Supplementary Material

https://osf.io/nuc5r/?view_only=aa7bc4183d8f4aaba3d89826b325884b

Fact-Checkers as New Journalistic Mediators: The Role of Global News Agencies' Verification Units and their Interplay with Tech Platforms

Table 1: Contingency Table of Debunked Themes and News Agencies (Figure 1 in manuscript)

Organizations		Domestic Politics	Economy- Political	Society, Social Issues, Cultural Wars, Zeitgeist	Science, Technology & Environment	Health	International Affairs & Politics	Human Interest	Total
DPA	N	30	6	11	10	32	65	29	183
	%	16.4%	3.3%	6.0%	5.5%	17.5%	35.5%	15.8%	100.0%
	sr	-0.2	0.5	-2.2	-0.2	0.3	-0.5	3.3	
Reuters	N	4	5	51	17	65	152	25	319
	%	1.3%	1.6%	16.0%	5.3%	20.4%	47.6%	7.8%	100.0%
	sr	-6.8	-1.2	2.3	-0.4	1.6	2.9	-0.5	
EFE	N	7	4	10	6	7	39	2	75
	%	9.3%	5.3%	13.3%	8.0%	9.3%	52.0%	2.7%	100.0%
	sr	-1.6	1.4	0.4	0.8	-1.5	2.0	-1.8	
AFP-BR	N	74	7	22	11	15	29	7	165
	%	44.8%	4.2%	13.3%	6.7%	9.1%	17.6%	4.2%	100.0%
	sr	8.8	1.2	0.6	0.5	-2.4	-4.2	-1.9	
AFP-AR	N	9	1	6	4	15	18	6	59
	%	15.3%	1.7%	10.2%	6.8%	25.4%	30.5%	10.2%	100.0%
	sr	-0.3	-0.5	-0.3	0.3	1.7	-0.9	0.4	
AFP-CL	N	21	0	0	2	9	21	6	59
	%	35.6%	0.0%	0.0%	3.4%	15.3%	35.6%	10.2%	100.0%
	sr	3.5	-1.3	-2.6	-0.8	-0.3	-0.3	0.4	
Total	N	145	23	100	50	143	324	75	860
	%	16.9%	2.7%	11.6%	5.8%	16.6%	37.7%	8.7%	100.0%

χ 2=227.182, df=30; α=.001

Table 2: Contingency Table of Types of Falsehoods and International Agencies (Figure 2 in manuscript)

Organiz ations		Satire & Jokes	Fabrica- tions	Sensatio- nalism	Mislea- ding	Conspiracy Theory	Hate Speech	Mistake	Online Scams	Others	Do not apply (true)	Total
DPA	N	8	85	0	44	17	10	1	18	0	0	183
	%	4.4%	46.4%	0.0%	24.0%	9.3%	5.5%	0.5%	9.8%	0.0%	0.0%	100.0%
	sr	-0.8	-3.3	-0.5	1.3	4.4	5.4	-0.2	6.9	-0.5	-0.5	
Reuters	N	34	201	1	71	6	0	4	1	1	0	319
	%	10.7%	63.0%	0.3%	22.3%	1.9%	0.0%	1.3%	0.3%	0.3%	0.0%	100.0%
	sr	3.6	-0.8	1.0	1.0	-1.5	-1.9	1.2	-2.3	1.0	-0.6	
EFE	N	1	60	0	14	0	0	0	0	0	0	75
	%	1.3%	80.0%	0.0%	18.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%
	sr	-1.6	1.4	-0.3	-0.2	-1.6	-0.9	-0.7	-1.3	-0.3	-0.3	
AFP-BR	N	5	134	0	23	1	0	1	0	0	1	165
	%	3.0%	81.2%	0.0%	13.9%	0.6%	0.0%	0.6%	0.0%	0.0%	0.6%	100.0%
	sr	-1.5	2.3	-0.4	-1.7	-1.9	-1.4	-0.1	-1.9	-0.4	1.8	
AFP-AR	N	1	48	0	5	5	0	0	0	0	0	59
	%	1.7%	81.4%	0.0%	8.5%	8.5%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%
	sr	-1.3	1.4	-0.3	-2.0	2.1	-0.8	-0.6	-1.1	-0.3	-0.3	
AFP-CL	N	1	45	0	13	0	0	0	0	0	0	59
	%	1.7%	76.3%	0.0%	22.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%
	sr	-1.3	0.9	-0.3	0.4	-1.4	-0.8	-0.6	-1.1	-0.3	-0.3	
Total	N	50	573	1	170	29	10	6	19	1	1	860
	%	5.8%	66.6%	0.1%	19.8%	3.4%	1.2%	0.7%	2.2%	0.1%	0.1%	100.0%

