

Is toxicity towards Italian politicians gendered? A multi-level analysis of hate speech on Twitter during election period



Keywords: political discourse, misogyny, hate speech, Twitter, communities

Executive summary: This study addressed toxicity in tweets targeting Italian politicians during the propaganda campaign preceding Italian snap elections to surface differences in gender. A large sample of replies (163, 544 replies) from the 20 most popular tweets posted by 123 Italian women and 121 men politicians is considered; through network analysis, semantic analysis and natural language processing, trends in the way toxicity is crafted and distributed across communities are identified. Main takeaways are the following: women politicians are targeted more than man politicians, especially through threats or attacks to their identity; women politicians are target of attacks from the same groups of users more than man, suggesting the presence of a misogynistic group of haters; women haters like men haters tend to attack women more than man; “being a woman” is topicalized in toxic speech against women and associated to fearmongering words.



Co-financed by the Connecting Europe
Facility of the European Union

This report was written as part of the Italian Digital Media Observatory project under EC grant agreement INEA/CEF/ICT/A2020/2394428. The information, documentation and figures available in this deliverable were written by the Italian Digital Media Observatory project consortium and do not necessarily reflect the views of the European Commission. The European Commission is not liable for any use that may be made of the information contained herein.

Background

Sadly, it is common ground knowledge that women in politics constitute privileged targets of toxic language which ends up in abusive behaviour. No doubt this trend has been exacerbated by social networks. Such a misogynistic phenomenon is cross-national: Hunt, Evershed and Liu (2016) have, for example, shown that Hillary Clinton and Julia Gillard received twice as much abuse on Twitter compared to male candidates¹, while the platform results to be a toxic place for women MPs in the UK (Dhrodia 2017², Southern and Harmer 2021³). The consequences are easy to imagine, with women pushed to disengage from politics as a place to which they do not belong, or they are not suitable for. Gendered online harassment has turned out to be a major cause for women to censor their political views or abandon platforms (Sobieraj 2020⁴); moreover, when women are in leadership positions and they receive insults and abuses completely distinct from those that affect men (<https://www.esteri.it/wp-content/uploads/2022/09/LUISS-Come-individuare-e-contrastare-operazioni-coordinate-di-disinformazione-in-Italia.pdf>).

To counter the phenomenon of online toxicity, social media platforms such as Twitter have created tailored policies and codes of conducts:

“Hateful conduct: You may not promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease. We also do not allow accounts whose primary purpose is inciting harm towards others on the basis of these categories.

Hateful imagery and display names: You may not use hateful images or symbols in your profile image or profile header. You also may not use your username, display name, or profile bio to engage in abusive behaviour, such as targeted harassment or expressing hate towards a person, group, or protected category.”

(<https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy>)

However, content moderation, as showcased by our research, is still challenging especially when it comes to languages other than English for which automatic content moderation tools are still not sophisticated enough to cope with the sheer amount of toxic language.

¹Hunt, E., N. Evershed, and R. Liu. 2016. “From Julia Gillard to Hillary Clinton: Online Abuse of Politicians around the World.” *The Guardian*, June 27. <https://www.theguardian.com/technology/datablog/ng-interactive/2016/jun/27/from-julia-gillard-to-hillary-clinton-online-abuse-of-politicians-around-the-world>.

²Dhrodia, A. 2017. “Unsocial Media: Tracking Twitter Abuse against Women MPs.” *Amnesty Global Insights*. Accessed 1 July 2019. <https://medium.com/@AmnestyInsights/unsocial-media-tracking-twitter-abuse-against-women-mps-fc28aeca498a>.

³Southern, R., & Harmer, E. (2021). Twitter, incivility and “everyday” gendered othering: An analysis of tweets sent to UK members of parliament. *Social science computer review*, 39(2), 259-275.

⁴Sobieraj, S. (2020). *Gender, Digital Toxicity, and Political Voice Online*.

Data and the Research Design:

To collect a relevant data sample, we focused on Twitter since it constitutes the social media platform conceived as the most used for political campaigning and propaganda. As a result, while every political figure in the Italian landscape has a Twitter account, parties' representation across Facebook, Telegram and Tik Tok is imbalanced. We retrieved through the Twitter API all the tweets published in the time span 20-07-2022 - 22-08-2022 by the official accounts of the 123 women politicians belonging to candidate parties to the snap electoral elections. We matched the accounts with 121 from male politicians selected to obtain a balanced distribution across parties. We then identified the 20 most popular posts for each account according to engagement metrics (total number of likes and retweets) and collected all the public replies per post in compliance with a max cap of 200 replies per post. The composition of the entire dataset is detailed in Table 1 in Appendix. To investigate the relevance of gender among haters, we inferred users' gender from the language used in their tweets/ profile, managing to match 54% of the total number of users of the original dataset (22529 users).

