Mediterranean Digital Media Observatory

D5.1

MedDMO Fact-Checks Archive Analysis for monitoring the media landscape in Cyprus, Greece, and Malta

2nd Report of platform practices and national authorities support 2024
Part A

Project Title	MedDMO
Project Number	101083756
Thematic Priority	DIGITAL-2021-TRUST-01-EDMO
Start of Project	1 December 2022
Duration	30 months

Deliverable title	2nd Report of platform practices and national authorities support –2024 – Part A
Deliverable number	D5.1
Deliverable version	V1.0
Contractual Date of delivery	n/a
Actual Date of delivery	n/a
Nature of deliverable	Report
Dissemination level	Public
Partner Responsible	CUT
Author(s)	Pantelitsa Leonidou, Nikos Salamanos, Paschalia Spyridou, Michael Sirivianos
Reviewer(s)	CERTH, UCY
EC Project Officer	Mark Vella Muscat

Abstract	This is the MedDMO annual report detailing the results of monitoring online platform practices to support the national authorities of Cyprus, Greece, and Malta in combating active disinformation campaigns. This report presents findings from our analysis of the MedDMO fact-check archive (1 January–30 June 2024). We mapped the channels—social media platforms, local news outlets, and other online or offline channels—through which disinformation circulated in Cyprus, Greece, and Malta, and evaluated how each platform's moderation tools (warning labels, removals, context panels, and community annotations) were deployed in response. These insights will both inform our assessment of the platforms' Code of

	Practice on Disinformation reports and strengthen ongoing monitoring of the media landscape across all three countries.
Keywords	Monitoring Social media platforms, disinformation campaign, fact-checking, Code of Practice on Disinformation

Copyright

● Revision History

Version	Date	Modified By	Comments
V0.1	01/11/2024	CUT	First Draft Table of Contents
V0.2	08/11/2024	CUT	Final Table of Contents
V0.3	02/12/2024	CUT	First Draft
V0.4	18/01/2025	CUT	Second Draft
V0.5	25/02/2025	CUT	Final Second Draft
V0.6	25/04/2025	CUT	Final Second Draft Review
V0.7	05/05/2025	AUTH, UCY	Review
V1.0	12/05/2025	CUT	Final Document

● Glossary

Abbreviation	Meaning
CoP	Code of Practice on Disinformation
VLOPs	Very Large Online Platforms
VLOSEs	Very Large Online Search Engines
AFP	Agence France-Presse
EH	Ellinika Hoaxes
ToM	Times of Malta
FCC	Fact-Check Cyprus
UoM	University of Malta

Table of Contents

Index of Figures

Index of Tables

• Executive Summary

This report examines over 1,200 instances of false or misleading claims circulating in Cyprus, Greece, and Malta between January and June 2024, as documented by MedDMO's fact-checking partners. It aims to measure how these flagged claims were treated by very large online platforms and by online news outlets that shared them. For each fact-checked item, we tracked platform responses—whether a warning label was applied, content removed, contextual information provided, or no visible action taken—as well as any subsequent corrections or clarifications issued by the news sites themselves. Our findings show significant variation in how platforms handle identical content and reveal that news outlets seldom follow up with corrections, leaving many debunked claims accessible without context. These patterns underscore gaps in the practical enforcement of the EU Code of Practice on Disinformation and inform MedDMO's ongoing Task 5.1 monitoring efforts.

1 Introduction

Disinformation monitoring requires not only the identification of false claims but also rigorous tracking of how those claims are managed once exposed. The EU Code of Practice on Disinformation (CoP) **[European Commission, 2022]** is a voluntary, self-regulatory framework through which major online platforms pledge to curb the spread of false or misleading content. Its core objectives are to empower users with clear labelling and context, boost the visibility of authoritative sources, remove or demote harmful disinformation, and ensure greater transparency and collaboration with independent fact-checkers. Under the EU Code of Practice on Disinformation, signatory platforms submit periodic self-assessment reports that outline the policies, technical mechanisms, and partnerships they have established to counter false content, from fact-check labels and removal workflows to demotion algorithms and contextual info panels. Each report also provides quantitative metrics (e.g., numbers of labelled posts, takedowns, demotions, and context exposures) for the covered period, helping stakeholders assess the real-world impact of these measures. To validate and complement those self-reports—and to monitor platform practices more broadly—we analysed the MedDMO fact-check archive (1 January–30 June 2024), tracing how false claims circulated across Cyprus, Greece, and Malta and evaluating exactly how each platform's moderation tools were applied to content flagged as disinformation. For the purposes of this report, we adopt the EU's original 2018 Code of Practice definition of "**Disinformation**" as follows: *"Verifiably false or misleading information which, cumulatively, (a) Is created, presented and disseminated for economic gain or to intentionally deceive the public; and (b) May cause public harm, intended as threats to democratic political and policymaking processes as well as public goods such as the protection of EU citizens' health, the environment or security."* **[European Commission 2018]**.

MedDMO's fact-checking consortium publishes fact-checking articles for disinformation spreading in the three countries, Cyprus, Greece, and Malta. These fact-checking articles often mention where the false claims were spread (which platform, online news outlets, online or offline spaces) and give the links to the archived or original content carrying the false claim. In our effort to monitor the platforms and generally other channels that contribute to the false information circulation and their practices to combat disinformation and assess the platforms' Code of Practice on disinformation reports, we decided to analyse the information we can find from the MedDMO fact-checks archive. We have compiled a dataset of 1,400+ links to content with verified falsehoods circulating in Cyprus, Greece, and Malta. This report focuses exclusively on that dataset, mapping each debunked claim to the visible actions—or inactions—taken by four signatories of the EU Code of Practice (Meta, YouTube, TikTok, and X) and by online news outlets or other media used to circulate disinformation that often serve as "sources of information" that are mentioned on the social media posts for online audiences. By analysing which posts were labelled, removed, received additional context, or left unaltered, we provide an evidence-based view of platform compliance and editorial accountability.

The report is structured as follows. First, we outline the methodology used to analyse the MedDMO fact-check archive, detailing our inclusion criteria for fact-check articles, as well as platform and publisher tracking procedures. In Section 3, we aggregate and analyse fact-checks for Greece and Cyprus, treating them as a combined corpus due to extensive overlap in language and circulating narratives, highlighting how each platform and online news outlet responded to shared disinformation content. Section 4 then turns to Malta, presenting the treatment of locally false claims by the same set of online platforms, news outlets, and other channels where false claims were circulating. Finally, the report closes with a summary of our key findings and recommendations.

