

Tracking down extremist narratives in German Telegram channels

Andreas Blombach^{1, 2}, Philipp Heinrich¹

¹Lehrstuhl für Korpus- und Computerlinguistik, Department Germanistik und Komparatistik

²Department Digital Humanities and Social Studies

Düsseldorf, 15 November 2023

1. Introduction

2. Telegram corpus

3. Categorisation and manual annotation

4. Automatic identification of conspiracy narratives

5. Conclusion

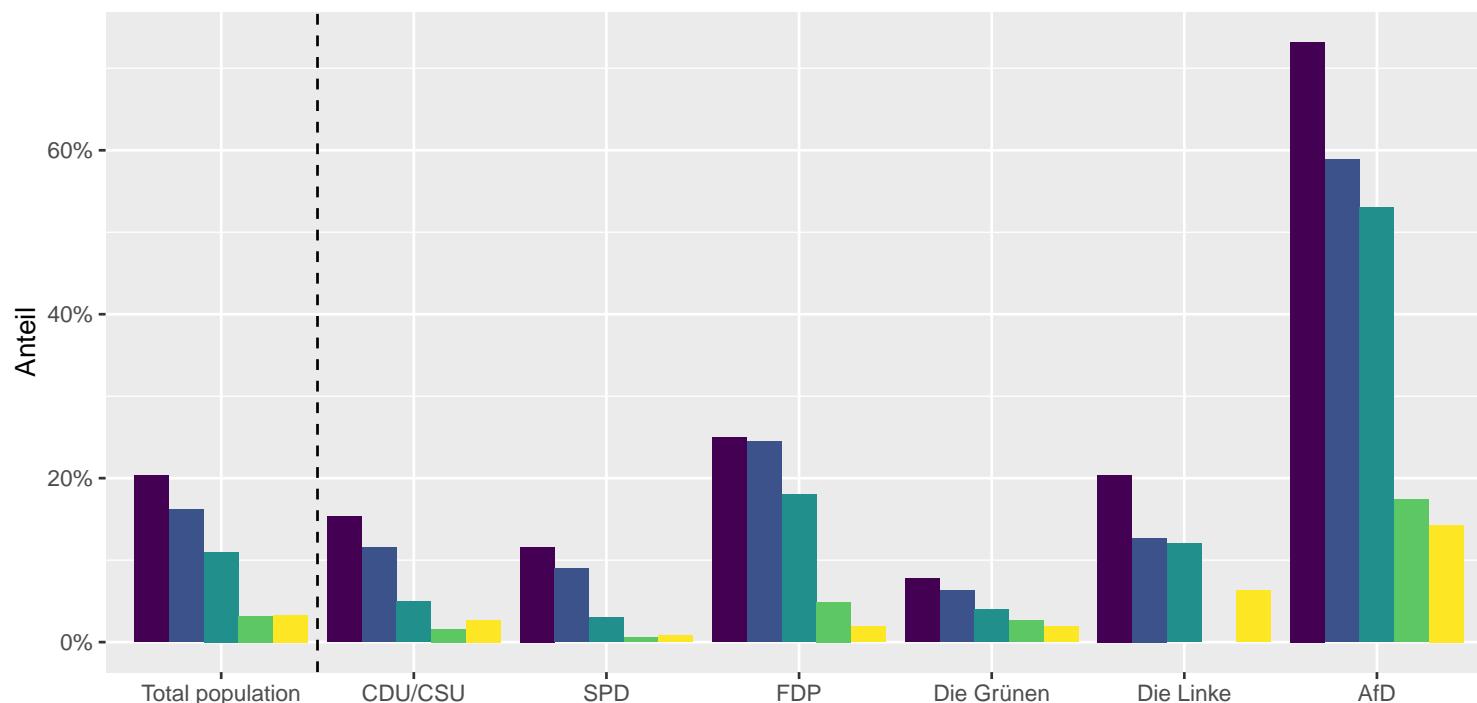
- In early 2020, the WHO warned that “we’re not just fighting an epidemic; we’re fighting an infodemic”.¹
 - Misrepresentation of medical facts and political processes surrounding the pandemic
 - Disinformation
 - Conspiracy narratives
- Somewhat predictable:
 - Mostly unmoderated online discussions
 - “Belief in conspiracy theories is stronger under conditions of uncertainty” (Douglas et al. 2019)
- One in five citizens in Germany believes
 - that the dangers of SARS-CoV-2 have been intentionally exaggerated to deceive the public (Infratest dimap 2020) and
 - that many figures and statistics on COVID-19 are falsified to spread fear (Allensbach 2022)

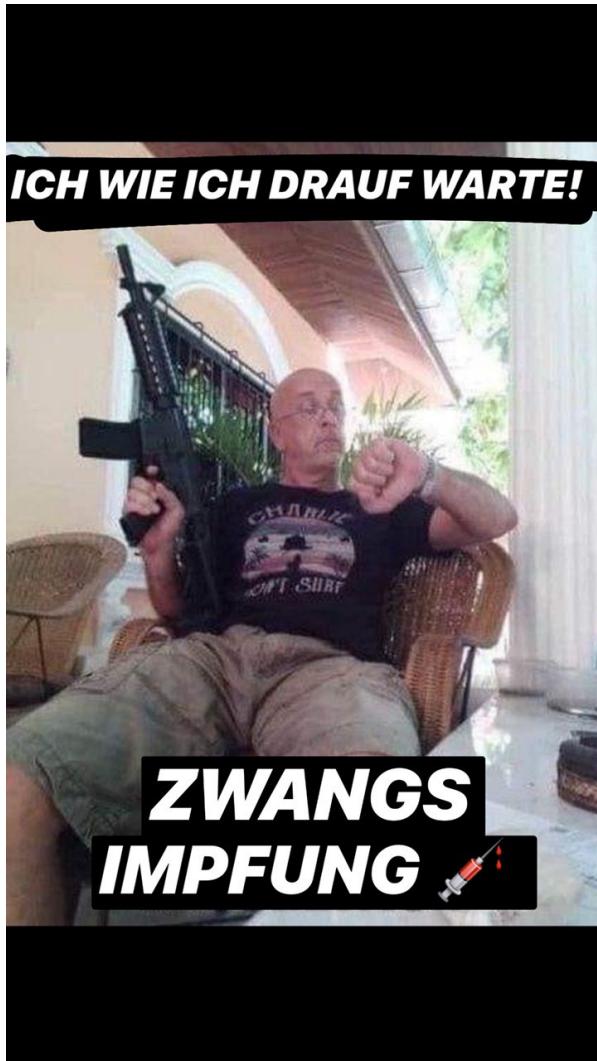
¹<https://www.who.int/director-general/speeches/detail/munich-security-conference>

Agreement with selected conspiracy theories by party affiliation

Source: Allensbacher Archiv, IfD-Umfrage 12049 (2022)

- █ Many figures and statistics on COVID-19 are falsified to spread fear.
- █ COVID-19 vaccination is more dangerous than the disease itself. The vaccination causes severe damage.
- █ I am afraid that they want to implant microchips in all people so that they can control them better.
- █ The government is using the crisis as an excuse to restrict citizens' fundamental rights.
- █ Bill Gates is the real culprit in the Corona crisis. He is using the crisis to establish a new world order.





2021-2022 Tracking the Infodemic: Conspiracy theories in the corona crisis

2022-2025 Die Normalisierung rechtspopulistischer und neurechter Diskurse in Japan und Deutschland

Lehrstuhl für Korpus- und Computerlinguistik

- Prof. Dr. Stephanie Evert
- Andreas Blombach
- Bao Minh Doan Dang
- Nathan Dykes
- Philipp Heinrich
- Dr.-Ing. Besim Kabashi
- Dr. Leonardo Zilio

Lehrstuhl für Japanologie (Moderne und Gegenwart)

- Prof. Dr. Fabian Schäfer
- Tamara Fuchs
- Linda Havenstein
- Dr. Olena Kalashnikova
- Stevie Poppe

- analysis of distribution of conspiracy narratives on social media
- typical linguistic patterns and discursive strategies (→ Tamara)
 - especially w.r.t. the overlap with right-wing extremist and populist discourses
- **automatic identification of conspiracy theories and related narratives**

1. Introduction

2. Telegram corpus

3. Categorisation and manual annotation

4. Automatic identification of conspiracy narratives

5. Conclusion

German Telegram channels

<https://telemetr.io/en/channels?languages=de>



Name	Subscribers	Increase	Average views per post
Corona-Infokanal des Bundesministeri... @corona_infokanal_bmg	289.0k	-10.2k	~106.4k
reitschuster.de @reitschusterde	280.9k	-8.3k	~153.3k
Eva Herman Offiziell @evahermanoffiziell	210.1k	-2.4k	~36.2k
AUF1 @auf1tv	198.5k	+1.2k	~107.8k
Freie Medien @freimedientv	178.4k	-335	~49.0k
Free Fire Redeem Codes @freefireredeem1	170.1k	-12.2k	~39.8k
Alles Ausser Mainstream @allesaussermainstream	164.0k	-1.5k	~46.6k
Reiner Fuellmich 🇩🇪	158.6k	-2.5k	~90.3k
Oliver Janich öffentlich @oliverjanich	156.3k	-1.9k	~29.3k
Ken Jebson - Aufklärung und Informati... @kenjebson	149.8k	+1.1k	~67.3k
Freie Sachsen @freiesachsen	148.5k	+217	~79.1k
LION Media @lionmedianews	139.2k	-681	~38.2k
Frag uns doch! Just ask us! WWG1WGA @fragunsdoch_wwg1wga	138.8k	-432	~63.6k
Krypto News Deutsch 🇩🇪 @kryptokompassde	133.0k	+11.0k	~65.9k
Q7 4 YOU @q74you	132.7k	+1.0k	~46.2k
Covid-19-QR-Codes und Impfzertifikate @covid19digitalerimpfausweis	131.0k	0	~2.8k

#	Name	Subscribers	Growth (7d)	Post views
1	Markus Krall Offiziell ✅ @markuskrall_a	306 974	+ 7 386	43 157
2	AUF1 ✅ @auf1tv	252 573	+ 60% ▲ 39%	120 119
3	Reitschuster.de ✅ @reitschusterde	235 155	+ 391	110 869
4	Vivoterra ✅ @vivoterra	229 953	+ 10 501	19 937
5	Mahakal Official Book ✅ @mahakalonlineofficial	229 048	+ 1 443	25
6	Eva Herman Offiziell ✅ @evahermanoffiziell	185 284	+ 268	34 672
7	Aktien mit Kopf ✅ @aktienmitk	171 920	+ 16 880	6 879
8	Infokanal des Bundesministeriu... @corona_infokanal_bmg	169 074	+ 785	45 061
9	Tim K. ✅ @timmkellner	166 528	+ 644	97 605
10	Freie Sachsen ✅ @freiesachsen	148 746	+ 25	58 281
11	Freie Medien ✅ @freimedientv	144 667	+ 430	35 182
12	Ken Jebson - Aufklärung und In... @kenjebson	144 208	+ 88	47 946
13	Thomas Röper-Anti Spiegel ✅ @antispiegelOffizieller	139 970	+ 589	24 232
14	Holistische Gesundheit, Heilung... @holistischegesundheitheilung	137 240	+ 786	70 872
15	Wetteradler 🌎 Umwelt & Tec... @wetteradlerkanal	135 997	+ 99	71 632
16	LION Media ✅ @lionmedianews	132 647	+ 77	54 673
17	Mäckle macht gute Laune ✅ @macklemachtguteaune	130 747	+ 97	54 205
18	Alles Ausser Mainstream ✅ @allesaussermainstream	130 716	+ 230	31 964