χ 2= 195.77, df=45; α=.001

Table 3: Monitored channels by international news agencies' fact-checking units

Organizations	Monitored Channels	N	Percent
OPA	Facebook	147	64.8%
	Twitter	20	8.8%
	Blogs (no newsroom structure)	19	8.4%
	Telegram	11	4.8%
	Alternative Partisan Media Websites	7	3.1%
	YouTube	7	3.1%
	TikTok	3	1.3%
	Official Politicians, Parties, Government's Websites	2	0.9%
	Mainstream Media	2	0.9%
	Social Media (without specification)	2	0.9%
	Alternative Social Media (8kun, Gab, Parlet etc)	2	0.9%
	Offline	1	0.4%
	Public Speech & Political Rallies	1	0.4%
	Tabloids	1	0.4%
	Specialists' Platforms	1	0.4%
	Instagram	1	0.4%
	Total	227	100.0%
Reuters	Facebook	263	42.2%
	Twitter	251	40.3%
	Instagram	48	7.7%
	Mainstream Media	11	1.8%
	Alternative Partisan Media Websites	11	1.8%
	Alternative Social Media (8kun, Gab, Parlet etc)	9	1.4%
	Blogs (no newsroom structure)	7	1.1%
	TikTok	7	1.1%
	YouTube	5	0.8%
	Tabloids	3	0.5%
	Reddit	3	0.5%
	Public Speech & Political Rallies	2	0.3%
	Offline	1	0.3%
	Podcasts (Apples, Spotify)	1 1	0.2%
	Telegram	-	0.2%
EFF	Total	623	100.0%
EFE	Twitter	60	38.7%
	Facebook	54	34.8%
	TikTok	16	10.3%
	Telegram	10	6.5%
	Mainstream Media	4	2.6%
	Instagram	3	1.9%
	Public Speech & Political Rallies	2	1.3%
	Alternative Partisan Media Websites	2	1.3%
	WhatsApp	2	1.3%
	Social Media (without specification)	1	0.6%
	YouTube	1	0.6%
	Total	155	100.0%
AFP-BR	Facebook	163	32.3%
	Twitter	142	28.1%
	Instagram	74	14.7%
	TikTok	46	9.1%
	Kwai	33	6.5%
	Telegram	15	3.0%
	YouTube	14	2.8%
	Alternative Partisan Media Websites	7	1.4%
	Mainstream Media	3	0.6%
	WhatsApp	3 3	0.6% 0.6%
	**	3 2	
	Helo		0.4%
	Offline	1	0.2%
	Blogs (no newsrrom structure)	1	0.2%
	Pinterest	1	0.2%
	Total	505	100.0%
AFP-AR	Facebook	59	36.0%
	Twitter	54	32.9%
	Instagram	18	11.0%
	Telegram	15	9.1%
	TikTok	7	4.3%
	Mainstream Media	3	1.8%
	Alternative Partisan Media Websites	3	1.8%
	YouTube	3	1.8%
	Blogs (no newsroom structure)	1	0.6%
	WhatsApp	1	0.6%
	w natsApp Total	164	100.0%
	10(4)	104	100.0%
A ED. CI	E 1 1	=0	4= 00:
AFP-CL	Facebook Twitter	59 48	47.2% 38.4%

Instagram	9	7.2%
TikTok	3	2.4%
YouTube	2	1.6%
Telegram	2	1.6%
Mainstream Media	1	0.8%
Kwai	1	0.8%
Total	125	100.0%

Table 4: Verdicts provided by the fact-checking units of global news agencies.