2   A Methodology for Leveraging the MedDMO Fact-Check Archive in Platform's Practices and Media Landscape Monitoring

The objective of this analysis is to examine the nature, dissemination, and platform-level response to disinformation identified by fact-checking organisations participating in the MedDMO initiative across three countries. Specifically, the study seeks to understand how false claims were distributed across different channels, including social media platforms, messaging services, online news outlets, and other media formats, as well as to assess the extent to which these platforms engaged in moderation practices—such as content labelling, removal, or referencing fact-checks—to mitigate the spread of false information.

To support this investigation, we collected and analysed fact-checks from the MedDMO Fact-Check archive[1], covering the period from January 1st to June 30th, 2024[2]. The selected fact-checks were reviewed to identify the type and topic of disinformation, the platform or channel through which it spread, and any moderation actions taken by platforms, where identifiable. By systematically examining these dimensions, the analysis aims to contribute insights into the evolving dynamics of disinformation in the Mediterranean region and the responsiveness of digital platforms in addressing it. The results of the analysis of MedDMO fact-checks for the same period in 2023 are available in MedDMO's 1st Report of platform practices and national authorities support 2023 (for Cyprus, Greece & Malta) **[Leonidou et al., 2023].**

MedDMO Fact-Check archive has fact-checks in Greek and English. The fact-checks in Greek refer to disinformation that spreads in Greece and/or Cyprus. At the same time, fact-checks in English may refer to Malta, Greece, and/or Cyprus.

From the **MedDMO Fact-Checks archive,** we have the following information:

1. Fact-check Title
2. Fact-check Article Link
3. Author (and corresponding MedDMO fact-checking organisation that conducted the fact-check)
4. Category of fact-checked information (i.e., Health, Politics, Technology, etc.)
5. Date of fact-checking article publication

Through manual examination of each fact-checking article, we extracted the following data:

1. **Country** (Greece/Cyprus, Malta)
2. **Topic** (Israel Hamas war, Russia/Ukraine war, EU, Climate, Health & Covid-19, 5G, Migration, National elections, LGBTQ+, Other)
3. **Sources** where false claims were detected by fact-checkers included online platforms, news outlets, websites, TV and radio broadcasts, public speeches, and other media formats. Website domains were recorded when applicable.
4. **Links** to the content that contains false claims (if available)
5. **Post-fact-checking treatment** of the content by the source (disinformation label, link to fact-checking article, availability of the content (if it was removed)

---

[1] https://meddmo.eu/fact-checking/archives/

[2] *1st of January to 30th of June 2024 is the period covered in the Signatories of the Code of Practice on Disinformation reports No.4. that MedDMO will publish a report on the assessment of the reports of META, Google and TikTok soon.*

Following the manual examination of the fact-checks, we developed an extended version of the MedDMO Fact-Checks archive for the specified period. This enriched collection, referred to as the Fact-Checks Archive Extended Dataset, includes additional metadata and annotations derived from our analysis. For clarity, we will refer to this collection as "the dataset" throughout the remainder of this report.

Identifying the Source of Disinformation

To determine the medium through which disinformation was disseminated in each fact-checking article, we applied a manual review process based on the following criteria:

- Screenshots or embedded media within the article displaying the false claim on a specific platform (e.g., Facebook, TikTok).
- Direct links to the original or archived instances of the disinformation.
- Explicit references in the article text identify the platform or source of the claim.

In cases where a platform was mentioned or shown via a screenshot, we conducted targeted searches on the respective platform to retrieve the link to the content containing the false claim and to assess whether any moderation actions, such as labelling, removal, or fact-checking tags, had been applied. However, we did not carry out broader, non-specific searches for the claims across platforms, due to the time-intensive nature of such an approach. This methodology ensured that all attributions to platforms or sources were evidence-based and directly traceable to specific elements within the fact-checking articles.

Determining Whether Content Was Labelled or Linked to Fact-Check by the Platform

Each platform adopts a distinct strategy for addressing false information through a mix of labels, warning screens, removals, and contextual tools. In our analysis, we evaluated the platforms' practices to combat disinformation as reported by the platforms' CoP reports No.4. Meta's services (Facebook and Instagram) apply disinformation labels or warning screens to posts flagged by accredited fact-checking partners and link directly to the relevant fact-check articles for added context; any content judged to carry harmful disinformation, such as claims that could undermine democratic processes or public health, is removed outright. In January 2025, Meta announced that beginning in the United States, it would end its collaboration with third-party fact-checking organisations, who had been independently rating claim accuracy, enabling Meta to apply warning labels and demote misleading content, and would instead roll out a Community Notes system similar to X's approach [**Kaplan 2025, BBC 2025**].

Google's video platform, YouTube, takes a contrasting route by focusing almost exclusively on removing videos that violate its disinformation policies, which target content with demonstrably harmful false claims and do not affix "disinformation" labels to any remaining videos. TikTok likewise prioritises the removal of content that breaches its harmful-disinformation rules, but it also introduces an "Unverified content" label when moderator reviews or partner fact-checks are inconclusive; however, it does not apply explicit "false information" tags, choosing instead to pull down policy-violating material. Finally, X (formerly Twitter) relies on its community-driven context tool, Community Notes, to surface additional information: any user can contribute a note on a post they believe is misleading, and those notes become visible only when a diverse group of contributors agrees they are helpful. Unlike the other platforms, X neither removes content for false claims nor attaches "false information" labels; instead, it depends on community-sourced annotations to provide nuance and context.

To assess whether the identified content with false claims in our dataset had been labelled by the platform (see examples of platforms' techniques to label content in Figure 1 and Figure 2), we examined active and publicly accessible content whenever possible. This included cases where fact-checking articles provided direct links to the original content, or where related posts were identified through platform-specific search features. In these instances, we verified whether the content was accompanied by a disinformation label, a warning notice, or any form of platform intervention indicating the claim was false.