Data collection

- Identification of known conspiracy theorists and platforms
 - KenFM, Eva Herman, Attila Hildmann, Bodo Schiffmann, Oliver Janich etc.
- Scraping of selected public channels and associated groups
- Addition of frequently mentioned/cited channels to the corpus (iteratively)

Telegram corpus

- over 200 channels, approx. 100 public group chats
- over 13 million posts
- over 400 million tokens
- high proportions of forwarded messages ($M = .25$, $SD = .23$)
- high proportions of messages containing images ($M = .27$, $SD = .19$)
- videos also very common ($M = .16$, $SD = .16$)

Linguistic annotation

- JSON to XML
- Tokenisation: SoMaJo^a
- POS tagging: SoMeWeTa^b
- Lemmatisation: SMOR + heuristics^c

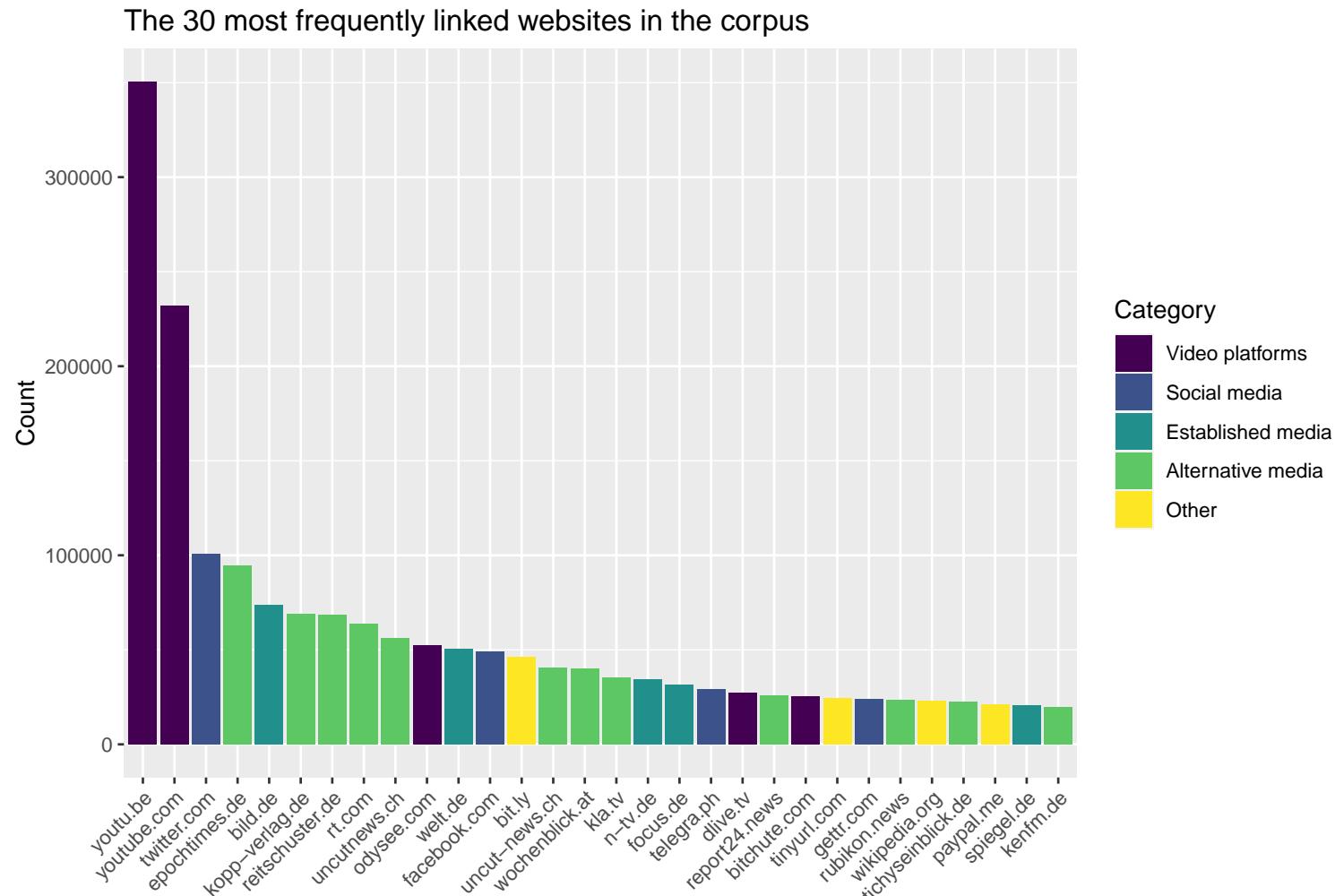
^a<https://github.com/tsproisl/SoMaJo>

^b<https://github.com/tsproisl/SoMeWeTa>

^cSee <https://github.com/fau-klue/empirist-corpus/tree/master/baselines>

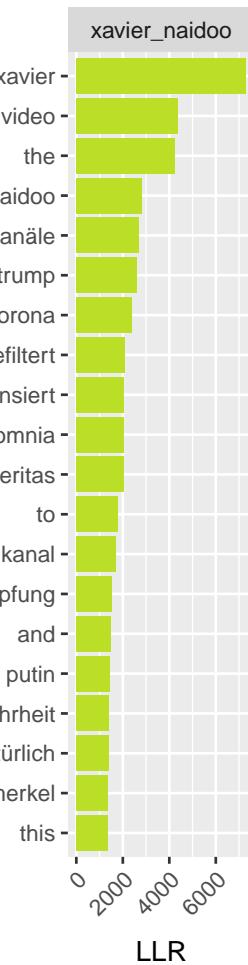
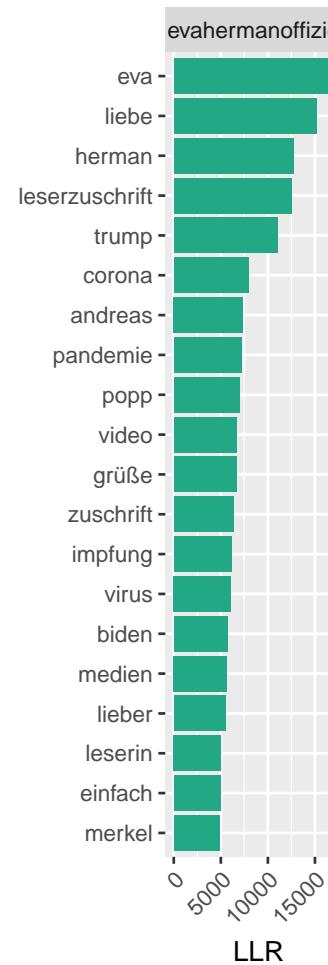
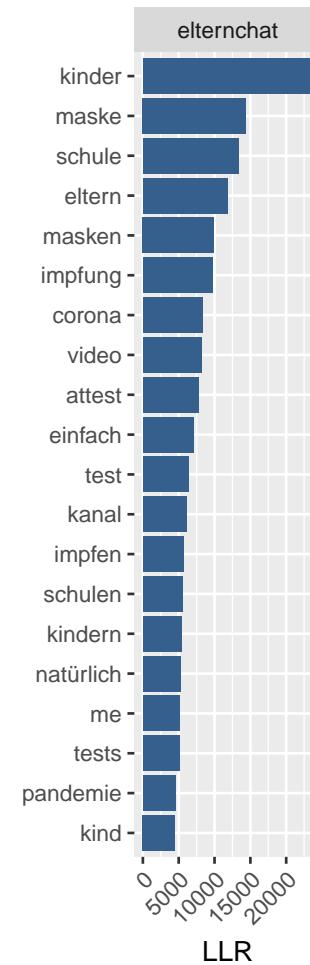
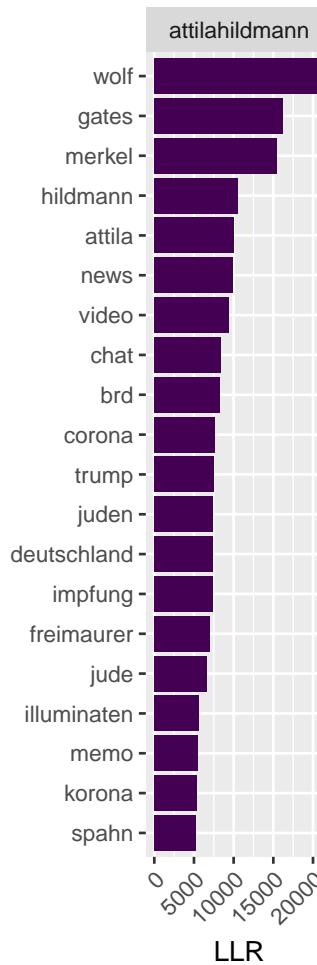
...erheit , statt seriosen zu das zu ? * warum ging die „ normale Grippe“ wieder im Jahr 2020 mit so „ zurück ? “ wenn der Rückgang der ...der PCR-Test auch diese Mutationen problemlos findet , warum sollte dann der PCR-Test nicht auch auf " normale " Coronaviren einer Diese Tatsache ist auch der Grund warum als Unterschied zwischen Covid-19 und der + offiziell im Schweizer Fernsehen und Presse : Virologen , Chefarzte usw erklären Covid-19 zur " Keiner kennt einen , der einen kennt 😅 absolut ! Ich kenne 2 positiv getestete (was ja nix heisst) , war eine Das alles wegen einer Stimmt genau . Die Geimpften können sich aber auch an Geimpften anstecken . Nach der Impfung sind Er sagte zu mir , dass ihm bekannt ist , dass covid 1 und covid 2 sich verhält wie eine Kauf 3 schnell tests das ist eine Ebenfalls laut WHO sterben aber jährlich an der ...erdem werden alle Teststellen für die Pandemie verwendet . Jeder mit Verdacht auf Covid wird in Quarantäne gesteckt , auch mit einer Bei uns waren schon 8 Leute positiv , 7 hatten keine Symptome und einer hatte eine Erinnert ihr euch noch ? Wie war das nochmal ? Eine ioniert und wie man die Viren verändert ... es wird noch viele geben . Bin derzeit auch erkrankt , soll die britische sein ... aber ist eher eine ...t dann vielleicht etwas zu spritzen was mit irgendwelcher Elektronik versetzt ist ? Es wird immer neue " Mutanten " geben (wie bei einer 😱 wahrscheinlich So nochmal zum denken Wissenschaftlich ist belegt das das coronavirus genau so harmlos ist wie eine ...login , Angela Speisberg , Ex-Frau von SPD-Gesundheitspolitiker Karl Lauterbach . Das Coronavirus in Deutschland sei wie eine ganz OMG natürlich gibt es Viren was ist denn eine ...n sogenannten virus aber der ist nicht annähernd so giftig wie sie es den leuten weiss machen wollen es was wie eine Angina oder eine Erinnerung : Das CDC bestätigt auf seiner Webseite , dass der Antikörper-Test nicht zwischen einer " Long Covid " : Auch eine ...lust des Geruchs- und Geschmackssinns wird derzeit ja exklusiv der Mikrobe zugeordnet , ist jedoch ein bekanntes Symptom für die Viel weniger Tote , als bei einer Schiss ? Du hast dich doch impfen lassen , weil du schiss vor einer Da kann ich fast mit jedem Punkt mitgehen . Nur halte ich das Virus für weitaus gefährlicher als eine Die Man konnte das Corona-Virus gar nicht zuverlässig von der Wie aus einer Alle , die corona hatten und das waren sehr viele !! Hatten zu 99 % NULL Symptome , nur positiven Test ... die anderen hatten ...r . Ich denke auch das sich der ganz normale Grippe Virus auch auf die Lunge legen kann . Wieso auch nicht . Und somit werden zur der	normale Grippe normalen Erkältung normale Grippe harmlosen Grippe normale Erkältung leichten Grippe harmlose Erkältungen normale Grippe normale Grippe NORMALEN GRIPPE normalen Grippe leichte Erkältung leichte Erkältung leichte Grippe normalen Grippe normale Grippe normale Grippe normale Grippe normalen Erkältung normale Grippe normale Grippe harmlosen Grippe normale Grippe normalen Grippe leichten Erkältung leichte Grippe normalen Grippe	durch die wichtigsten zu ständige Vorkommen seit 2011 , warum breitet sich corona dann immer noch aus ? * wenn ein starkes imm... reagieren ? hauptsächlich dieser trockene Husten angeführt wird . Da es nichts gibt was der Körper durch Husten aus der Lunge transportieren ... " mit niedrigerer Mortalitätsrate als andere Grippeverläufe , Zahlen seien unkorrekt , Durchsuchungsquoten seien verzerrt und di... 😊 Dieses Lügengebäude ist unerträglich ! für Geimpfte lebensgefährlich (Organversagen) . Man impft so zu sagen das Immunsystem weg !!! .. Lieber eine Weile auf alles ver... . Davor bräuchte man eigentlich keine Angst haben , wenn man ein intaktes Immunsystem hätte und keine Vorerkrankungen . corona gibts nicht bis zu 500.000 Menschen einer pro Minute und das ist KEINE Epidemie ! , es wird nicht getestet und daher nicht diagnostiziert . Darüberhinaus können Menschen , die schon eine Corona-Erkrankung hatten un... ^^ . Durch diese Tatsache ist vielen bei uns schon ein Licht aufgegangen , dass da etwas nicht stimmt Aber ich glaube dir auch , das... ... und was hat die Regierung draus gemacht ? Einen Fear-Porn-Virus . Nichts anderes . Aus der Mücke nen Elefanten gemacht ... ich ch... , nur mit extremer Mattigkeit ... fang bei kleiner Anstrengung an zu schwitzen ... auch) und man etabliert eine routinierte Denkweise " Ach die Regierung will uns ja nur helfen , schließlich haben uns die älteren Impfu... Es ist auch belegt das der test nicht funktioniert so auch kein covid19 nachweisen kann Es ist nicht mal belegt das es covid19 in Deuts... verlaufen , erklärt sie im RTL-Interview ihre umstrittene These . (...) (Viren) Es gibt auch Bakterien , und was bist du für ein Trampel hier Kindern was zu wünschen unfassbar hau ab sonst vergess ich mich ... alles total übertrieben und dem Coronavirus differenziert . https://www.cdc.gov/coronavirus/2019-ncov/faq.html kann Langzeitfolgen haben https://de.rt.com/inland/112813-long-covid-auch-normale-grippe/ . , dafür Milliarden-Gewinne für die Pharma-Konzerne . hastest !! Du hast hier nichts verloren du Wahrheitsleugner ! Verlass lieber die Gruppe !! . Und solange Sie nicht schreiben , dass das ein geheimer Plan einer Elite sei , die Kinder in unterirdischen Tunneln foltert und durch eine... ist anscheinend an Corona verstorben . unterscheiden . Trotzdem (oder gerade deshalb ?) verzeichnete man einfach jeden mit Grippe-Symptomen als Corona-Kranken . So k... eine " tödliche Pandemie " konstruiert wurde . oder Erkältungssymptome Saison jetzt , auch garantiert viele Influenza Fälle als corona abgestempelt .
--	---	--