Organizations	Verdict	N	Percent
DPA	False	79	43.2
	No labels (narrative)	55	30.1
	Indeterminate	25	13.7
	Partially False	21	11.5
	Satire-Jokes	3	1.6
	Total	183	100.0
Reuters	False	164	51.4
	Partially False	122	38.2
	Satire-Jokes	19	6.0
	Indeterminate	7	2.2
	No labels (narrative)	6	1.9
	Partially True	1	0.3
	Total	319	100.0
EFE	False	72	96.0
	Partially False	2	2.7
	Partially True	1	1.3
	Total	75	100.0
AFP-BR	False	150	90.9
	No labels (narrative)	12	7.3
	Indeterminate	2	1.2
	True	1	0.6
	Total	165	100.0
AFP-AR	False	56	94.9
	Partially False	3	5.1
	Total	59	100.0
AFP-CL	False	49	83.1
	Partially False	7	11.9
	Partially True	2	3.4
	Satire-Jokes	1	1.7
	Total	59	100.0

Codebook

• General Instructions

- a. Read the entire verification article.
- b. Save the article as a PDF (not print), giving the identification number: e.g., "253.pdf". The links provided in the articles should also work in the PDF document.
- c. Read it a second time, paying attention to the presence/absence/intensity of the several attributions/variables from the codebook. Highlight the indicators in the PDF file. You must completely understand and internalize the meaning of each variable to code it correctly.
- d. Please give the corresponding numbers in the Excel Sheet. For each story, one expects you to enter a code for each variable, i.e., nothing will be left "empty". Use "99" for cases where "there is nothing to be coded" or if it does not apply.
- e. If a link to the verification article is not working, please try to google the title. If you do not find, please let me know and I will send another one.

Formal Categories

- 1. Identification Number
- 2. Title of the debunking-fact-checking article
- 3. Link for quick identification (URL of the fact-check)
- 4. Date of the Fact-Check.

5. Fact-Check Organizations

- 1. DPA Faktenchecks
- 2. Reuters Fact Check
- 3. EFE Verifica
- 4. AFP Checamos
- 5. AFP Factual
- 6. AFP Chile

Content Categories

6. Genre: Type of Correction or Target of Fact-Checking

1. Debunking of information circulating on social networks

Correction of online disinformation, viral hoaxes, and rumors on social media spread by anonymous sources or bots (viral rumors, images, and videos) in the "anonymous web". Verification of content published by **unofficial** sources or without a clear author. Horizontal fact-checking/debunking. Aim: promote a more truthful information environment in a broader sense.

2. Fact-Checking (Political Statements).

Inspect rhetoric and claims made by political actors and public figures (political representatives, pundits, and even journalists) in the public sphere (established media, Twitter, party conferences, etc.). Vertical fact-checking. Aim: safeguard accountability and correctness in the public debate. In the rare cases of correction of inaccurate media reports (NOT blogs without newsroom structure), consider the correction a fact-check.

3. Other types of articles without adjudication:

Contextualization, explanations (climate basics: C02 explained), meta-discussion on disinformation, compilations of most spread disinformation (without verdict), tutorials, in-depth investigative reporting, data journalism, reportage with forensic elements, reportage about disinformation, background information about a specific discussion, service piece (e.g. what to read for good information on Russian's invasion of Ukraine), opinions etc., without correction process and adjudication. The article presents no verification, fact-checking or debunking process, i.e. there is no verdict of any kind, also no narrative adjudication. Some organizations (BBC) do not work with labels, but the verdict is presented through narrative. In this case, you must code the article in the categories above. STOP THE CODING.

7. Thematic Area of Verification and topic of disinformation.

The thematic area under verification by fact-checking articles: issues, agenda of disinformation. Contexts of false information. The main area and topic are in the foreground of the statement being checked. Usually, it appeared within the headline, subtitle, or lead of a fact-checking article. Only one option is possible. Keep in mind: what triggered the correction articles? What are false information's thematic area and topic (independent of the target)? For false allegations, consider the topic/theme of the attributed claim. For instance, the allegation that Lula (target)

will remove the name of Jesus from the Bible if elected should be coded within the topic of religion (not election). Do not mix context with the theme of false information. Although the disinformation was circulating within an electoral context, the topic of the spread of disinformation is about religion. Search for a subtopic within the main thematic area selected in the previous category. A decisive point for the coding decision is to focus on the information's false part. For instance, consider this article: "The images are not from the murder of a Paraguayan prosecutor but from the movie 'Sundown.' (EFE). The topic of false information here is "crime and delinquency" and not "culture and entertainment." The film image is accurate, and the crime attributed to it is false.