To analyse the data, we followed a two-tiered approach, to i) identify toxic language in the replies targeting politicians across genders and political parties ii) investigate whether toxicity trends significantly differ across targeted-genders as well as haters' gender with a focus on topicalized themes. To achieve i), we leveraged the Perspective API (<https://perspectiveapi.com/>), a machine-learning fuelled classification system that enables the identification of toxic language broadly defined as “rude, disrespectful or unreasonable comment that is likely to make someone leave the discussion”. We run the request to classify each reply according to the available attributes “Insult”, “profanity”, “threat”, “severe toxicity”, “toxicity” and “identity attack”. A description of each of the attributes is available at <https://tinyurl.com/yscscm4y>. We then aggregated the results for the attributes “insult”, “severe toxicity” and “profanity” on one side (toxic_language) and “identity attack” and “threat” on the other side (toxic_content) in order to allow for the observation of differences in the type of lexicon (e.g. sexual insults) as well thematic features used in the attacks (e.g. being incompetent vs. being dishonest).

To answer ii) we adopted social network analysis, a methodology used to monitoring online users' interactions through networks and graph theory (Knoke and Yang, 2019), and semantic analysis. More specifically, we employed state of the art methodologies for community detection and we analysed words' frequencies divided as to part of speech (e.g. adjective, noun etc.) and their word sketches. A word sketch processes the word's collocates (words which co-occur within a given word) and other words in its surroundings. In the network analysis, the user is represented as a node in the graph and relationships among users as edges. The politicians are nodes and by edge with weights equal to the number of users who responded to a couples of them. The key objectives were to observe gendered and political-orientation patterns in the two communities of politicians subject to hatred speech and groups of haters. To test the significance of attested trends we adopted the Kruskal test on both the medians aggregated per account (to look at features related to the individual politicians) and per single tweet.

Results

Women politicians are privileged targets of hate speech, especially when it comes to “Content-toxicity”

The classification of the dataset of 163,544 replies according to the toxicity metrics reveals that women are across the board more frequently the target of toxic behaviour. As visualized in Table 1., the medians of toxicity are higher in correspondence of women regardless the unit of analysis (account/single tweets) and the type of toxicity (language toxicity and content toxicity):

Table 1: Statistical tests for the difference in toxicity received between men and women

	Language Toxicity			Content Toxicity		
	Median M	Median F	p-value	Median M	Median F	p-value
Single Tweets	0.113	0.135	6.7×10^{-68} *	0.292	0.317	5.3×10^{-74} *
Medians aggregated by account	0.095	0.106	0.0524	0.266	0.295	0.0117 *
Means aggregated by accounts	0.189	0.208	0.0264 *	0.317	0.336	0.0041 *

P-values of Kruskal-Wallis tests that compare the level of toxicity of replies to tweets of male and female politicians. The null hypothesis is that the median toxicity addressed at men and women is the same. Zooming into the type of Toxicity, we can see that the gender difference is more significant in correspondence of Content _toxicity. In other words, women are attacked significantly more frequently than men through threats (e.g. example1) and attacks which have to do with their identity (e.g. example 2).

Example 1:

@meb @ItaliaViva Divento volontario solo per annientarvi. Non deve rimanere una sola pietra della vs casa fondata su interessi personali e menzogne. E che il procuratore Durham dia al tuo padrone tutto quello che egli merita: un bel calcio nel culo.