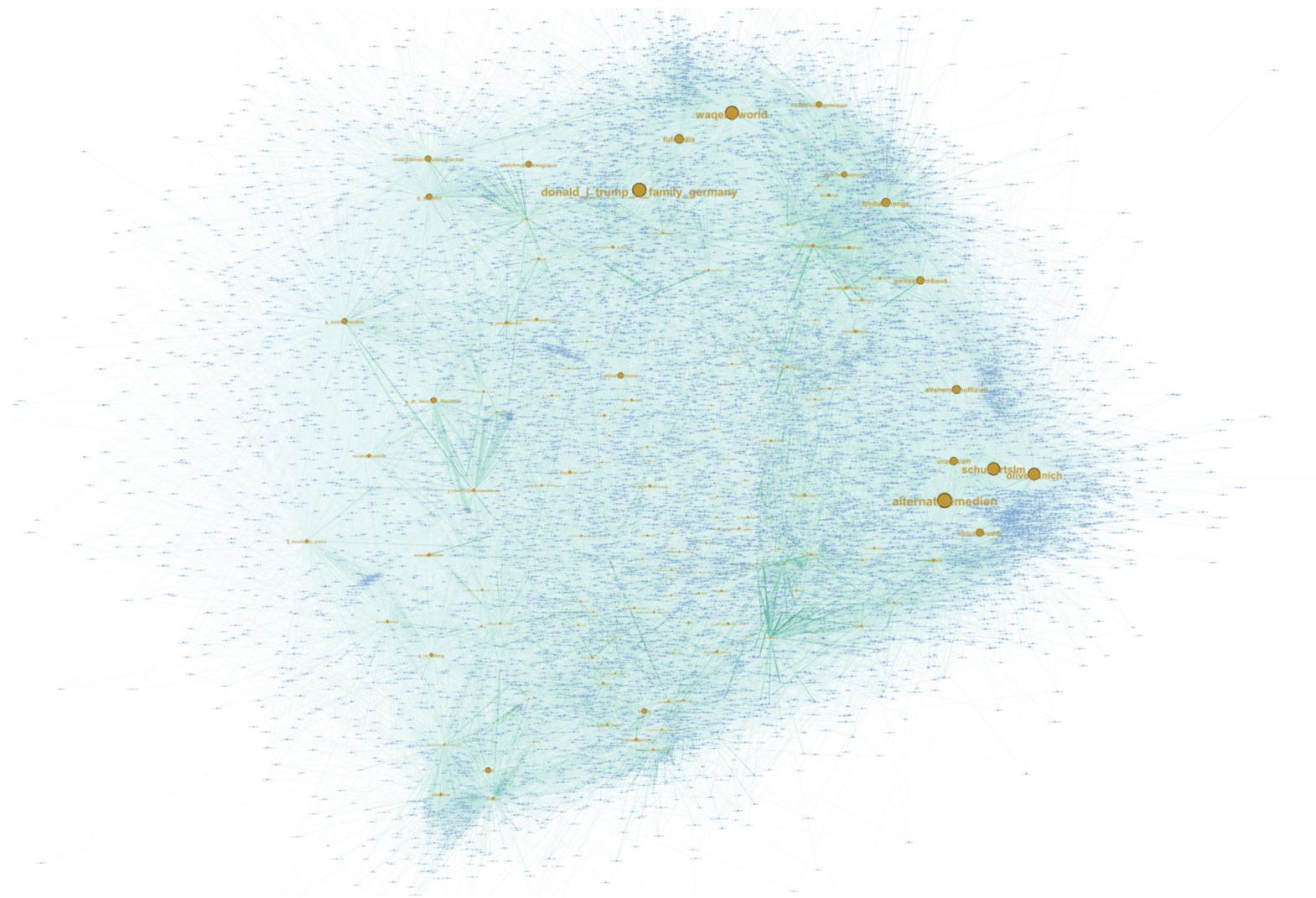
https://corpora.linguistik.uni-erlangen.de/cqpweb/schwurpus_v2 (access upon request)