100. Domestic Politics (dis)information/National Issues

200. Economy-Political mis-(dis)information

300. Society, social issues, cultural wars, and zeitgeist

400. Science, Technology, and Environment

500. Health Disinformation

600. International Affairs and Politics

700. Human Interest: Culture & Entertainment, Tourism, Food, Sports, Celebrities, and (false)

Advertisements

900. Others

8. Levels of debunking veracity (code only for debunking)

Oriented on the (Allcott & Gentzkow, 2017; House of Commons, 2018, p. 7). To the best of your ability, what type of false information is the article dealing with? Select the most appropriate one. If the article presents more than one verification, code the first one.

1. 100% Fabrication (Completely made-up content)

The invention of facts. Absurd, ridiculous, baseless, ludicrous, surreal fabrications (100% made-up content with no basis in reality): Fabricated content, news fabrication, and false statement, i.e., completely false content. Here one refers to the invention of complete falsehoods, such as pizzagate or bottle in the shape of a dick distributed in schools, flat earth claims, and insertion of a 5G ship in the brain after vaccination —100% wrong; designed to deceive. The entire message (text + image + audio) has to be created. E.g.:

• (...) "by reaching the age of five, the child becomes the property of the State. It is up to us to decide if a boy will be a girl and vice-versa. It is up to the parents to respectfully accept our decision! We know what is best for the children". Sentence attributed to Fernando Haddad, candidate for the presidency in 2018, in Brazil.

2. Fabricated content based on manipulation

Genuine information/imagery is manipulated to deceive. It refers to digital alterations, i.e. it alters authentic visual and audio content. Non-existent elements are added to photographs or videos to reinforce a message. This type of fabrication distorts <u>truthful</u> information or <u>imagery</u>, for example, deepfakes, fake subtitles or other audio and visual manipulation types. Manipulations and alterations of videos from national broadcasting newspapers and channels should be coded here.

3. Fabricated content based on imposter material.

Impersonation of genuine sources. This type of fabrication impersonates genuine sources by creating a fake page or profile (legitimate news outlets, government agencies and companies), for example, by using the branding of an established news agency and recognized organization to support their (false, fabricated) claims. In this case, the genuine content of established media is not manipulated as above, but the brand's logo and design are incorporated into the (false) message to give it credibility. It is a fabricated message pretending to be a news story published by an established media outlet or an organization's official website. The fabricated message could also pretend to be an "official" Twitter account. Only imitation, no alteration of the original content.

4. Fabrication content based on decontextualization and attribution of context.

False context: content taken out of the original context and used in another so that its meaning changes. Decontextualization (inserting distorted contextual data into authentic content) or fabricating (false) context for a real image. Genuine content is shared with the wrong contextual background or missing context (e.g., authentic images or discourses taken out of contexts, a claim used in a different context than the original). Attributing another (false/distorted) meaning or reasoning to an actual piece of information. E.g., Exhibit an image of an Indonesian

Earthquake to show the suffering of an earthquake in Papua New Guinea or use Kosovo war images as if they were from the Ukraine war. Key word: attribution, creation, invention of contexts

5. Misleading content (Exaggeration, Cherry-Picking Facts, Distortions, omission, flawed methods, misinterpretation of scientific/statistical methods)

Distorting information to favor an individual or issue to fit a particular partisan narrative. For example, presenting a comment as fact, misreading scientific processes and results, incorrectly interpreting statistical data or test results, citing statistics, laws, or facts <u>selectively</u>, misreading official numbers, ignoring others, etc. Fact and information are rhetorically misrepresented or skewed (partisanship). The <u>information is not entirely false</u>, it contains correct data, but it is **exaggerated or twisted** to reinforce an argument and fit a narrative. Important elements are **ignored or mixed with incorrect** data that produce a false impression. ATTENTION: if there is attribution of false meaning/reasoning to a factual piece of information, you probably should code 4 above. **Here there is no creation, but distortion**.

6. False connections and/or poor journalism.

Discrepancies between text, headlines, images, and captions that do not support the content. When headlines, visuals, or captions do not support or correspond to the content (sensationalist headlines used by tabloids and partisan media outlets).

