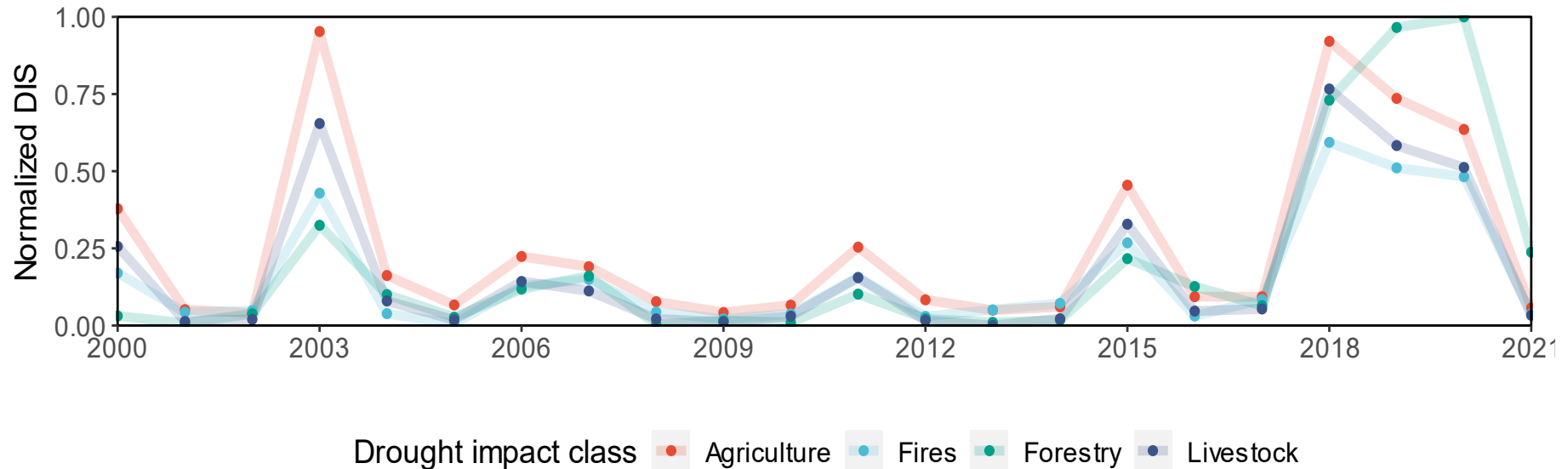
Sorting potential locations



Mapping the locations to NUTS units

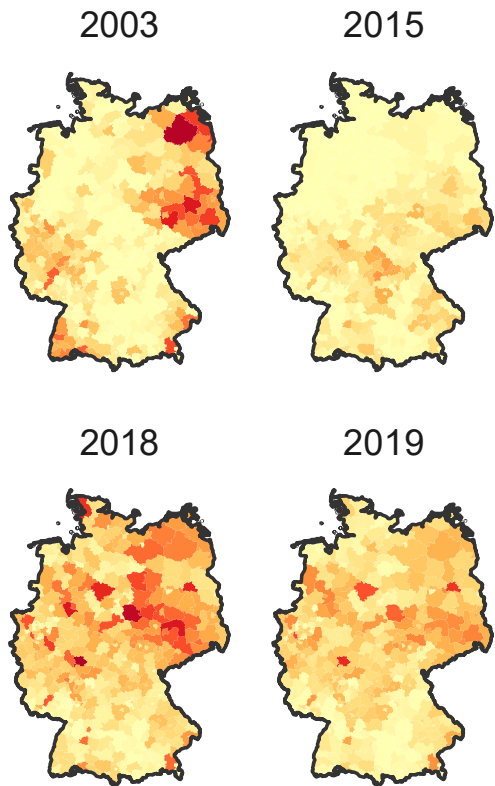


Drought impact statements (DIS)