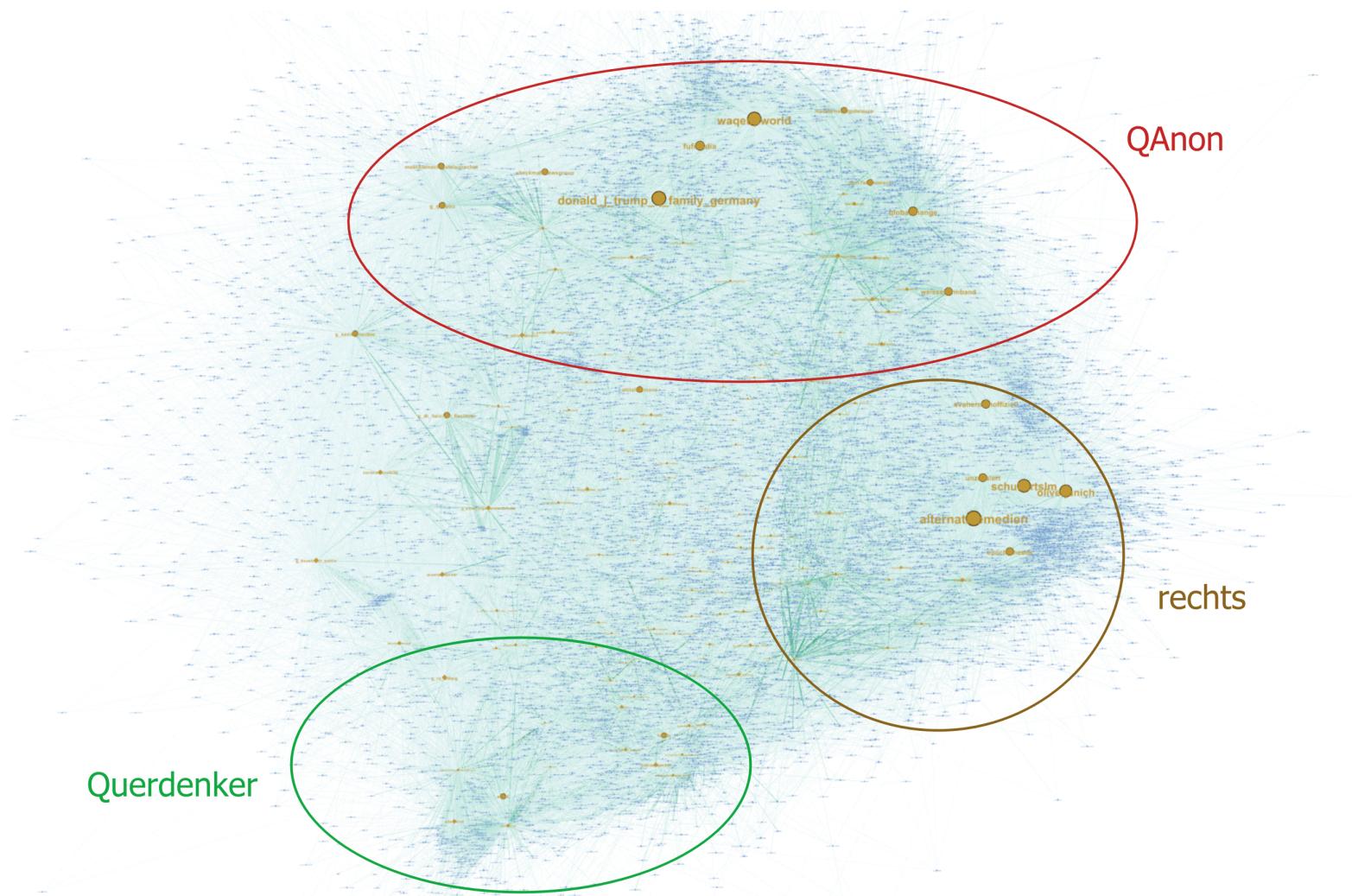
Keywords in selected channels

Top 20 Keywords (LLR)

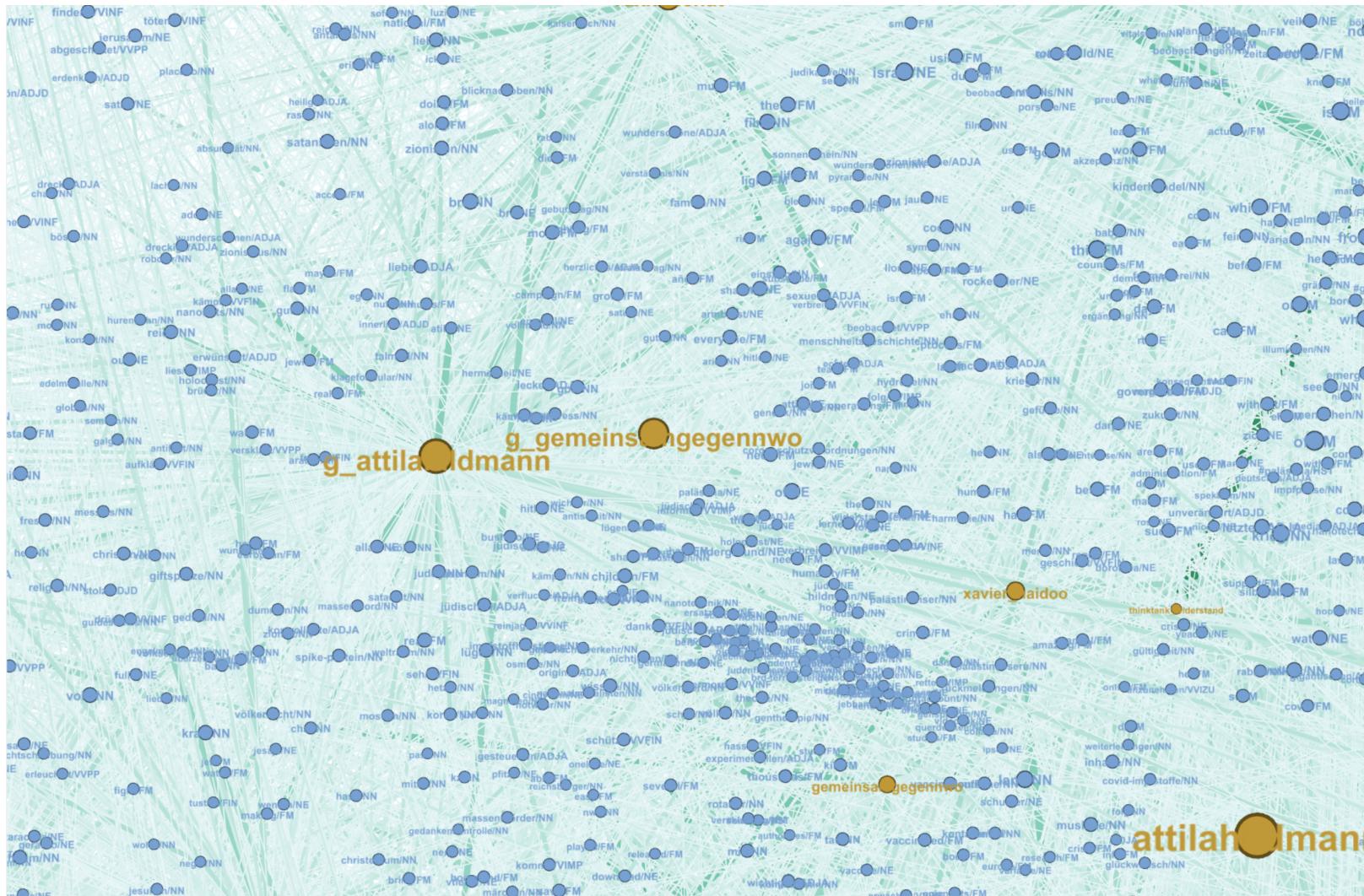




Keyword network

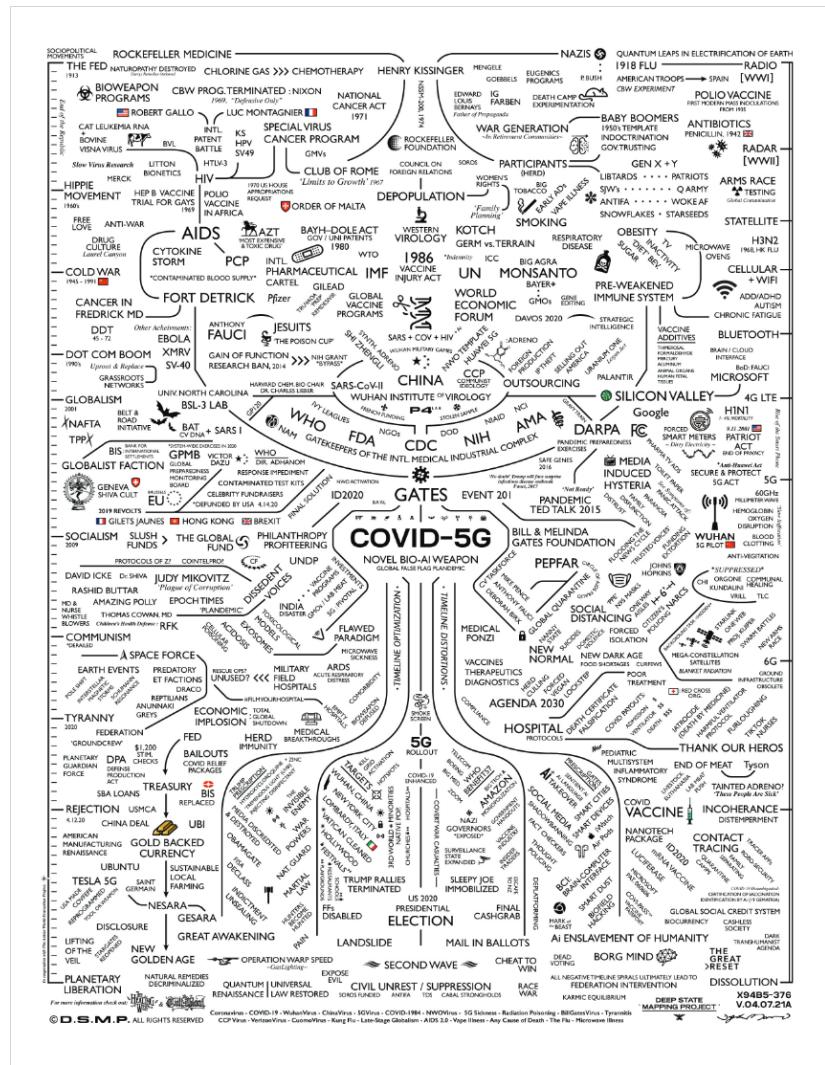


Keyword network: zoomed in



-
1. Introduction
 2. Telegram corpus
 3. Categorisation and manual annotation
 4. Automatic identification of conspiracy narratives
 5. Conclusion

Categorisation: even conspiracy believers can lose track of all the connections



Source:

<https://deepstatemappingproject.weebly.com/>

- Each post either contains “Geschwurbel” or not
- “Geschwurbel”: ‘drivel’, i. e. conspiracy-related or conspiracy-adjacent content
- “Geschwurbel” needs to be specified:
 - Currently 63 fine-grained categories
 - Based on previous work (e. g. Kuhn et al. 2021), domain knowledge and close reading prior to sampling
 - Narrative groups for similar categories
 - Categorisation scheme (including examples) available here: <https://github.com/fau-klue/infodemic>