We also specifically checked whether the platform included a reference or direct link to a fact-checking article that debunked the claim, either as part of a moderation label or embedded within the content interface. In the case of online news articles or websites, we considered it a form of labelling when the article explicitly stated that a claim was false, such as including phrases like *"the statement that [X] is true is fake news"*—even if no technical or visual moderation label was present. These textual disclaimers were treated as an editorial form of disinformation recognition.
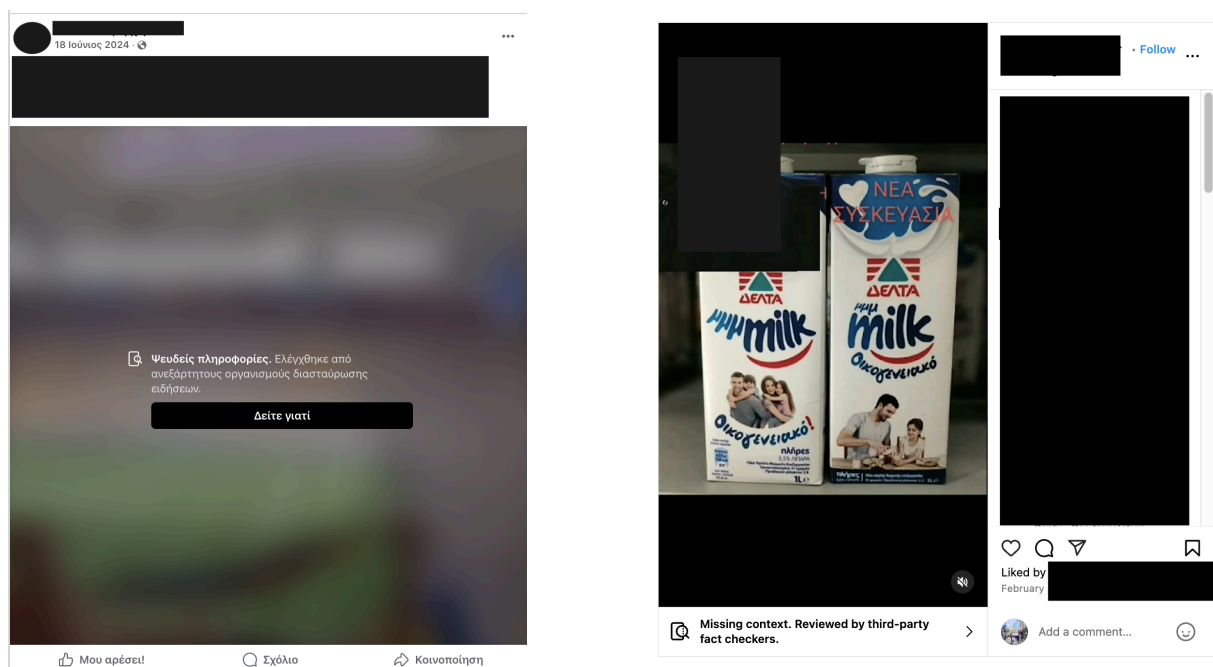

Figure 1: Examples of Meta applying labels to content with false information to empower users (left: Facebook, right: Instagram)
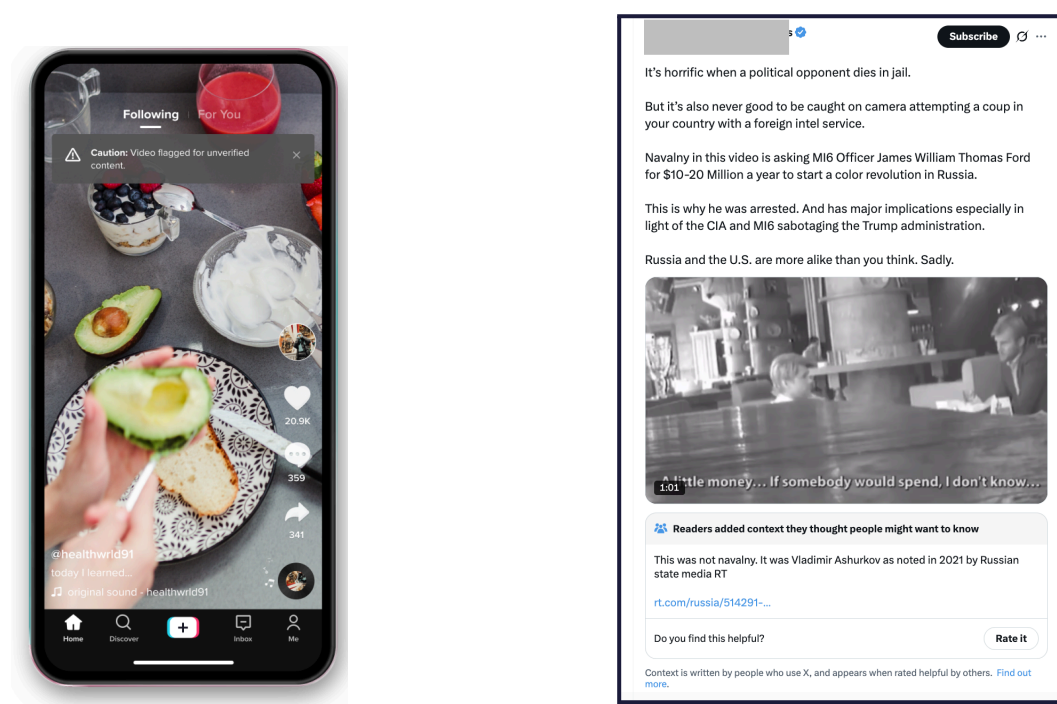
Figure 2: Examples of TikTok and X mechanisms to empower users. TikTok's label on unverified content[3] on the left and X's Community Notes on the right

Determining Whether Content Was Removed

In cases where the content was **no longer accessible on the platform**—as determined through direct or archived links provided in the fact-checking articles—we considered the content to have been **removed**, either by the platform, by the user, or as a result of **modified privacy settings** that made the post no longer publicly visible. This classification was based solely on the inaccessibility of the original content at the time of review.

The extended MedDMO Fact-Checks archive Dataset:

In our dataset, we analysed 150 fact-check articles covering Greece and Cyprus, 138 of them in Greek, 12 in English, and 34 fact-checking articles covering Malta (in English). There are **1226 links** to online content (content from social media platforms or news outlets' websites) that contain false information. Below we present the results of our analysis separately for Malta, and Greece, and Cyprus together.

---

[3] Image source: https://newsroom.tiktok.com/en-us/new-prompts-to-help-people-consider-before-they-share

3 Analysis of Greece and Cyprus Fact-Checks

As a result of the common language in Greece and Cyprus (Greek), it is very difficult to distinguish between false claims that are spread in the two countries. Usually, false claims spread in Greece also reach the Cypriots and vice versa. Due to this fact, we present the results of this analysis for Cyprus and Greece together.
Our dataset includes fact-checks coming from three different fact-checking organisations, Agence France-Presse (AFP), Ellinika Hoaxes (EH), and Fact-Check Cyprus (FCC), only from the period of January 1st to June 30th, 2024. In total, there are 150 fact-checks; the number of fact-checks per fact-checking organisation is depicted in Figure 3.