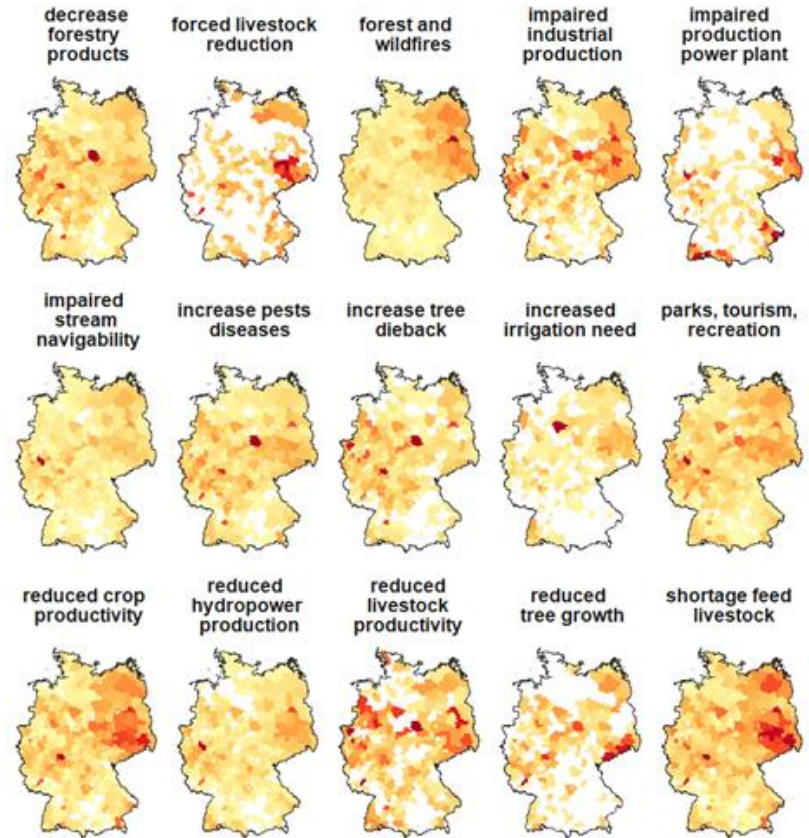
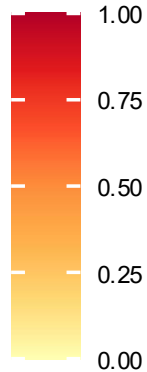


Sodge, J., Kuhlicke, C., de Brito, M.M. (2023) Automated spatio-temporal detection of drought impacts from newspaper articles using natural language processing and machine learning. *Weather and Climate Extremes*