Taxa	Kategorie	Narrativ	Beschreibung
Skepsis / Falschinformation	Gefahr	Corona wird verharmlost	Corona ist viel gefährlicher, als offizielle Stellen behaupten. Die potentiellen Auswirkungen des Virus werden heruntergespielt, um Angst vor Corona zu verhindern.
	Coronaverharmlosung	Corona ist harmlos	Man muss keine Angst vor Corona haben. Vom Coronavirus geht keine oder zumindest keine große Gefahr aus. Corona ist nicht gefährlich.
		Corona gibt es nicht	Die Existenz des Coronavirus ist nicht belegt. Corona gibt es nicht.
		Pseudopandemie	Die offizielle Darstellung der sogenannten Pandemie ist unberechtigte Panikmache. Die Virusverbreitung ist viel geringer als offiziell behauptet.
		Pseudoexperten	Wichtige Stimmen und Entscheidungsträger in der Pandemie, wie Gesundheitsminister Jens Spahn, RKI-Präsident Lothar Wieler und andere Experten sind Pseudoexperten.
	5G	5G	Die Strahlung des neuen 5G-Netzes ist schädlich für Menschen. Sie schwächt das Immunsystem, schädigt menschliche Zellen und führt zu anderen gesundheitlichen Problemen.
	Maßnahmenkritik	Maßnahmen illegal	Die staatlichen Maßnahmen zur Eindämmung der Pandemie sind nicht mit dem Grundgesetz vereinbar und deshalb illegal. Der Umgang mit den Maßnahmen ist unzulässig.
		Diskriminierung	Menschen, die sich den Maßnahmen der Regierung widersetzen, indem sie keine Masken tragen, sich nicht testen oder gegen COVID-19 impfen, werden als Rassisten, Antisemiten oder anderen verhöhnt und diskriminiert.
		Tests / Masken sind gefährlich	Corona-Tests und/oder Masken sind gefährlich. Der Stoff der Masken und Teststäbchen enthält Morgellons (bewegliche schwärzliche Partikel), die in die Haut eindringen und zu Schäden führen.
		Gegenmaßnahmen gefährlicher als Virus	Die psychischen und physischen Belastungen durch Lockdown, Maskenpflicht, Tests, Schulschließungen, Geschäftsschließungen und andere Maßnahmen sind gefährlicher als das Virus selbst.
	alternative Gegenmittel	Bleich- und Desinfektionsmittel	Die Einnahme oder Injektion von Bleich- oder Desinfektionsmitteln (wie Chlordioxid, CDL, CDS, MMS usw.) tötet das Coronavirus.
		umfunktionierte Medikamente	Bewährte Medikamente wie Chloroquin/Hydroxychloroquin, Azithromycin oder Ivermectin helfen auch gegen COVID-19.
		sonstige Wundermittel	Diverse einfache Mittel helfen gegen COVID-19, u.a. Vitamine, Heißluft, Knoblauch, Zwiebeln, Alkohol, Zistrosenextrakt, Kuhdung, Honig, Salz und andere.
		wirksame Mittel werden uns vorenthalten	Es gibt wirksame, günstige und weniger schädliche Gegenmittel und Impfstoffe gegen Corona, die von Medien und Politik verschwiegen werden.
	Ursprung / Verbreitung	Labortheorie	Das Coronavirus ist künstlicher Herkunft, also menschengemacht. COVID-19 wurde in einem Labor gezüchtet, beispielsweise in China.
		Fremde als Pandemietreiber	Die Pandemie wird durch Ausländer (wie Chinesen), Migranten, Asylbewerber und/oder Muslime eingeschleppt und verbreitet.
	Gefährlichkeit der Impfung	abgetriebene Embryonen für Impfstoffherstellung	Für die Herstellung neuartiger mRNA-Impfstoffe gegen das Coronavirus (wie von BioNTech oder Moderna) werden menschlich abgetriebene Embryonen verwendet.
		Versuchskaninchen	Die Impfung gegen COVID-19 ist nicht ausreichend getestet worden und deshalb nicht sicher. Mögliche Langzeitschäden durch die Impfung sind unbekannt.
		Impfung ist schädlich	Die Impfung gegen COVID-19 ist sehr gefährlich für den Menschen. Sie zerstört das Immunsystem, das Lungengewebe, führt zu Entzündungen und anderen gesundheitlichen Problemen.
		Graphenoxid / Impfung macht magnetisch	Die Impfung enthält Graphenoxid, einen schädlichen Zusatzstoff, der das Immunsystem schwächt oder Menschen direkt krank macht.
		Shedding	Die negativen Auswirkungen der Impfung können sich über den Ausstoß des Spike-Proteins des Coronaviruses bei Geimpften durchsetzen.
		Zwangsimpfung	Die Bevölkerung soll zur Impfung gezwungen werden. Der Staat wird eine Zwangsimpfung einführen.
		Impfung macht unfruchtbar	Die Impfung macht unfruchtbar. Die Antikörper gegen das Spike-Protein des Coronaviruses, die sich nach der Impfung bilden, richten sich auf die menschliche Erbgut.
		Genveränderung	Die neuartigen mRNA-Impfstoffe verändern das menschliche Erbgut. Die Impfung führt zu Genmutationen bei den Geimpften.
		Plandemie	Die Pandemie ist eine Plandemie. Ein geheimer Plan steckt hinter der Pandemie. Im Hintergrund geht es nicht um die Bekämpfung des Virus, sondern um politische und wirtschaftliche Ziele.
Verschwörung / Plandemie	Ziel der Pandemie / Maßnahmen	Profit	Die Pandemie dient dem Profit und der Marktdominanz von Großkonzernen und mächtigen Einzelpersonen oder versteckten Eliten.
		Wirtschaftsschädigung	Die Pandemie dient dazu, der deutschen Wirtschaft und hiesigen Unternehmen zu schaden. Die Wirtschaft wird durch die Maßnahmen geschädigt.
		Klimalockdown	Die Pandemie dient der Durchsetzung von Klimaschutzzügen, indem der Auto- und Flugverkehr und die damit verbundenen Emissionen reduziert werden.
		Pharma industrie als Drahtzieher	Hinter der Pandemie steht in Wahrheit Big Pharma (oder aber die Pharmaindustrie steckt zumindest mit den eigentlichen Drahtzweibern zusammen).
		Juden als Drahtzieher	Israel, Zionisten, die Rothschilds oder die Juden stecken in Wahrheit hinter der Pandemie.
		Biowaffe	Das Coronavirus oder die Impfung ist in Wirklichkeit eine Biowaffe.
		Konflikt	Die Coronapandemie ist ein indirekter Angriff eines Staates (z.B. China) auf einen oder mehrere andere Staaten. COVID-19 dient als Instrument des Konflikts.
		Kontrolle	Die Pandemie (COVID1984) und der Lockdown dienen der Regierung dazu, die Bevölkerung zu kontrollieren und zu überwachen.
		Ablenkung	Die Pandemie dient in Wirklichkeit nur dazu, von anderen Skandalen oder Vorgängen abzulenken, beispielsweise dem Zusammenbruch der Weltwirtschaft.
		Great Reset	Die Pandemie dient in der Wirklichkeit der Umsetzung des Great Reset, also der Erlangung politischer oder wirtschaftlicher Weltmacht.
		Überbevölkerung	Das Coronavirus, die Maßnahmen zur Bekämpfung der Pandemie und/oder die Impfung dienen dazu, die Weltbevölkerung zu reduzieren.
		Mikrochips	Mit der Impfung gegen Corona oder über die Corona-Tests werden den Menschen Mikrochips oder Nanobots injiziert.
		Gates als Drahtzieher	In Wirklichkeit hat Bill Gates das Coronavirus entwickeln lassen oder ist zumindest in die vorsätzliche Verbreitung von COVID-19 involviert.

VTs	Allgemeine VTs	NWO	Eine geheime Elite aus Politik und Wirtschaft plant die NWO (neue Weltordnung, New World Order), die den Umsturz der derzeitigsten Weltordnung anstrebt.
		Geheimbünde	Eine Geheimbund wie die Freimaurer, Illuminaten oder Rosenkreuzer hat weltweit Politik und Wirtschaft infiltriert und lenkt im Interesse dieser Gruppen.
		Gedankenkontrolle	Es wird Technologie eingesetzt, die die Gedanken oder das Verhalten von Menschen kontrolliert oder überwacht.
		Chemtrails	Kondensstreifen von Flugzeugen sind in Wirklichkeit Chemtrails, absichtlich freigesetzte Chemikalien. Chemtrails dienen dazu, die Öffentlichkeit zu beeinflussen.
		False Flag	Öffentlich wirksame Ereignisse wie die Anschläge auf das World Trade Center (9/11), das Massaker von Butscha u.ä. sind in Wirklichkeit geplante Falle.
		Bargeld	Bargeld und nicht zurückverfolgbares Geld sollen abgeschafft werden.
		Enteignung	Das private Vermögen, das Eigentum, Besitztümer oder Geschäfte der Bevölkerung sollen zerstört, enteignet, verstaatlicht oder verboten werden.
	QAnon	Anti-Christentum	Eliten in Politik und Wirtschaft sind Satanisten oder vom Teufel beeinflusst oder arbeiten heimlich gegen das Christentum.
		QAnon	Q, ein anonymer Patriot mit Zugang zu streng geheimen Informationen der US-Regierung, teilt chiffrierte Informationen, Einschätzungen und Theorien mit.
		Kinder-/Frauenschädigung	Ein internationaler Kinderhändlerring unter Beteiligung von Prominenten aus Politik und Hollywood sowie von bekannten Organisationsgruppen.
	Rechte VTs	Großer Austausch	Es gibt einen geheimen Plan namens großer Austausch oder Bevölkerungsaustausch, in dem es darum geht, die weiße, christliche Bevölkerung zu verdrängen.
		BRD GmbH	Deutschland ist kein souveräner Staat, sondern wird noch immer von den alliierten Besatzungsmächten kontrolliert. Einen Frieden kann es nicht geben.
Allgemeine Narrative	Staat als Gegner	GMF	Ausländer, Juden, Muslime, Nicht-Weiße, Homosexuelle, Behinderte, Frauen, Gutmenschen, Linke oder andere Gruppen sind menschenfeindlich.
		Klimageschwurbel	Der menschengemachte Klimawandel ist nicht real, sondern eine große Lüge. Wenn es überhaupt Klimawandel gibt, so ist er nicht menschengemacht.
		Schlafschafe	Ein großer Teil der Bevölkerung ist gehirngewaschen und hat keine Ahnung, was wirklich in der Welt vor sich geht. Schlafschafe sind Menschen, die leicht manipuliert werden können.
		Millenarismus	Der Tag des großen Umsturzes, der großen Veränderung oder der Abrechnung steht bevor. Auch wenn es nicht so aussieht, verbergen sich dahinter geheime Pläne.
		Wahlbetrug	Die Wahlen sind nicht sicher und werden im großen Stil manipuliert.
	Indoktrination	Deep State	Es gibt eine versteckte Machtstrukturen innerhalb des Staates, den tiefen Staat, Deep State (DS) oder Staat im Staate.
		Demokratür	Deutschland ist kein Rechtsstaat, sondern eine Diktatur, in der politikgelenkte Korruption und juristische Willkür herrschen. Die Demokratie ist nur eine Fassade.
		Systemlinge	Eine Person oder Gruppe handelt vorgeblich aus freiem Antrieb, ist in Wahrheit aber ein Systemling, der vom System oder Staat bestimmt wird.
	Pseudowissenschaft	Lügenpresse	Die Medien oder der Mainstream lügen die Bevölkerung an, sie verschweigen Wahrheiten und Zahlen oder betreiben Gehirnwäsche.
		Systemmedien	Die Medien oder der Mainstream werden von einer Gruppe oder der Politik kontrolliert und gesteuert. Die Öffentlichkeit wird so manipuliert.
		Meinungsdiktatur	Kritische Stimmen werden im öffentlichen Diskurs und in sozialen Medien unterdrückt oder unsichtbar gemacht. Es herrscht eine Meinungsdiktatur.
	Esoterik		Es gibt Engel, Einstellungen, Heilkristalle, Gebete, Schwingungs- und Lichttherapien, die bei Erkrankungen helfen. Impfungen können schädlich sein.
		Medischwurbel	Die Schulmedizin oder westliche Medizin ist sehr fehlerhaft und sogar gefährlich. Alternativmedizinische Ansätze wie Entgiftung und Detoxification sind besser.