“I volunteer only to annihilate you. There must not be a single stone of your house founded on personal interests and lies. And may DA Durham give your master everything he deserves: a good kick in the ass’

Example 2 :

@MonicaCirinna @msgelmini @mara_carfagna LA LESBICONA CON I SOLDI DENTRO LA CUCCIA DEL CANE, MA NON PROVI VERGOGNA, CRETINA SIETE IL MALE ASSOLUTO DI QUESTA NAZIONE PIDIOTI LADRONI NEL DNA NASCETA LADRI E MORIRETE LADRI E INFAMI. I PRIMI A INSULTARE LE DUE CRETINE POCO DI BUONO ERAVATE VOI E CMQ MARSILIO HA RAGIONE

‘THE BIG LESBIAN WITH MONEY INSIDE THE DOG’S KENNEL, BUT DON’T FEEL SHAME, JUMP YOU ARE THE ABSOLUTE EVIL OF THIS NATION IDIOT THIEVES IN THE DNA BORN THIEVES AND YOU WILL DIE THIEVES AND MISERABLE. MARSILIO IS RIGHT’

The results of the analysis confirm for Italy what attested in other political scenarios, such as Japan (Fuchs & Schäfer 2021)⁵ and USA (e.g. Carlson 2017)⁶. We go one step forward showing that the type of hate speech seems to be tailored to reduce women’s free speech and, anyways, their political participation in leadership position. Threats and identity attacks are in fact likely to have stronger chilling effects than general insults.

Women politicians are the most central as toxicity targets in their communities

The social network analysis has revealed that there are communities emerging from the way replies to the accounts are distributed. However, the communities seem to be clustered around political coalitions rather than gender. This appears to be the case both if we consider all the tweets and toxic tweets. In other words, it seems that political affiliation rather than gender is a key driver when it comes to the creation of online communities. However, if we consider the composition of each network filtering for toxicity (networks of politicians which are target of toxicity), an interesting pattern emerges: the actors which score highest for network centrality are women, as visualised in Fig 1:

⁵Fuchs, T., & Schäfer, F. (2021, October). Normalizing misogyny: hate speech and verbal abuse of female politicians on Japanese Twitter. In Japan forum (Vol. 33, No. 4, pp. 553-579). Routledge.

⁶Carlson, C. R. (2017). Misogynistic Hate Speech and its Chilling Effect on Women’s Free Expression during the 2016 US Presidential Campaign. J. Hate Stud., 14, 97.

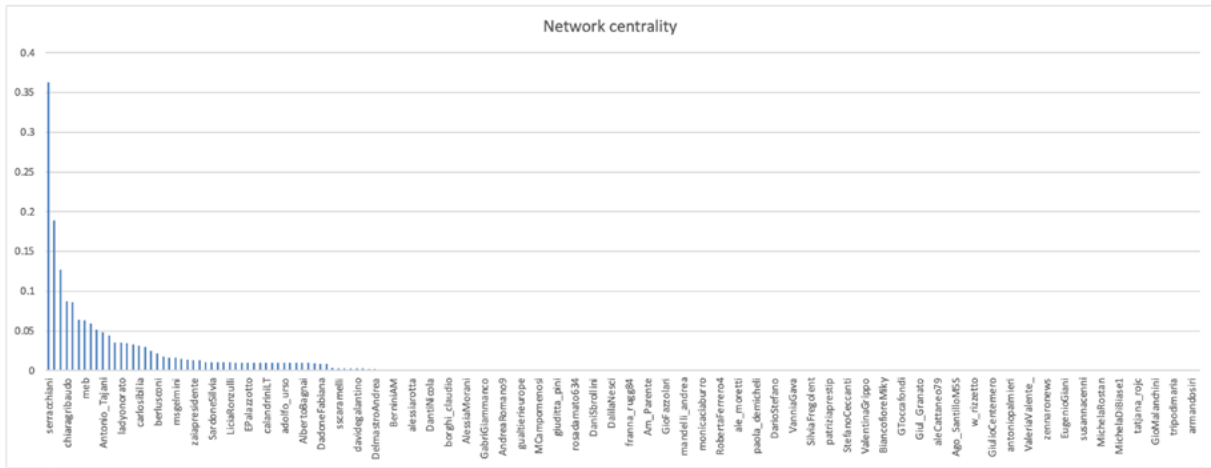


Fig. 2: Networks' centrality scores

This pattern suggests that attacks towards women politicians cater activities from more cohesive groups of users. The 10 accounts which are most central, namely targeted the most by a given group of haters, are the following:

Account	Number of attacks
chiaragribaudo	26
SimonaMalpezzi	26
DSantanche	22
serracchiani	21
bragachiarara	18
TeresaBellanova	17
AlessiaMorani	15
RossomandoPd	15
MonicaCirinna	15
meb	13

Fig.3 Top 10 female politicians attacked by users who attacked more than 5 politicians