Greece & Cyprus Fact Checks (in Greek & English) 1 Jan - 30 Jun 2024
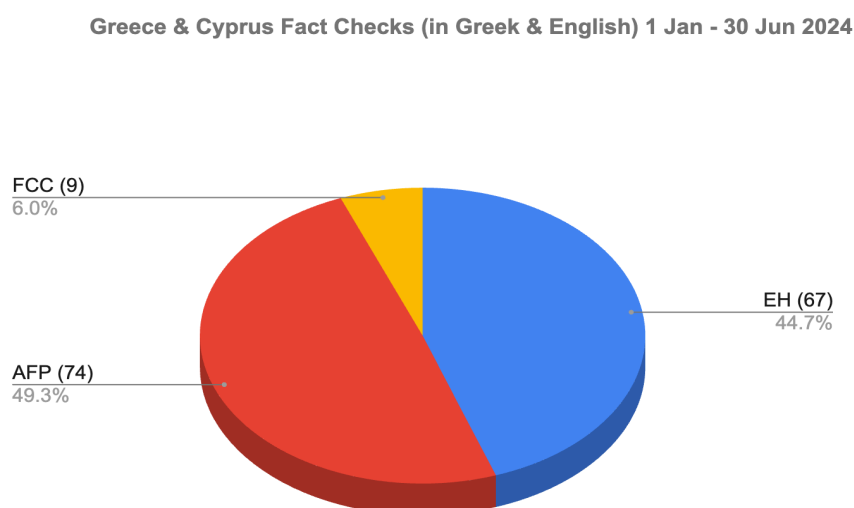
Figure 3: The proportion of Fact-Checks published by each of the three fact-checking organizations (Fact-Check Cyprus, Ellinika Hoaxes, and Agence France-Presse) covering Greece and Cyprus in our dataset for the period of January 1st to June 30th, 2024.

The fact-check articles' topics in our dataset are presented in Figure 4. It is important to note that the fact-checks can have more than one topic, i.e., Health, Covid-19. A great number of fact-checks are relevant to Society and Politics. Several claims related to COVID-19, and other health-related topics were spread during this period. Furthermore, false claims relevant to the Russia and Ukraine war, LGBTQ+, and migration are part of this dataset since those topics' enhanced interest during this period in Cyprus and Greece. Claims relevant to science, technology, environment, climate, and economy are also included.

Greece/Cyprus | Fact Checks Topics
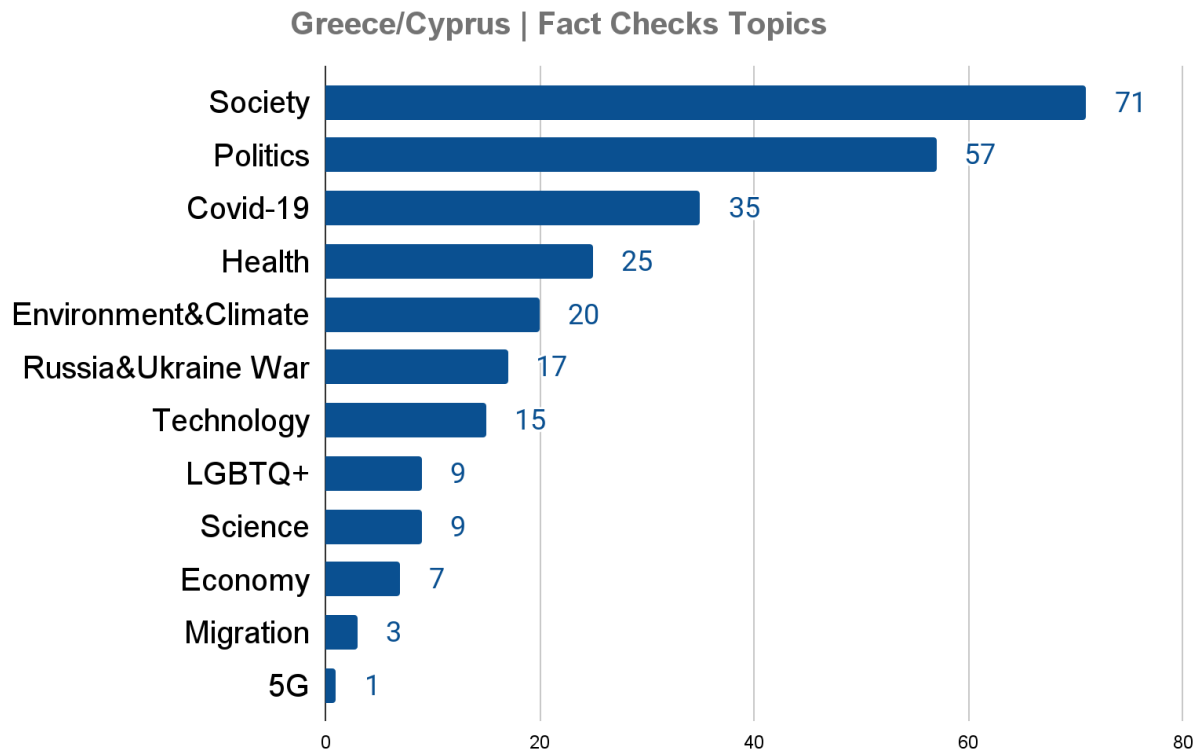
Figure 4: Fact-Check Topics for Greece and Cyprus for the period of this analysis

False Claims in Cyprus/Greece and Social Media Platforms:

Fact-checks often include information, either in their text or accompanying visuals, indicating where the false claims were detected. Based on our analysis, we present in Figure 5 how many times each platform was referenced in fact-checks as a source where relevant false claims were spreading. A 91% of fact-checks in our dataset mentioned that the false claim was identified in Facebook, and 42% reported false claims found in X (prior Twitter). There are a few references to Instagram, YouTube, and TikTok. There is also a high percentage, 64%, of references to other online news outlets in Greece and Cyprus or platforms that we do not cover in our analysis (i.e., Telegram, LinkedIn, Threads) that are referred to as "Other" in Figure 5. Each fact-check may refer to multiple platforms spreading the false claims (see Figure 6). "Facebook and Other," and "Facebook, Other and X" are the two most frequent combinations of platforms.