Pseudo-pandemic: **pseudo-pandemic**

Vaccine hazards: **genetic manipulation**, **infertility**, **guinea pigs**, **vaccine is harmful**

Alles wegen der Lügenpandemie! Kurz: **Impfstoff im 1. Halbjahr 2021 einsatzfähig!** Bei den Impfstoffen handelt es sich um einen **weltweiten Menschenversuch** mit **Stoffen, die unsere Gene verändern und nicht einmal richtig getestet wurden!** Wer sich impfen lässt, ist verloren! Unendlich viele werden sterben und krank werden und die Folgen sind noch gar nicht abzusehen! **Laut eines Phama-Whistleblowers werden 97% unfruchtbar!**

Pseudo-pandemic: **pseudo-pandemic**

Vaccine hazards: **infertility**

COVID-19 conspiracies: **Great Reset**, **depopulation**

Other conspiracies: **New World Order**, **chemtrails**

State as an enemy: **'Systemlinge'**

Other drivel: **climate change denial**

☒ ANSCHNALLEN! **BLICK IN DIE NWO GLASKUGEL**: Vorhersagen aus der Serie "Utopia" aus dem Jahre 2013! ☒

⚠ Sie sagen uns alles direkt ins Gesicht. Was kommen wird, wie sie es machen werden. Eine **Fake Pandemie**,
Impfstoffe die unfruchtbar machen sollen, **Bevölkerungsreduktion auf 500 Millionen**, **korrumptierte und erpresste**
Wissenschaftler, **Sprühflugzeuge**, **Geburtenkontrolle**, **Wirtschaftskollaps**, die Errichtung eines Weltstaates ohne Ethnien,
Klimalüge uvm.!

Pseudo-pandemic : narratives that downplay the danger of COVID-19, deny its existence or feed doubts about the official narrative of the pandemic

Criticism of countermeasures : narratives claiming that pandemic response efforts are illegal, more dangerous than the virus (e.g. masks causing illness) or that they discriminate against sceptics and people who refuse to wear masks, be tested or vaccinated

Alternative treatments : narratives about repurposed drugs or other ‘miracle cures’ against COVID-19 that are allegedly withheld from the population

Vaccine hazards : narratives portraying COVID-19 vaccines as insufficiently tested, unsafe, or even dangerous

COVID-19 conspiracies : conspiracy theories claiming some hidden agenda behind the pandemic, e.g. Bill Gates aiming to reduce the world population or to inject people with microchips to control them, or the pandemic serving to destroy the economy, to achieve climate change goals or to produce profit for big corporations and powerful elites

Other conspiracies : pre-existing conspiracy theories not specific to COVID-19 – chemtrails, claims about false flag operations, mind control, all-powerful secret societies etc.

QAnon : narratives about an anonymous individual called Q and his claims of insider knowledge about highly classified U.S. government documents, a Satanic cabal operating a global child sex trafficking ring, and Donald Trump’s secret fight against this cabal

Group-focused enmity : all forms of racism, xenophobia, Islamophobia, homophobia, misogyny etc., as well as the far-right conspiracy theories *Great Replacement* (claiming a plot to replace the ethnic white population with non-white immigrants, especially Muslims) and *BRD GmbH* (claiming that Germany never ceased to be controlled by the Allies after World War II, and rejecting the constitution and legitimacy of the modern German state in favor of the German Reich)

Sheeple covers the single narrative (often accompanying conspiracy narratives) that most people have no idea what is really going on, because they are brainwashed or choose to live in ignorance

Millenarianism : narratives about an upcoming day or time of great change or reckoning when the group's beliefs will be validated and/or its enemies will be defeated

State as an enemy : narratives questioning the status of democracy – by assuming a deep state that holds the real power, by accusing free actors of being covert agents of the system, by doubting election results or by accusing the state of having dictatorial features

Indoctrination : narratives claiming that the (mainstream) media is controlled by the state or a powerful group and/or that it lies, censors information and indoctrinates the population, as well as narratives about cancel culture and the death of free speech

Esotericism & pseudo-science : pseudo-scientific claims about medicine (e.g. disbelief in the existence of viruses in general) or alternative medicine, as well as various esoteric practices and beliefs (healing crystals, mediums, auras etc.)

Other drivel includes narratives with very low prevalence: climate change denial and narratives about a man-made origin of COVID-19

- Goal: annotate a sample of messages likely to contain relevant textual content
- Filter #1: no forwarded messages and posts containing images, videos or polls
- Filter #2: only consider longer posts (≥ 400 characters, excluding URLs)
- Sample: some channels and groups contain many more posts than others \Rightarrow stratify by month, channel/group and number of messages
- Result: 1099 posts from January 2020 to March 2022 (average post: 180 tokens, 11.4 sentences)
- Annotation: three (mostly) independent annotators

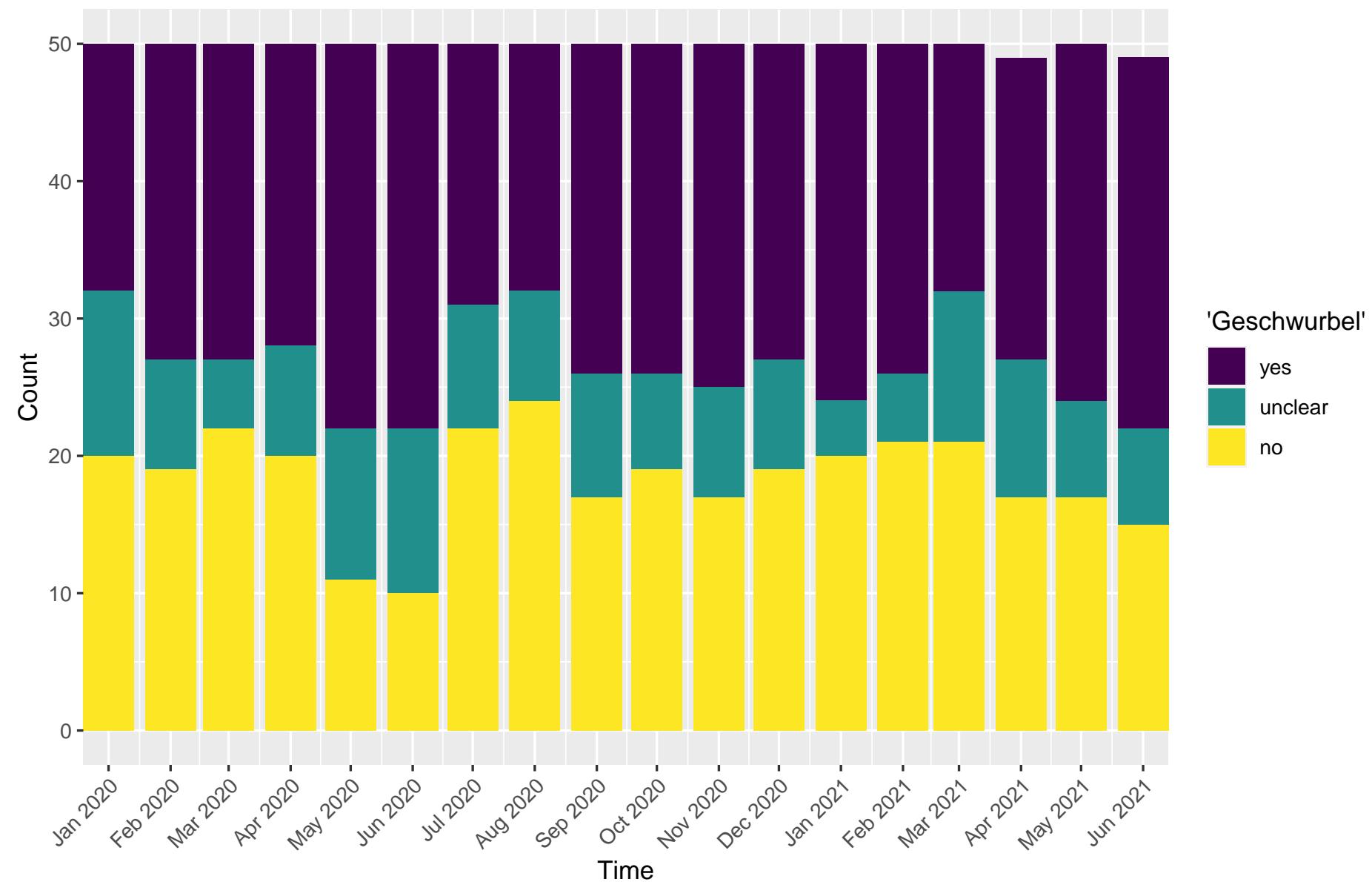
Pairwise Cohen's κ for the initial classification as 'Geschwurbel' or not (Fleiss' κ excluding the gold standard: .61, indicating substantial agreement):

	Rater 1	Rater 2	Rater 3
Gold	0.84	0.78	0.67
Rater 1		0.67	0.57
Rater 2			0.60

Fleiss' κ for narrative groups ranges from .32 (millenarianism) to .83 (other drivel); $M = .59$, $SD = .13$.

Telegram sample: how many posts contain 'Geschwurbel'?

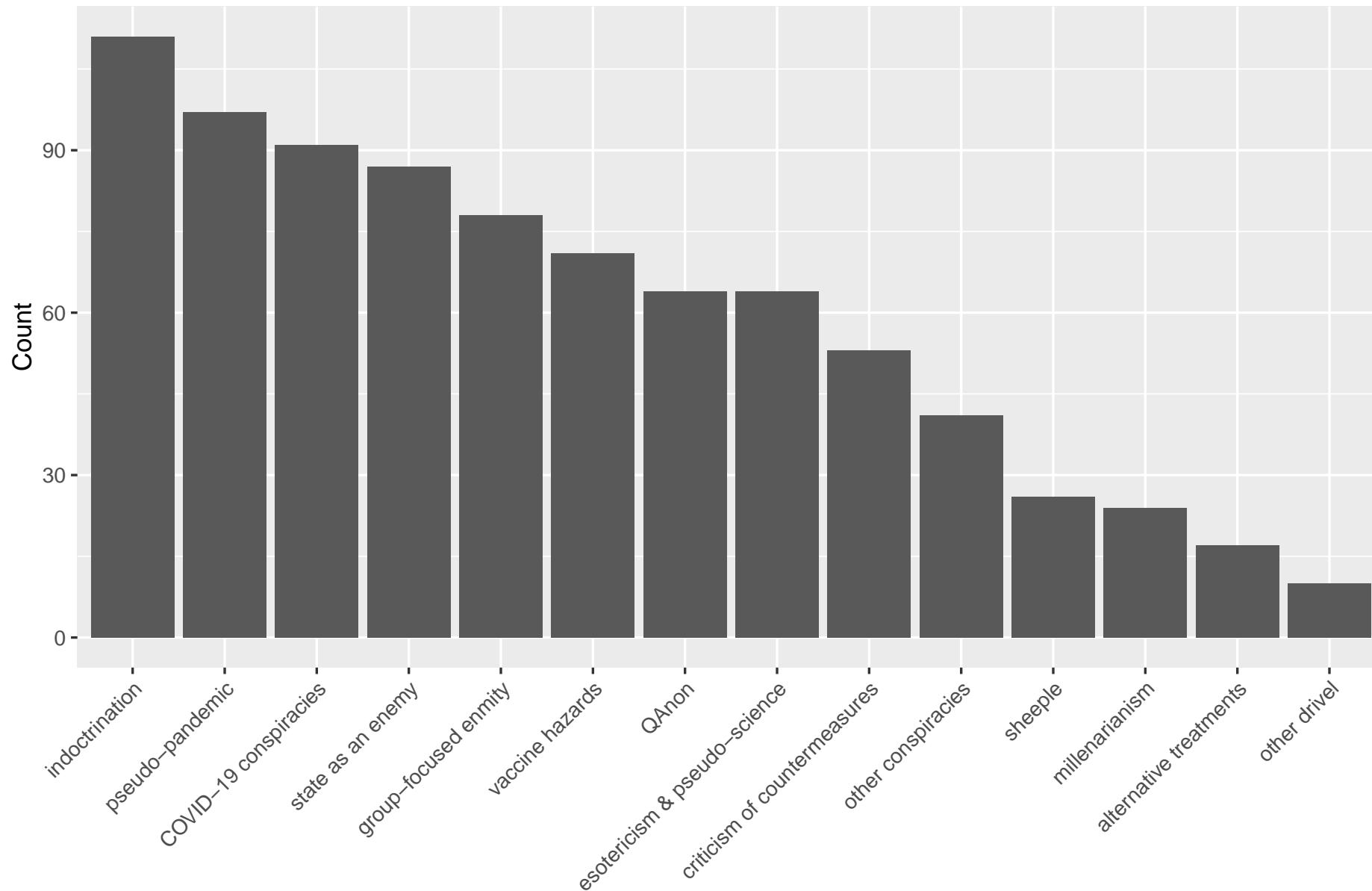
'Geschwurbel': conspiracy-related or conspiracy-adjacent content