7. Satire, parody, and internet jokes and pranks (humorous content):

Presenting humorous but false stories as accurate. Although not usually categorized as fake news, this may unintentionally fool readers."

8. Branding, advertising, commercial, rent-seeking

Consumer misinformation. Dubious products, goods, and services advertising circulating online. E.g., "Beware of fake Facebook ads for weight loss and diet products. "Lose five pounds in seven days!" - that promises such rapid weight loss success can be untrustworthy or fake.

99. Not applicable (rumor is true or online claim is unverifiable).

9. General types of (false) information oriented and expanded from (Humprecht, 2018).

1. Satire or artistic interventions:

Humorous content that aims for social or political criticism. Readers can mistake it for conventional news reporting (no intention, but potential to harm)

2. All types of fabrication from V24 included (encapsulates complete fabrication, manipulation, imposter and attribution of false context).

Fabrication, manipulation, creation, adulteration of content, the invention of false context. False information, ridiculous claims (pizzagate case, bottle in the shape of a dick distributed in schools, dealers are present during ballot to force people to vote). The keyword to decide is invention. Is there any sort of it in the entire message? If you selected V24 options 1,2,3, or 4, you probably will code this option here!

3. Sensationalism and bad journalism. Combination of misinformation and elements of sensationalism and its online version, clickbait:

It is provocative and aims to attract the audience's attention. The emotional and dramatic elements of a story are emphasized. It can be understood as sensational stories originating from online rumors. Content focuses mostly on non-political sources and presents information sensationalistically (evoke emotion, use extreme circumstances to grab attention, simplify and trivialize complex issues, promote shock value, tabloid presentation). Employment of a sensational narrative. Headlines that provoke curiosity alluring the reader to click to fill the "curiosity gap" (forward-referencing). E.g., "ELECTRIC CARS: THE BIGGEST SCAM THE WHOLE WORLD HAS EVER SEEN!"

4. Misleading information: bias, employment of a partisan narrative, distortions, cherry-picking data, misinterpretation of scientific methods/statistics or employment of flawed methods (similar to V24-5)

One considers here statements lacking important information (omissions), observations based on a partial interpretation of data, or exaggerating a specific aspect of reality (cherry-picking facts). It presents events or issues in a different context or a peculiar manner to fit a specific narrative (eventually ideologically motivated information). Misreading scientific processes and

results, incorrectly interpreting statistical data or test results, citing statistics, laws, or facts <u>selectively</u>, misreading official numbers, ignoring others, etc. Fact and information are rhetorically misrepresented or skewed (partisanship). The <u>information is not entirely false</u>, it contains correct data, but it is **exaggerated or twisted** to reinforce an argument and fit a narrative. Important elements are **ignored or mixed with incorrect** data that produce a false impression. Key word here is **distortion/misinterpretation/misunderstanding** and not creation/invention/false attribution.

5. Combination of Disinformation/Fabrication and Conspiracy Theory¹ (double check footnote 7 before selecting here, no interpretation!)

Information, claims, or statements that explain an event or practice based on the false and **paranoid perception of machinations secrete connections and tactics employed by powerful** people and institutions (e.g., anti-vaccination, 9/11 truth movement, flat-earth beliefs, plans to diminish world population, climate-change denial, QAnon, 5G implantations, etc.). **Machinations must stay in the foreground of the verification; do not look under the lines trying to detect subtle indicators**.

6. Combination of Disinformation/Fabrication and (ethnic, religious) Hate Speech.

The information contains insults, violent threats, hatred, and discrimination. It targets people based on their ethical or national origin, religion, gender, sexual orientation, and political beliefs. E.g., a manipulated image of president Zelensky with the subtitle "The social world cannot be ruled by homosexuals".