Greece/Cyprus | # Platform Mentions in Fact-Checks Related to False Claims

Platform	Percentage
Facebook (127)	93%
Other (87)	64%
X (58)	42%
Instagram (7)	5%
YouTube (6)	4%
TikTok (4)	3%
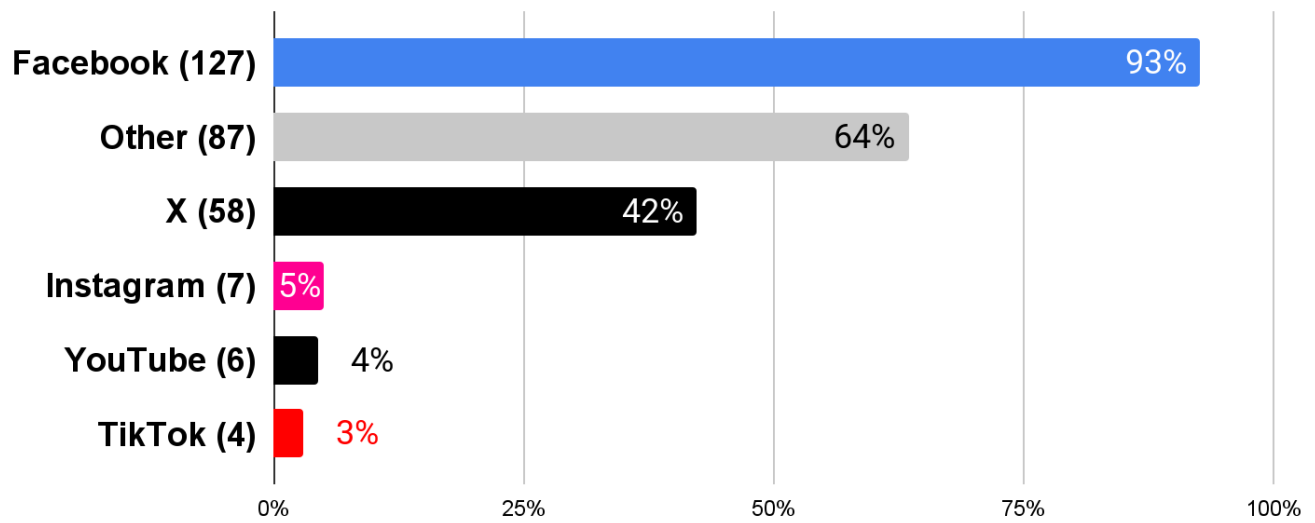
Figure 5: Platforms mentioned in Greece and Cyprus' Fact-Checks

*Other refers to online news outlets, online platforms that are not covered in this survey, and other media

Combination of Platforms Mentions in Fact Checks Related to False Claims

Combination	Percentage
Facebook, Other (39)	28.47%
Facebook, Other, X (prior Twitter) (31)	22.63%
Facebook (31)	22.63%
Facebook, X (prior Twitter) (13)	9.49%
Other, X (prior Twitter) (4)	2.92%
Other (3)	2.19%
Facebook, Instagram, X (prior Twitter) (3)	2.19%
Facebook, Other, X (prior Twitter), YouTube (2)	1.46%
X (prior Twitter) (2)	1.46%
Facebook, Other, YouTube (2)	1.46%
Facebook, Other, TikTok, X (prior Twitter), YouTube (1)	0.73%
Facebook, Instagram, Other, TikTok (1)	0.73%
Facebook, Other, TikTok, X (prior Twitter) (1)	0.73%
Other, YouTube (1)	0.73%
Facebook, Instagram (1)	0.73%
Facebook, Instagram, Other, X (prior Twitter) (1)	0.73%
Facebook, TikTok (1)	0.73%

Figure 6: Combination of Platforms' mentions in Greece and Cyprus' Fact-Checks

In total, we found **1242** instances containing false claims from the fact-checks in our dataset. ~59% of these instances, 729 instances, come from Facebook. There are 80 instances from X, eight from Instagram, and fewer from other platforms such as YouTube and TikTok (see Figure 7).
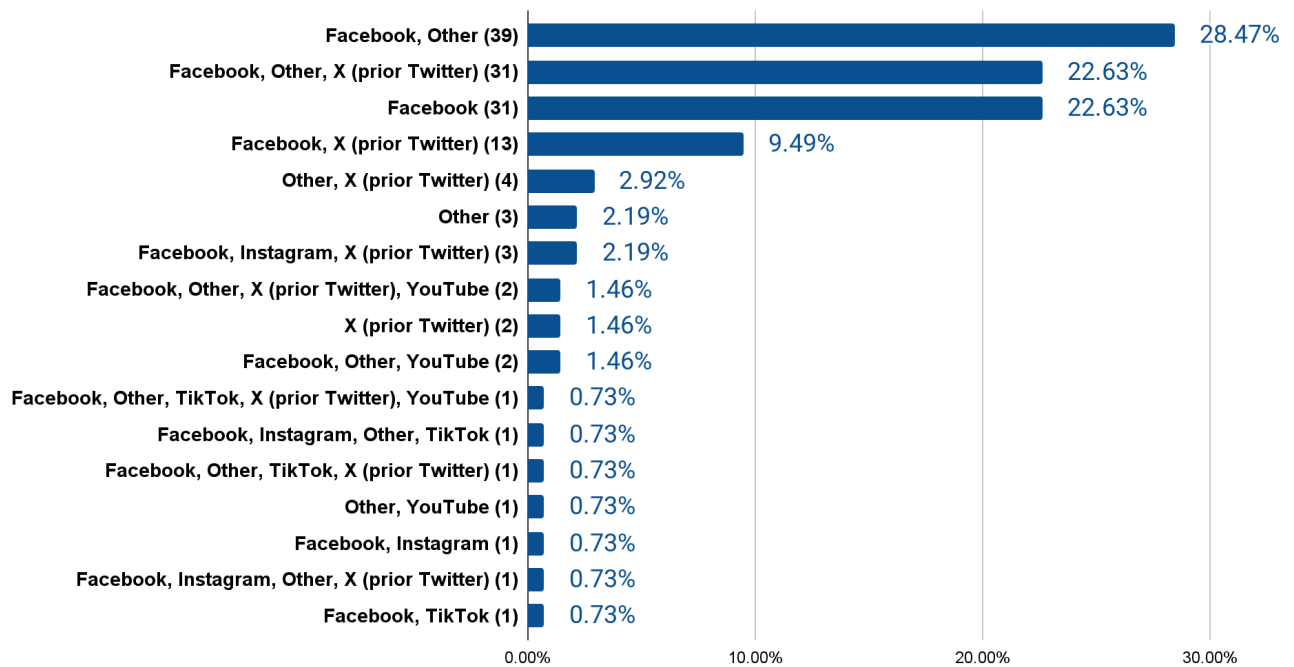
Greece/Cyprus | # Instances containing false claims per platform