Most active haters are also disinformation spreaders

Focusing on the users which interact toxically with the first 5 accounts in Fig 3, such accounts, 13 happened to be especially active (they published 33000 posts overall starting from July); out of these accounts, 12 follow Debora Serracchiani, Teresa Bellanova and the daily national newspaper LaRepubblica. More than by the accounts that they follow, these users are linked by the type of content that they publish which fuels conspiratorial thoughts, with a focus on the two topics of Covid-19 and the Ukraine Russia war (see Appendix for examples). Their perceived social role of revealing hidden truths and calling for a disruption is suggested by the display names on their accounts which contain sentences along the lines of “I liked to think of a revolution taking place. Unfortunately, the strong powers are still there”, “fuck off with your no / pro vax, green pass” or even worrying keywords such as “counter-information”, “red tick pesticides”⁷. Certain images in users' profiles, such as the following, express fearmongering vibes:

⁷“Zecche rosse” (red ticks) is a denigratory phrase used to make reference to people politically aligned with left wing parties



Women politicians are the most hated online both by men and by women haters

To get insights into haters rather than hated, we focused on the gender dataset in relation to the distribution of the toxicity metrics emerging from their replies targeting both men and women politicians.

Table 2: Medians of the different toxicity metrics distributions and p-values from the Kruskal test

	Content Toxicity			Language Toxicity		
	Male figures	Female figures	p-values	Male figures	Female figures	p-values
Male users	0.2947	0.3163	10^{-26}	0.1134	0.1317	$4 \cdot 10^{-24}$
Female users	0.2873	0.3193	$3 \cdot 10^{-18}$	0.1104	0.1423	$3 \cdot 10^{-17}$
p-values	0.0098	0.3430		0.0110	0.3902	

As shown by the values of the medians, there is no substantial difference in how men and women haters address political figures of a given gender. A surprising result appears when we read the table horizontally rather than vertically, focusing on how users of a given gender behave towards women and man politicians: both female and male users tend to address female political figures with higher toxicity levels both in content and in language.

are some party-dependent tendencies: swear terms such as affanculo/vaffanculo, cagare/cagata and cazzata are keywords in toxic replies towards women from the democratic parties. They are frequently accompanied to other, more covert cases of racist hate speech, where reference to ethnic groups is used to signify low social status: e.g. “@pdnetwork Ma andate affanculo , tu co ‘sto caschetto che pari na filippina @chiaragribaudo” (‘But go fuck yourself, you with that bob haircut look like a person from the Philippines @chiaragribaudo’). When it comes to Brothers of Italy, the most frequent accusation is that of being a “fascist”, commonly paired with other insults in a climax: e.g. “@FratelliItalia Buffona, ignorante, pagliaccia, fascista, schifosa, maleodorante, ladra” (‘clown, ignorant, clown, fascist, disgusting, foul-smelling, thief’). As to ItaliaViva, there is a stress on slurs which make metaphoric reference to excrements (e.g. “feccia”, ‘feccia’ and “liquame”, ‘sewage’) and on the idiomatic expression “andare a zappare la terra”(e.g. “@TeresaBellanova Ma andate a zappare, parassiti!”, ‘but go hoeing, pests’), which is used as an insult to mark the incapacity of dealing with complex matters.

Recommendations and implications

From our data analysis it emerges that Twitter content moderation has failed to identify toxic content towards Italian women politicians during the election period. The retrieved content is explicitly condemned by the platform hateful conduct policy since it i) incites violence, ii) it spreads malign stereotypes towards the category of women politicians, regardless of their political orientation, iii) it propagates “non consensual slurs, epithets and sexist tropes”, iv) it spreads hateful imaginary. The reasons underlying struggles for a successful moderation are multiple and they are difficult to pinpoint due to the lack of complete transparency in the design of automatic tools used by Twitter for content moderation. Transparency, thus, appears among our main recommendations:

- Describe and disclose the types of automatic content moderation tools used for tweets written in Italian and for multi-modal content to allow for joint efforts to improve them beyond the pure lexical level
- Reflect upon the category of women in leadership position as one to be explicitly protected
- Consider that hate speech does not only take the form of slurs but it can be covert and expressed through idiomatic expressions which call for language-specific knowledge
- Use the detection of online haters as a proxy for the identification of disinformation and networks of fake news spreaders