- 7. (Genuine) Mistakes (only if clearly stated on the article, no assumptions).
- 8. Online scams.
- 9. Others
- 99. Do not apply if it's true!
- 10. Verdict/Judgement. The fact-checking or debunking outcome according to the organization assessment (not yours)². Search for the conclusion, final verdict given by the organization.
 - 1. True. Confirmation (it is true)
 - 2. Partially true. Relativization tends to the veracity (partially truth, inaccurate but true)
 - **3. Partially false**. Relativization tends to falsity (not quite like this, incomplete, inaccurate, exaggerated, ambiguous, confusing, miscaptioned, misleading, missing context)
 - **4. False.** Refutations (it is false, not right, deceptive, erroneous, this isn't true, not real, manipulated, false context)
 - **5. NO EVIDENCE.** Lack of evidence to reach a veridiction, e.g., contradicting legal interpretations (unverifiable/lack of evidence/inconclusive/debatable)
 - 6. Satire, jokes
 - 7. Mixed verdicts (rare! Do not use without base)
 - **8.** No labels or scales. The organization presented no explicit labels.

Verification is based on evidence and argumentation (narration) but without a straightforward confirmation/refutation of a claim. Pieces of evidence speak for themselves, and the organization does not conclude if the speaker is right or wrong (common for BBC when fact-checking politicians).

- 11. Transparency of practices of information gathering. Are there indicators in the fact-checking article why the verified rumor, statement, or content was selected? In other words, why a particular claim deserves to be checked. Code 0=No/absent and 1+ Yes/present
 - 1. (Social) Media Viralization

Strong resonance on social media regarding likes and comments. Due to viralization/widely disseminated content on social media networks (reach or repetition in different platforms, i.e. cross platform circulation, cross-border disinformation, also checked by several international fact-checking organizations). Increasing topic circulation in the press or on the Internet (e.g., "publication shared

¹ In general, in those cases, fact-checkers can only provide context by presenting the facts beyond the narrative, but not proving or disproving the narrative itself since the power of conspiracies theories relies outside the meticulous adherence to the facts.

² Pay attention with BBC (generally no designed labels, but verdict within the text). For example: **Presented claim:** Liz Truss has referred to this polling before in Parliament, telling MPs on 22 May: "An overwhelming proportion of people in Northern Ireland - 78% - agreed that the protocol needed to change in polling conducted in December 2021." "<u>It is simply not true</u> to say that a majority of people in Northern Ireland support the protocol." **The fact-checker verdict:** You get to 78% if you add up the people who thought it should be scrapped, those who thought it needed serious reform and those who thought it would be acceptable with some adjustments. But that figure encompasses a very wide range of opinions on the protocol (BBC Reality Check, 15.06.2022)

more than 5 thousand times" or "a widely shared photograph"). Just code if there is explicit mention do dissemination.

2. The prominence of (disinformation) sources, spreader or speaker

Due to the prominence of the disinformation agent or source of utterance, i.e., who issues the claim (e.g., nationally prominent personalities or media outlets). Code if the source of utterance or rumor has at least a regional prominence. One should also consider traditional media outlets here as sources of false information/replication and distribution.

3. The prominence of the target of disinformation

Due to the prominence of the target of false information. Code if the affected actor has higher (national, international) prominence or if the false information affects a national/international company, organization, or institution. One understands "targets" as individuals, groups of actors, or institutions accused of specific problems or actions attributed to their names (NOT the direct audience who the source of false information is trying to disinform). Please pay considerable attention to this difference. We also consider "secondary" targets here. For instance, actors could also be blamed/accused of an action, e.g. media organizations that had their layout imitated. In this case, one attributes the media organization for reporting something they did not. In case of doubt, try to respond to the question — who would write a statement to assure the public that the information is false?

4. User requests or denounce on social media.

The audience requested the verification; readers have sent a suggestion of dubious information (only if explicitly said, e.g. "we have been asked"). Large numbers of users request asking to evaluate these stories.

5. Timely (Hot Topic). Media Events. What is happening right now (2022) in the news!

It refers to what is happening in the current regional political debate or is under discussion in the media (select if the original claim/audio-visuals was taken from the mass media, public speech in Parliament, or the verification is related to a world event – Covid, Ukraine War, Queen's death, etc.). The article can also mention current events indirectly: "The misinformation spread after Europe has been hit by a heatwave that has caused several forest fires".

6. Social Relevance/Impact/Public Interest

Importance of verification for the public interest and agenda. Issues of general interest (that affect and impact as many people as possible). E.g., false information about election campaigns, approval of new constitutions, law debates, policy making process, **public decision making process**, policy making, wars, health issues or verifications to protect citizens against online scams.