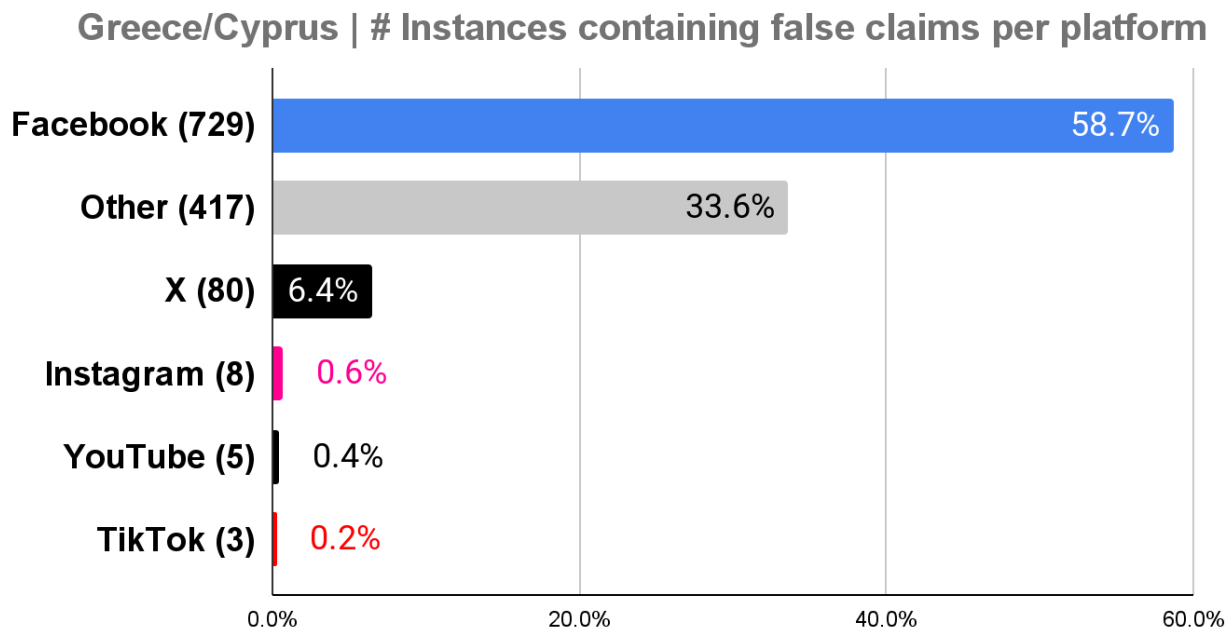| Platform	Percentage
Facebook (729)	58.7%
Other (417)	33.6%
X (80)	6.4%
Instagram (8)	0.6%
YouTube (5)	0.4%
TikTok (3)	0.2%

Figure 7: Number of instances containing false claims per platform found in Greece and Cyprus' Fact-Checks

Other than the aforementioned platforms, there were instances of content with false claims from Telegram (5), Threads (1), LinkedIn (1), and Rumble (1).

Platforms Handling Content with False Claims

Each entry of the dataset, i.e., instance containing a false claim, is categorised by whether it was labelled as disinformation, linked to a fact-check article, or no longer publicly accessible. Importantly, disappearance from public view ("Removed") may result from various user-side actions (e.g., post deletion, account suspension, or privacy setting changes) and should not always be interpreted as platform-driven moderation. See the summary of the results of this analysis in Table 1.

Facebook had the largest number of flagged posts, with 55.6% labelled as disinformation and linked to fact-checks. However, only 5.5% of the entries were no longer publicly accessible. Facebook (and Instagram) applies labels, giving context by adding a link to fact-checks, and downranking based on third-party fact-checking, rarely removing content unless it violates specific harm-related policies. This aligns with the observed pattern of high labelling but limited disappearance of posts. Instagram showed similar trends to Facebook, with 11.1% labelled and 33.3% of content no longer publicly available.

X (prior Twitter) displayed a minimal intervention pattern: only 4.2% of posts were "labelled", and just 3.2% became inaccessible. While X maintains policies against civic disinformation, its enforcement relies primarily on user-driven Community Notes. So, labelled in the case of X's instances means displaying a community note. In practice, few instances received labels or were removed, consistent with its reduced reliance on centralised moderation during this period.

TikTok had very few entries, and none of these three entries received a disinformation label or a link to a fact-checking article.

YouTube had five entries, none labelled, and all remaining visible. YouTube addresses disinformation primarily through policy-based removals (e.g., health or election disinformation) and non-targeted information panels. The lack of labelling and continued visibility of flagged content was expected since YouTube moderation policies are not focused on labelling and removing content with false claims.

Table 1: Platforms' moderation applied to instances containing false claims

Platform (total number of instances)	# Instances (%) Labelled as False	# Instances (%) Links to Fact-Check	# Instances (%) Removed from/ Not Publicly Accessible Platform
Facebook (729)	Yes: 418 (57.3%) No: 255 (35.0%) NA: 56 (7.7%)	Yes: 418 (57.3%) No: 255 (35.0%) NA: 56 (7.7%)	Yes: 41 (5.6%) No: 670 (91.9%) NA: 18 (2.5%)
X* (80) *Community notes*	Yes: 4 (5.0%) No: 66 (82.5%) NA: 10 (12.5%)	Yes: 0 (0.0%) No: 70 (87.5%) NA: 10 (12.5%)	Yes: 3 (3.8%) No: 70 (87.5%) NA: 7 (8.8%)
Instagram (8)	Yes: 1 (12.5%) No: 3 (37.5%) NA: 4 (50.0%)	Yes: 1 (12.5%) No: 3 (37.5%) NA: 4 (50.0%)	Yes: 3 (37.5%) No: 4 (50.0%) NA: 1 (12.5%)
YouTube (5)	Yes: 0 (0%) No: 5 (100%) NA: 0 (0%)	Yes: 0 (0%) No: 5 (100%) NA: 0 (0%)	Yes: 0 (0%) No: 5 (100%) NA: 0 (0%)
TikTok (3)	Yes: 0 (0%) No: 1 (33.3%) NA: 2 (66.7%)	Yes: 0 (0%) No: 1 (33.3%) NA: 2 (66.7%)	Yes: 1 (33.3%) No: 1 (33.3%) NA: 1 (33.3%)

*NA: cases where we lacked access to the content and therefore could not draw any conclusions.

The case of Facebook's false information labels and links to fact-checking articles:
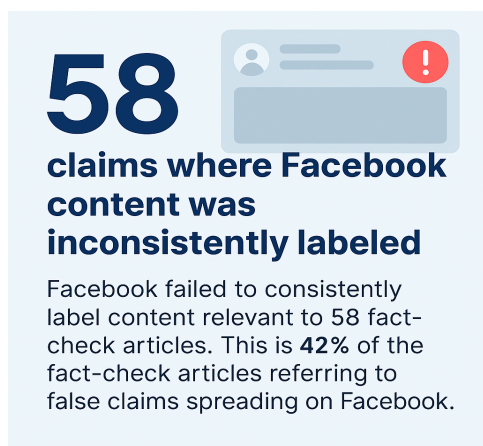