DIS over time and space



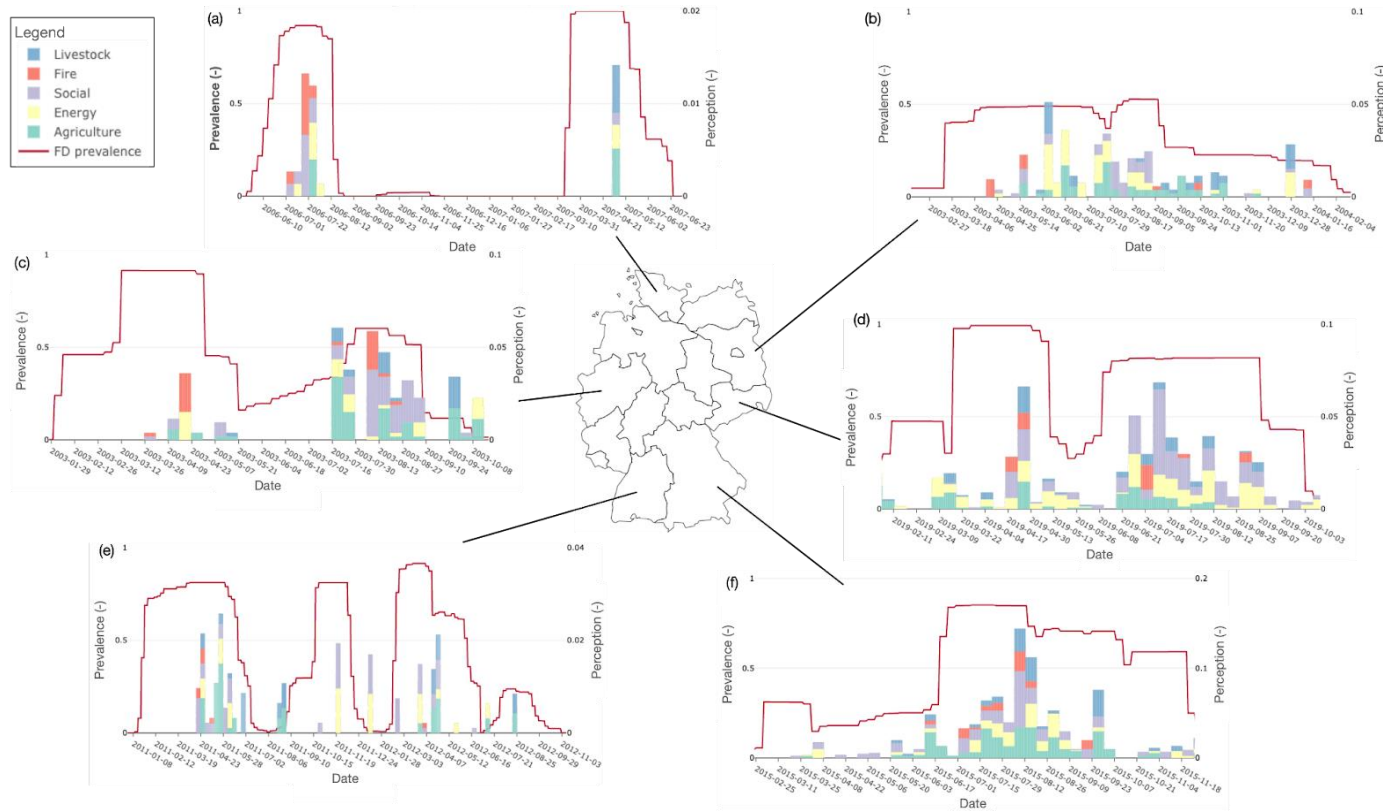
Normalized DIS



@JSodoge

Relative distribution of DIS
0.000 250.500 751.00

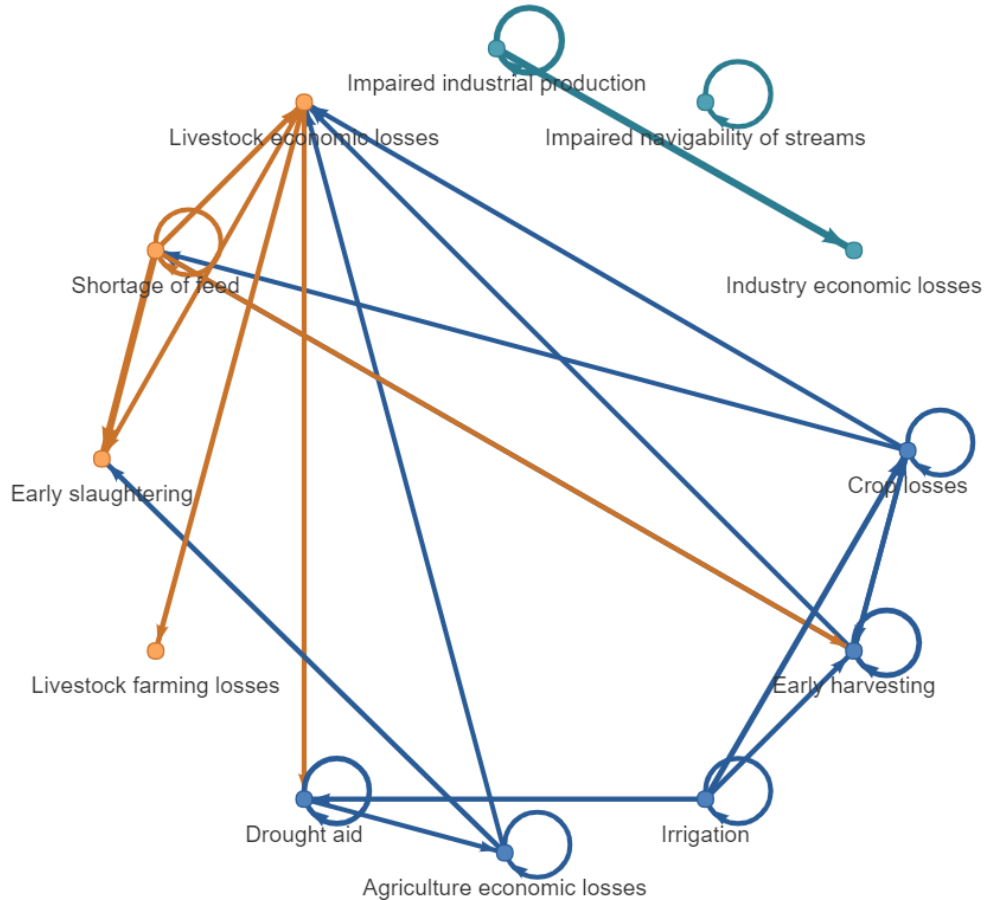
Flash drought occurrence vs impact reporting



Alencar et al. Flash droughts and their impacts – using newspaper articles to assess the perceived impacts of rapidly emerging droughts (under review)



Cascading impacts



de Brito, M.M. (2021) Compound and cascading drought impacts do not happen by chance: a proposal to quantify their relationships. *Science of the Total Environment*.