Appendix

A_Gusmeroli	10	60	ale_moretti	19	1337
a_viscomi	15	46	Ale_ToddeM5S	20	986
adolfo_urso	12	115	AlessiaMorani	20	2793
aggravivda	3	5	alessiarotta	20	645
Ago_SantilloM5S	5	29	Am_Parente	7	223
Alberto_Cirio	10	45	AnnaAscani	20	1976
AlbertoBagnai	20	1673	annalisabaroni	5	7
ale_villarosa	7	35	ariccardi9	4	4
alebenvenuto	2	4	augustamontarul	20	485
aleCattaneo79	10	356	AvvRossello	3	5
AMorelliMilano	19	1147	AzzolinaLucia	6	917
andrea_cangini	20	2231	BaldinoVittoria	17	319
AndreaRomano9	12	618	BeaLorenzin	20	1614
AngeloCiocca	20	1215	bennyfiorini	6	40
AngeloTofalo	2	8	BerniniAM	8	611
Antonio_Tajani	20	2675	BiancofioreMiky	3	25
AntonioDePoli	8	27	bragachiarra	18	843

antoniopalmieri	15	201	c_appendino	19	1326
armandosiri	8	417	CalabriaTw	13	544
arrigoni_paolo	17	359	carlaruocco1	20	2685
augussori	2	2	caterinabiti	8	103
BarillariDav	20	2550	ceciliadelia	19	821
bartolopietro1	6	99	ChiaraColosimo	7	39
beppe_grillo	9	1539	chiaragribaudo	20	1788
berlusconi	20	4233	ClaudiaPorchiet	5	15
borghi_claudio	20	4228	CollotMarta	20	347
BrunoAstorre	16	131	DadoneFabiana	16	2311
BuompaneG	7	24	DalilaNesci	8	109
calandriniLT	20	681	Dani_Rondinelli	12	20
CandianiStefano	6	20	DaniSbrollini	19	450
CarloCalenda	20	4318	DeborahBergamin	20	658
carlosibilia	20	2709	DonaConzatti	8	160
DalessandroCam	4	18	DonazzanElena	5	14
DamianiDario	1	1	DSantanche	20	4062
DanieleLeodori	2	2	ElenaCarnevali	7	27
DanieleValle3	1	1	ElenaLucchini	11	19
DantiNicola	20	909	EleonoraEvi	20	1208
DarioStefano	11	71	eleonoramattia1	11	23
davidefaraone	20	2080	EmmaPetitti	4	97
davidegalantino	20	335	erica_rivolta	9	77
DelmastroAndrea	20	731	erikastefani71	15	40
Donzelli	5	456	federicadieni	16	259

edorixi	19	384	federicazanella	7	30
Emagorno	14	146	FiammettaModena	8	14
EmilioCarelli	5	195	francesca_flati	5	83
EnricoBorghi1	20	811	FrancescaLaMar	7	10
EnricoLetta	20	4407	franna_rugg84	6	103
enzopresutto_	3	4	FrassinettiP	10	137
EPalazzotto	11	286	GabriGiammanco	14	247
Ettore_Rosato	20	3134	GiorgiaMeloni	20	4279
eugenio_zoffili	10	30	giuditta_pini	8	536
EugenioGiani	19	155	gualminielisa	5	36
fabiano_amati	7	14	ilariafontana5s	3	182
FerriCosimo	5	12	Irene_5s	4	6
fulviomartuscie	3	5	isabellarauti	20	806
galeazzobignami	20	962	isatovaglieri	12	21
gasparripdl	20	1186	itinagli	16	1087
GBenamati	2	4	LaCavandoli	7	42
GiacomoPossamai	6	23	ladyonorato	20	3519
giamma71	12	83	LauraGaravini	19	305
gianlucarizzo46	4	11	lauraravetto	20	419
GianniGiroto	14	214	LiciaRonzulli	17	2958
GiGraziano	4	5	lisanoja	20	1883
GioFazzolari	9	177	MaiorinoM5S	11	72
GioMalanchini	5	9	mantolalla	5	29
giorgiomule	20	778	MaresaBellucci	12	260
giovannidicaro	19	541	maria_spena	7	126