Figure 8: Meta (Facebook)'s cases with inconsistent labelling/moderation

From our analysis, 127 fact-check articles referenced false claims that had been disseminated via Facebook content. Among these, as explained in Figure 8, at least 58 fact-checks/claims were linked to multiple Facebook posts where the same false claim appeared. In those 58 cases, we found that labelling was applied inconsistently: some posts received a disinformation label while others did not, despite referring to the same debunked claim (see Figure 9). While we cannot confirm consistent labelling across all remaining fact-checks, these 58 clearly illustrate that Facebook's application of disinformation labels is not uniformly enforced, even within identical claim contexts.
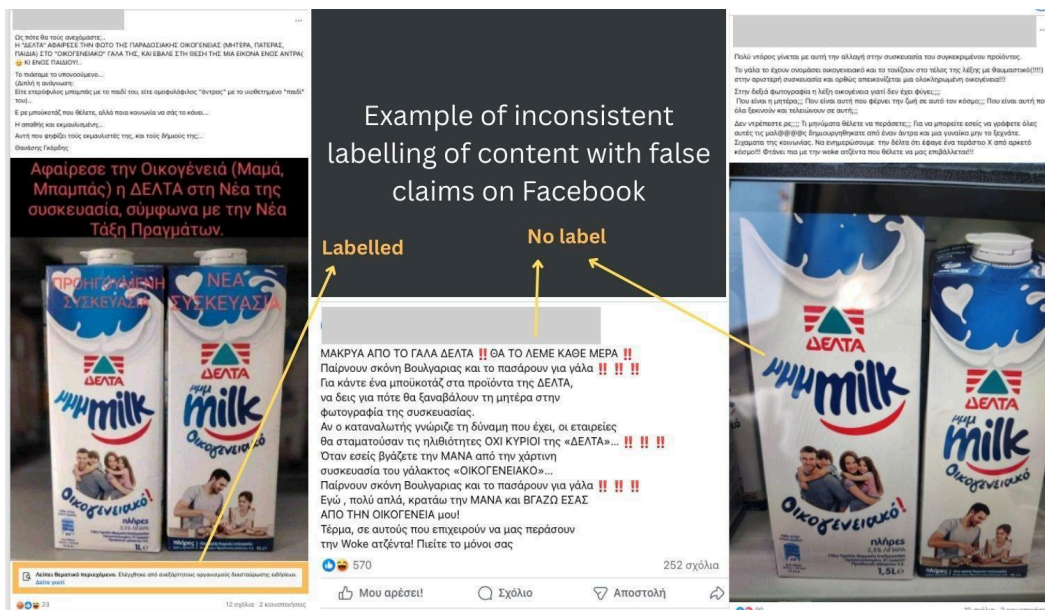

Figure 9: Example of Facebook's inconsistency in labelling content with identical false claims

Greek News Websites that spread false claims through their published articles in the period of January-June 2024:

In our dataset, there are 417 instances with the value "Other" as the platform that spread the false claim. These instances are coming from Greek news websites (406), other platforms that are not covered in our analysis (i.e., Telegram, Threads, LinkedIn), or do not mention a specific platform. From the analysis of the MedDMO Fact-Checks archive in the period of January-June 2024, we found from our data **235 unique Greek news websites/media** that published articles containing false claims[4]. Figure 10 shows the word cloud of the websites' names/domains and demonstrates the websites in our data that published articles with false claims; the font size of the names/domains increases if the websites publish a greater number of false claims.

---

[4] The websites listed have been identified in fact-checking articles as having published false or misleading information. Our listing is based solely on MedDMO fact-check sources and does not imply that these websites consistently or intentionally share false claims. This list is not exhaustive, and the presence or absence of any website does not indicate a blanket judgment about its overall credibility or intent. Our goal is to reflect documented cases of disinformation as flagged by fact-checkers — not to make final claims about the reliability or motives of these sources.

Figure 10: Word cloud[5] of Greek and Cypriot Online News Outlets mentioned in the Fact-Checks for the period under analysis

Only 20 out of the 406 articles (~5%) with false claims were corrected (i.e., altered the text or media, or mentioned the claim was false). More specifically, only five (~1.2%) mentioned that the claim was false, and only three (~0.75%) had a link to the fact-checking articles. In our analysis, we found 34 articles (16.5%) that we could not find online, and we assumed that they had been removed by the site admin. A ~84% (341) of the articles with false claims are still publicly available with no correction. For 12 instances, we were not able to reach the website (i.e., the website was down, the domain was for sale, we got the "forbidden" error response, etc.). In Figure 11, we list all the websites found in our dataset, categorised by the number of articles that contain false claims. The figure also illustrates whether a site was unreachable, corrected at least one article, or removed at least one article that contains disinformation.

---

[5] The Word Cloud was generated with the online tool from https://www.wordclouds.com/

News websites sharing false claims
(found in MedDMO Fact Checks Archive during 1st of January to 30th of June 2024)

30-5 TIMES

pronews.gr (25)
anazitiseis.gr (12)
oparlapipas.gr (8) ◆
press-gr.com (7) ◆
amazonios.net (6)
choratouaxoritou.gr (6)
katohika.gr (6)
makeleio.gr (6)
triklopodia.gr (6) ◆
attikanea.info (5)
karditsastakra.com (5)
newsbreak.gr (5)
sinomosiologos.blogspot.com (5)

4-2 TIMES

anixneuseis.gr (4)
bankingnews.gr (4)
hellasultras.gr (4)
korakas-news.gr (4) ◇
newsit.gr (4)
primenews.press (4)
taxalia.blogspot.com (4)
apagoreuetai-group.blogspot.com (3) ◇
epilekta.com (3)
europost.gr (3)
greeknewsondemand.com (3)
happenednow.gr (3)
iskra.gr (3)
megatv.com (3)
newsbomb.gr (3)
odysseiatv.blogspot.com (3) *
olympia.gr (3)
aftodioikisi.gr (2)
alphatv.gr (2)
cnn.gr (2)
cretalive.gr (2)

cyprustimes.com (2) ◇
dekeleianews.gr (2)
diaforetiko.gr (2)
dimosio.gr (2)
dimpenews.com (2)
efenpress.gr (2)
egerssi.gr (2)
enimerotiko.gr (2)
ethnos.gr (2)
greece24.gr (2) ◇
iapokalipsi.gr (2)
ieidiseis.gr (2)
indicator.gr (2)
kapa-news.gr (2)
kefaloniapress.gr (2)
kinima-ypervasi.gr (2)
kourdistoportocali.com (2)
larissanet.gr (2)
lifo.gr (2)
magazinenews.gr (2) ◇
meaculpa.gr (2)

molwnlave.net (2)
neakriti.gr (2)
newslife.gr (2) ◇
oxafies.com (2) ◇
patrisnews.com (2)
pelop.gr (2)
pentapostagma.gr (2)
ptolemaidanews.gr (2)
romioitispolis.gr (2)
skai.gr (2)
star.gr (2)
thebest.gr (2)
topontiki.gr (2)
tzafnews.gr (2)
vimaorthodoxias.gr (2)