'Geschwurbel'

- yes
- unclear
- no

Telegram sample: narrative groups



-
- 1. Introduction**
 - 2. Telegram corpus**
 - 3. Categorisation and manual annotation**
 - 4. Automatic identification of conspiracy narratives**
 - 5. Conclusion**

multi-label

- task: distribute set of labels to each post
 - one for each narrative group
- evaluation: separately for each narrative group
 - find model with best trade-off between precision and recall

two information sources

- label descriptions (created by domain experts)
 - here: concatenation of description sentences for fine-grained categories
- examples of posts with corresponding labels (annotated training data)
 - few-shot: narrative groups have low prevalence

vaccine hazards Die Impfung gegen COVID-19 ist nicht ausreichend getestet worden und deshalb nicht sicher. Mögliche Langzeitschäden durch die Impfung lassen sich nicht ausschließen. Die Antikörper gegen das Spike-Protein des Coronavirus, die sich nach der Impfung bilden, richten sich auch gegen das ähnliche Protein Syncytin-1, das eine entscheidende Rolle bei der Bildung der Plazenta spielt; dies führt zur Unfruchtbarkeit.

pseudo-pandemic Die offizielle Darstellung der sogenannten Pandemie ist unberechtigte Panikmache. Die Virusverbreitung ist viel geringer als offiziell angegeben. Die Corona-Tests sind extrem unzuverlässig und damit wertlos und irreführend.

COVID-19 conspiracies ID2020, ein Bündnis aus der Impfallianz Gavi, der Rockefeller-Stiftung, Microsoft und weiteren Unternehmen, strebt globale Überwachung an, indem digitale Impfausweise mit biometrischen und anderen persönlichen Daten kombiniert werden sollen. Unter dem Vorwand der Pandemie sollen gezielt unerwünschte Bevölkerungsteile wie alte Menschen, Arbeitslose oder Sozialhilfeempfänger reduziert werden.

indoctrination Die Medien oder der Mainstream lügen die Bevölkerung an, sie verschweigen Wahrheiten und Zahlen oder betreiben Gehirnwäsche. Bestimmte Informationen werden von den Medien unterdrückt oder censiert. Es gibt keine Pressefreiheit.

:

- „Bei den Impfstoffen handelt es sich um einen weltweiten Menschenversuch mit Stoffen, die unsere Gene verändern und nicht einmal richtig getestet wurden!“ → {*vaccine hazards*}
- „Glaubt mir es gibt kein Coronavirus oder dergleichen, das ist kompletter Schwachsinn.“ → {*pseudo-pandemic*}
- „Eine Fake Pandemie, Impfstoffe die unfruchtbar machen sollen, Bevölkerungsreduktion auf 500 Millionen, korrumptierte und erpresste Wissenschaftler, Sprühflugzeuge, Geburtenkontrolle, Wirtschaftskollaps, die Errichtung eines Weltstaates ohne Ethnien, Klimalüge uvm.!“ → {*pseudo-pandemic, vaccine hazards, COVID-19 conspiracies, other conspiracies, state as an enemy, other drivel*}
- „Auch wer geimpft ist, kann trotzdem noch ansteckend sein, wenn er sich mit dem Coronavirus infiziert.“ → {*no drivel*}
- ⋮

general strategies

- outsource the problem
 - student assistants
 - ChatGPT4 via OpenAI's API
- supervised
 - classic ML
 - few-shot (SetFit)
- zero-shot
 - sentence similarity (SBERT)
 - natural language inference (Zero-Shot-NLI)

general strategies

- outsource the problem
 - student assistants
 - ChatGPT4 via OpenAI's API
- supervised
 - classic ML
 - few-shot (SetFit)
- zero-shot
 - sentence similarity (SBERT)
 - natural language inference (Zero-Shot-NLI)

implementation details

- straightforward: we can train different models for each narrative
 - vs. ensemble model (leverages label co-occurrence statistics)
- we can predict a separate label “no drivel”
 - vs. if None → “no drivel”
- integration of label descriptions
 - supervised prediction: as positive training examples
 - zero-shot: essential

ML baseline

- bag of n grams (words, lemmas)
- tf.idf
- feature selection
- logistic regression, support vector machine, ...
- other vectorisation techniques possible
 - e.g. sentence embeddings

Few-shot classification

- generalise from a small amount of training data
 - (for some or all labels)
- SetFit (Tunstall et al. 2022): a framework for few-shot learning
- based on Sentence Transformers (Reimers and Gurevych 2019)
- fine-tuned on contrastive pairs of labelled texts
- encode training data + text classification head
- model (Huggingface):
paraphrase-multilingual-mpnet-base-v2

- main idea: individual sentences of narrative descriptions represent different ways in which the narrative can be expressed
- split posts p and narrative descriptions n into sets of sentences S_p and S_n
- compare sentence pairs and calculate score (similarity / NLI)

$$s(s_i, s_j) \quad \text{for each sentence pair } (s_i, s_j) \text{ with } s_i \in S_p, s_j \in S_n$$

- primary interest: given conspiracy narrative in post?
- use maximum over all scores of all sentence pairs as overall score for post p and narrative n

$$\text{score}(p, n) = \max_{s_i \in S_p, s_j \in S_n} (s(s_i, s_j))$$

sentence similarity

- encode all sentences using SBERT model: SentenceTransformer Reimers and Gurevych (2019)
- cosine similarity between sentence embeddings \mathbf{e}_{s_i} and \mathbf{e}_{s_j}

$$s(s_i, s_j) = \cos(\mathbf{e}_{s_i}, \mathbf{e}_{s_j}) \quad \forall s_i \in S_p, s_j \in S_n$$

- model (HuggingFace): distiluse-base-multilingual-cased-v1

sentence similarity

- encode all sentences using SBERT model: SentenceTransformer Reimers and Gurevych (2019)
- cosine similarity between sentence embeddings \mathbf{e}_{s_i} and \mathbf{e}_{s_j}

$$s(s_i, s_j) = \cos(\mathbf{e}_{s_i}, \mathbf{e}_{s_j}) \quad \forall s_i \in S_p, s_j \in S_n$$

- model (HuggingFace): distiluse-base-multilingual-cased-v1

natural language inference (NLI)

- predict whether sentences of post entail label descriptions (Yin, Hay, and Roth 2019)
- often used with single- or few-word labels, here: detailed description sentences

$$s(s_i, s_j) = \mathbb{P}(s_i \rightarrow s_j) \quad \forall s_i \in S_p, s_j \in S_n$$

- models (HuggingFace): gbert-large-nli, xlm-roberta-large-xnli, mDeBERTa-v3-base-xnli, mDeBERTa-v3-base-xnli-2mil7

system prompt

You will be provided with German Telegram posts from people who are potentially spreading COVID-19 misinformation and conspiracy theories.

Posts will be delimited with ~~~~ characters. Each post will be preceded by a unique numeric id on the first line.

Your job is a multi-label classification task. Each post can belong to one or more of 14 different classes. If none of these classes apply, a post is to be labelled "kein_Geschwurbel".

For each class or label, there are multiple description sentences in German in the form "label: description". These label descriptions, one per line, try to encapsulate the essence of several subclasses.

The labels and their descriptions are:

Impfung_ist_gefährlich: Für die Herstellung neuartiger mRNA-Impfstoffe gegen das Coronavirus (wie von BioNTech oder Moderna) werden menschliche Embryonen oder abgetriebene Föten benutzt.

Impfung_ist_gefährlich: Die Impfung gegen COVID-19 ist nicht ausreichend getestet worden und deshalb nicht sicher. Mögliche Langzeitschäden durch die Impfung lassen sich nicht ausschließen.

...