Giul_Granato	20	558	mariaederaM5S	4	52
GiulioCentemero	11	112	Mariettatidei	6	122
GrimoldiPaolo	15	53	MarroccoPatty	2	13
grotondi	20	1897	martabonafoni	7	19
GToccafondi	19	178	MartaLeonori	5	29
gualtierieurope	20	1937	McGadda	19	273
guerini_lorenzo	5	108	meb	20	3828
GuidoDeMartini	20	1509	MichelaDiBiase1	10	65
ignaziocorrao	20	279	MichelaRostan	18	172
igoriezzi	11	83	mirellaliuzzi	5	139
iunioromano	13	104	Mlucialorefice	1	1
ivanscalfarotto	20	2750	monicaciaburro	7	25
leonardo_marras	2	3	MonicaCirinna	20	1655
Luca_Sut	5	19	msgelmini	20	3845
lucafrusone	7	15	paola_demicheli	14	623
lucarizzonervo	7	29	paolabocci	2	3
luigidimaio	6	1364	patriziaprestip	20	337
mandelli_andrea	19	177	Patty_LAbbate	7	115
ManlioDS	7	826	pinapic	6	194
marcodimaio	20	1133	pirroelisa	12	780
MarcoDreosto	17	145	Pres_Casellati	10	181
MarcoRizzoPC	20	4370	PucciarelliLega	14	32
mariopittoni	11	52	raffaellapaita	20	2444
massimobitonci	18	243	rebeccafrassini	3	34
MassimoCasanov3	13	22	renatapolverini	7	28

matteosalvinimi	20	4306	RobertaFerrero4	19	90
MCampomenosi	12	71	robertalombardi	8	103
Michele_Anzaldi	20	1400	robertapinotti	18	382
MicheleGubitosa	20	763	RominaMura	18	190
mircocarloni	1	1	rosadamato634	12	35
mruspandini	5	188	RossomandoPd	8	718
nzingaretti	20	3045	sabrina_decarlo	2	6
PasqualeCiaccia	3	3	Sandra_Savino	7	13
PediciniEu	5	34	sara_moretto	20	384
pfmajorino	20	3075	sarafoscolo	1	2
raffaelebruno	9	61	SardoneSilvia	15	509
Rinaldi_euro	20	2049	serracchiani	20	4079
RobBagnasco	2	3	SilviaFregolent	13	221
roderiu	10	26	SimonaMalpezzi	19	2531
sandrogozi	20	659	stefaniapezzopa	3	24
Sergio_Vaccaro1	3	5	SusannaCeccardi	20	377
sergiopuglia	3	4	susannacenni	7	119
severino_nappi	16	40	TardinoAnnalisa	1	5
SianiPaolo	14	101	tateo_annarita	6	9
silviopaolucci	1	2	tatjana_rojc	5	13
simo_billì	5	6	TeresaBellanova	20	3318
simonebaldelli	19	186	Tiziana_Drago	5	19
sscaramelli	20	1127	v_casa_camera	1	1
StefanoCeccanti	20	400	valecorrado86	4	6
tripodimaria	14	74	ValentinaGrippe	8	17

Vito_DeFilippo	2	2	valeriefedeli	20	1198
w_rizzetto	13	60	ValeriaValente_	8	55
zaiapresidente	20	340	vannaio	5	63
zennaronews	19	334	VanniaGava	6	7
			Vincenza_Lab	3	26
			ylucaselli	20	482
TOT		86311			77233

Example 1:

#Crosetto manda altre armi al pazzo #Zelensky...ahhhhh...gli sono avanzate...le pagate voi...contenti??

#Crosetto sends other weapons to the madman # Zelensky... ahhhhh... they are lef-oversyou pay for them... happy ??

Example 2 :

In agosto SpaceX di Elon Musk ha stretto una partnership con T-Mobile, di proprietà di Deutsche Telekom, che da febbraio è sotto contratto per costruire un'applicazione globale di verifica dei vaccini COVID per l'OMS.

(‘In August Elon Musk’s SpaceX partnered with T-Mobile, owned by Deutsche Telekom, which has been under contract since February to build a global COVID vaccine verification application for WHO.’)

<https://reuters.com/business/healthcare-pharmaceuticals/deutsche-telekom-build-global-covid-vaccine-verification-app-who-2022-02-23/> -

<https://t-mobile.com/news/un-carrier/t-mobile-takes-coverage-above-and-beyond-with-spacex>

“We thank Reset for sharing with us insights and criticism, while discussing and debating the focus of this research. We also thank Wateronmars (<https://www.wateronmars.it/>) for their help in the process of data collection”.

Elena Musi, Lorenzo Federico, Ayoub Mounim e Gianni Riotta