**Global database
of natural hazards
impacts reported in
scientific literature**



Query

ScienceDirect



drought

flood

storm

heatwave

coldwave

mass
movement

wildfire

AND

social

water availability

food production

health and
wellbeing

cities, settlement
and infrastructure

economic sectors



Duplicates = 5,716

n = 38,862

Expected outcome

Hazard type

Location

Date

Social impacts

Qualitative

Decrease in income

Decrease in GDP

Water impacts

Qualitative

Number of affected water sources

Food prod. impacts

Qualitative

Total production loss (per crop type)

Health impacts

Qualitative

Disease incidence (per type of disease)

Infrastructure
impacts

Qualitative

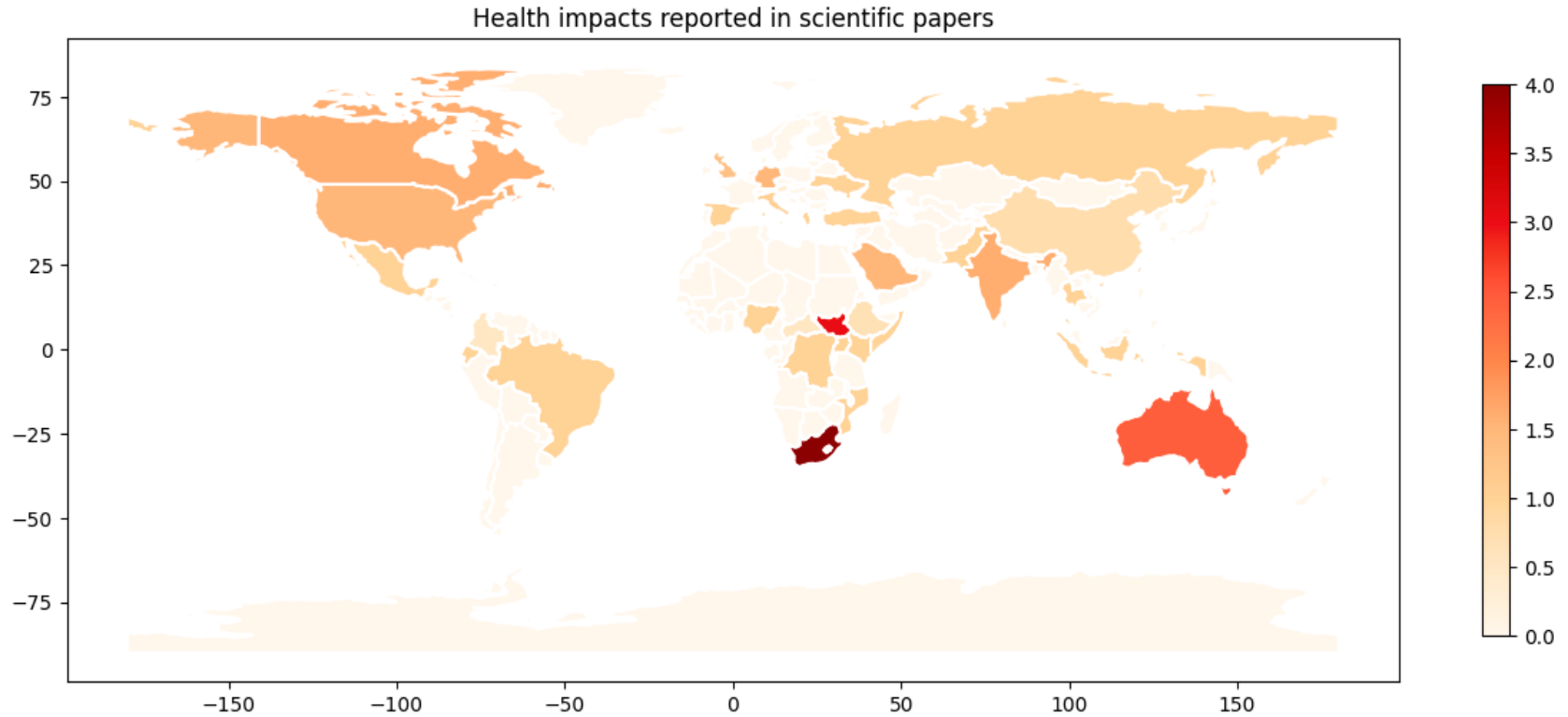
Number of damaged infrastructures

Economic impacts

Qualitative

Monetary losses (per sector)

Preliminary results



Text as data

- Unprecedented abundance
- Investigation of social issues that are normally hard to quantify
- Complementary data source



@m_de_Brito