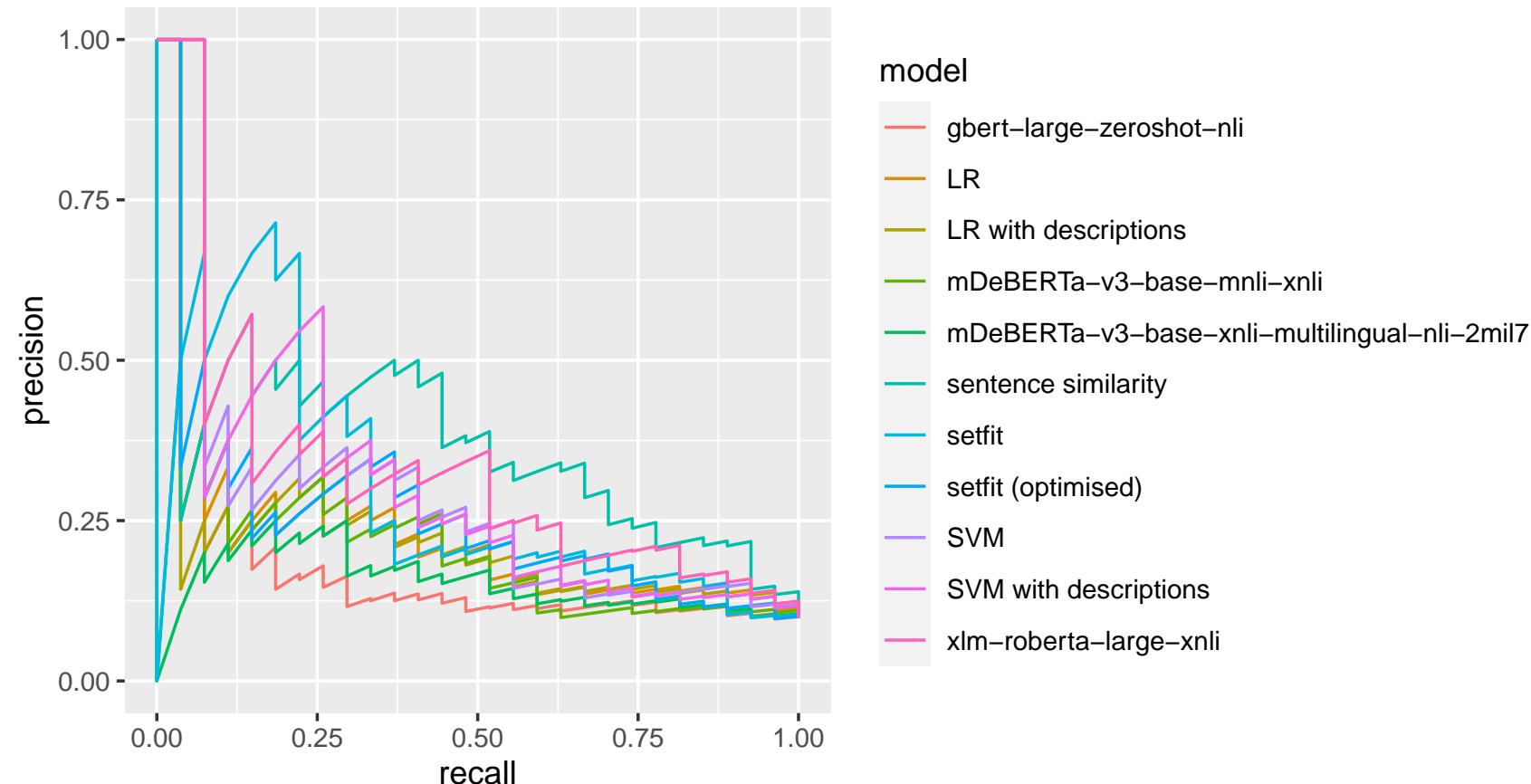
Millenarismus: Am Tag X wird das Volk sich erheben, die Regierung stürzen und zur Rechenschaft ziehen.

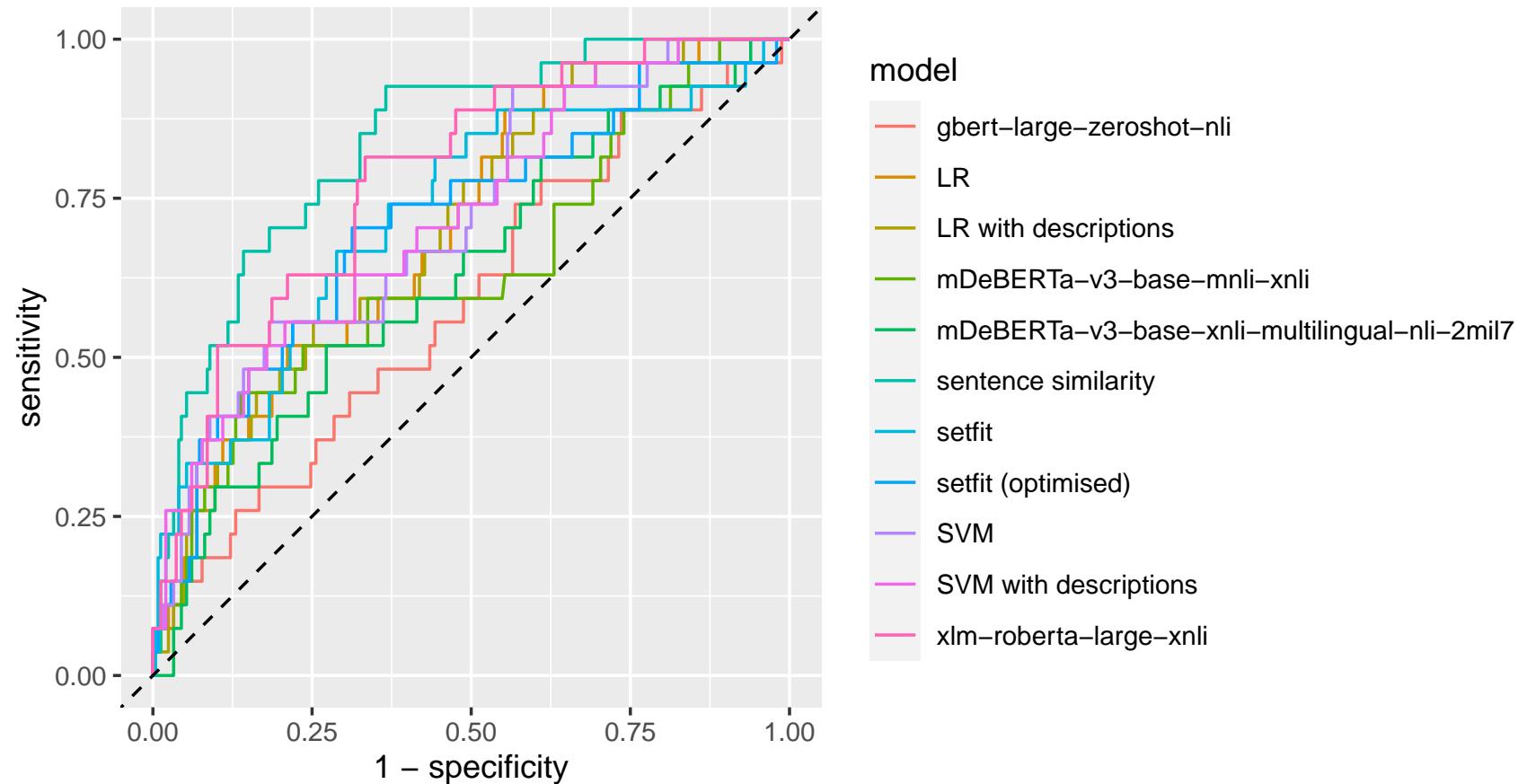
Schlafschafe: Ein großer Teil der Bevölkerung ist gehirngewaschen und hat keine Ahnung, was wirklich in der Welt vor sich geht.

Schlafschafe: Schlafschafe lassen sich von den Medien blenden und erdulden alles wie brave Schafe, anstatt sich zu wehren.

Classify each post according to the labels and descriptions above. Use only these labels. Multiple labels per post are possible.

Provide your output in json format with ids and all applicable labels.





model	ROC-AUC	ROC-AUC (all)
sentence similarity	0.86	
xlm-roberta-large-xnli	0.82	
SVM with descriptions	0.81	0.81
LR with descriptions	0.80	0.80
SetFit (optimised)	0.79	0.78
SVM	0.78	0.78
LR	0.77	0.77
SetFit	0.76	0.75
mDeBERTa-v3-base-mnli-xnli	0.70	
mDeBERTa-v3-base-xnli-multilingual-nli-2mil7	0.69	
gbert-large-zeroshot-nli	0.58	

model	PR-AUC	PR-AUC (all)
SetFit (optimised)	0.34	0.36
SVM with descriptions	0.32	0.34
sentence similarity	0.31	
SetFit	0.31	0.34
LR with descriptions	0.28	0.31
xlm-roberta-large-xnli	0.28	
SVM	0.28	0.31
LR	0.27	0.30
mDeBERTa-v3-base-mnli-xnli	0.21	
mDeBERTa-v3-base-xnli-multilingual-nli-2mil7	0.17	
gbert-large-zeroshot-nli	0.14	

model	F_1 (micro)	F_1 (macro)
Carina	0.76	0.72
Laura	0.71	0.66
ChatGPT4	0.52	0.43
SetFit (optimised)	0.51	0.37
SetFit	0.50	0.36
SVM	0.48	0.33
SVM with descriptions	0.47	0.38
LR with descriptions	0.45	0.35
sentence similarity	0.40	0.32
LR	0.40	0.28
xlm-roberta-large-xnli	0.39	0.31
mDeBERTa-v3-base-xnli-multilingual-nli-2mil7	0.22	0.21
mDeBERTa-v3-base-mnli-xnli	0.20	0.21
gbert-large-zeroshot-nli	0.15	0.15

-
1. Introduction
 2. Telegram corpus
 3. Categorisation and manual annotation
 4. Automatic identification of conspiracy narratives
 5. Conclusion

- overall: difficult task
 - moderate IAA
 - no automatic approach comes close to human performance (yet)

- overall: difficult task
 - moderate IAA
 - no automatic approach comes close to human performance (yet)
- classic ML models competitive (and fast to train)
- ChatGPT4 offers decent zero-shot performance
 - but: non-deterministic output, a pain to set up properly, and expensive
- sentence similarity offers decent zero-shot performance, too
 - plus: easy to use and understand
- NLI zero-shot approaches perform poorly
- SetFit only approach (aside from ChatGPT4) to beat baselines
 - but: hyperparameter optimisation takes a lot of time

- overall: difficult task
 - moderate IAA
 - no automatic approach comes close to human performance (yet)
- classic ML models competitive (and fast to train)
- ChatGPT4 offers decent zero-shot performance
 - but: non-deterministic output, a pain to set up properly, and expensive
- sentence similarity offers decent zero-shot performance, too
 - plus: easy to use and understand
- NLI zero-shot approaches perform poorly
- SetFit only approach (aside from ChatGPT4) to beat baselines
 - but: hyperparameter optimisation takes a lot of time
- outlook:
 - more annotated training data
 - domain adaptation for base LLMs (e.g. Schwurpert: <https://huggingface.co/ausgerechnet/schwurpert>)

Thanks for listening!

Questions!

- Douglas, K. M., J. E. Uscinski, R. M. Sutton, A. Cichocka, T. Nefes, C. S. Ang, and F. Deravi (2019). "Understanding conspiracy theories". In: *Political Psychology* 40.S1, p. 7.
- Infratest dimap (2020). *Corona und die Medien - Eine Studie im Auftrag des NDR – Magazin zapp - Mai 2020*.
- Institut für Demoskopie Allensbach (2022). *Politischer Radikalismus und die Neigung zu Verschwörungstheorien*.
- Kuhn, S. A. K., R. Lieb, D. Freeman, C. Andreou, and T. Zander-Schellenberg (2021). "Coronavirus conspiracy beliefs in the German-speaking general population: endorsement rates and links to reasoning biases and paranoia.". In: *Psychological Medicine*, pp. 1–15.
- Reimers, N. and I. Gurevych (2019). "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks". In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.
- Tunstall, L., N. Reimers, U. E. S. Jo, L. Bates, D. Korat, M. Wasserblat, and O. Pereg (2022). "Efficient Few-Shot Learning Without Prompts". In: *arXiv preprint arXiv:2209.11055*.
- Yin, W., J. Hay, and D. Roth (2019). "Benchmarking zero-shot text classification: Datasets, evaluation and entailment approach". In: *2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019*. Association for Computational Linguistics, pp. 3914–3923.