* The site cannot be reached　　　**underlined text:** Sites that corrected their articles at least once　　　◇ site removed at least one article with false claims

ONCE

24h.com.cy
94fm.gr
adieksodos.gr
aegeanews.gr
agonaskritis.gr
agrinio24.gr
agriniosite.gr
alfavita.gr
alldaynews.gr
alphanews.live
amfipolinews.blogspot.com ◇
amianow.gr
amina-politiki.gr
anatropi.substack.com
ant1live.com
antenna.gr
antipliroforisi.blogspot.com ◇
athanasios12.blogspot.com
athensmagazine.gr
avgi.gr
candiadoc.gr
childrenshealthdefense.eu
city.sigmalive.com
conserva.gr
coollife.gr
crash.gr ◇
cyprusnews.eu ◇
cyprusnews.live
daddy-cool.gr
daynight.gr
defencenet.gr
difernews.gr
differentnews.gr ◇
discovercreta.gr ◇

grtimes.gr
huffingtonpost.gr
huli.gr *
iefimerida.gr
ihunt.gr
in2life.gr
irafina.gr
iribeyond.com
iskra.gr
kirkinews.gr
kosmodromio.gr
koutipandoras.gr
ksipnistere.com
lamiareport.gr
lawandorder.gr
libre.gr
loutrakiodusseas.blogspot.com
lumi-news.gr
madata.gr
magnesianews.gr ◇
makthes.gr
maleviziotis.gr
medlabgr.blogspot.com
mikropragmata.lifo.gr ◇
mykosmos.gr
nafpaktos-online.blogspot.com
naftemporiki.gr
nea-news.gr
net-time.gr
newpost.gr
news.makedonias.gr
news.meteolive.gr *
news247.gr *
newsbase.gr

◇

newsbeast.gr
newsfire.gr
newshub.gr
newskamatero.eu
newslarisa.gr
newspepper.gr
newspistol.gr
newsthessaloniki.gr
nikolaosanaximandros.gr
nikosantoniadis.tv *
npress.gr
oimaskespeftoun.blogspot.com
ond.gr ◇
onvolos.gr ◇
orinadervenoxoria.com
parallaximag.gr
parapolitika.gr
patrasevents.gr
patriotiko.gr
patris.gr
pellain.com
periodista.gr
pierianews.gr
pineiosnews.gr
planetwebradio.gr
play-news.gr ◇
pointmedia.gr ◇
pontosnews.gr
press-gr
protothema.gr
radiogamma.gr
radiosecret.gr ◇
redskywarning.blogspot.com
refreshnews.gr
reporter.com.cy

reporter24.gr ◇
revolution.gr
RIK Cyprus
romios.online
salaminionvima.gr *
schizas.com
scinews.eu
sfedona.gr
simeiakairwn.wordpress.com
spoudazo.gr
startmediacorfu.gr
strategist.cy
synmag.gr
syrostoday.gr
tanea.gr
taxidromos.gr
tideon.org
tilestwra.com
tomsnews.io *
tovima.gr
trelogiannis.blogspot.com
trikalaweb.com ◇
trikalazoom.gr
tromaktiko.gr
vimaiteas.blogspot.com
voicenews.gr ◇
volosday.gr
wwwlefkaslive.blogspot.com

Figure 11: Online news outlets in Greece and Cyprus, ranked by the number of times they were cited in fact-checks, with markers indicating whether their content underwent any moderation actions.

4   Analysis of Malta Fact-Checks

For the period of January 1st to June 30th, 2024, there are 34 fact-checks for Malta. In our analysis, we excluded one fact-check since it was a summary of false claims shared in 2023, which is out of the period this analysis focuses on. The fact-checks for Malta in MedDMO are produced by the Times of Malta (ToM) and the University of Malta (UoM). In the specific period under study, only one fact-check was produced by the UoM (see Figure 12). Times of Malta was the first Maltese organisation conducting fact-checks under the MedDMO project. The University of Malta later began producing such articles to support the Times' efforts. It is important to note that the two organisations do not have any collaboration with VLOPs and VLOSEs to fact-check claims that are distributed on their platforms. Meta, TikTok, and Google CoP reports (covering the period of January 1st to June 30th, 2024) do not explicitly mention any collaboration with any organisation covering Malta or the Maltese language.

Malta Fact Checks | 1 Jan - 30 Jun 2024

University of Malta (1)
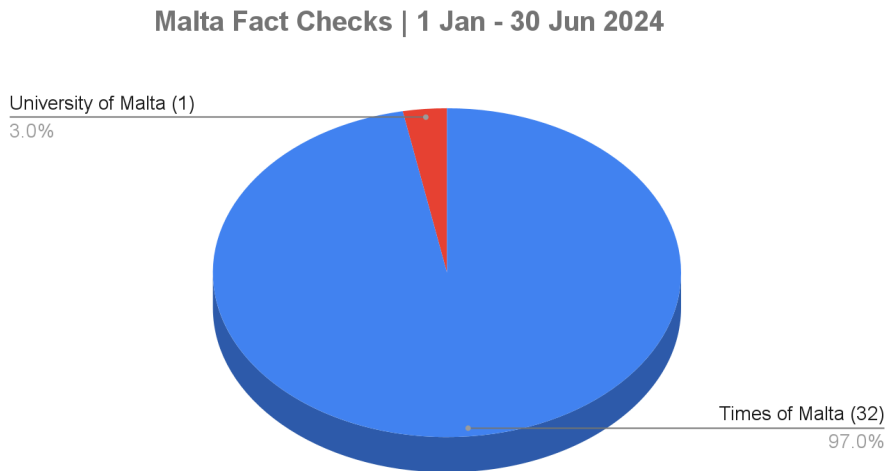3.0%

Times of Malta (32)
97.0%

Figure 12: The proportion of Fact-Checks published by each of the two fact-checking organizations covering Malta in our dataset for the period of January 1st to June 30th, 2024.

The fact-check topics in our dataset are shown in Figure 13. As with the case of Greece and Cyprus, it is important to note that a single fact-check may be associated with more than one topic. A significant number of fact-checks are related to Politics, which emerges as the most prominent category. This is followed by Society and Economy, indicating a focus on political discourse, social issues, and financial matters during the period in question. A number of fact-checks also addressed online scams, AI-generated content, and manipulated media, reflecting growing concerns about digital disinformation and deceptive practices online.

Fewer fact-checks were observed on topics such as Technology, Health, Environment and Climate, and Migration, suggesting either lower disinformation activity in these areas or less fact-checking focus. Nonetheless, these topics are still part of the dataset and reflect the broader range of narratives addressed in Malta. Topics like migration and climate, while less frequent, are of regional interest and could become more prominent depending on the socio-political context.
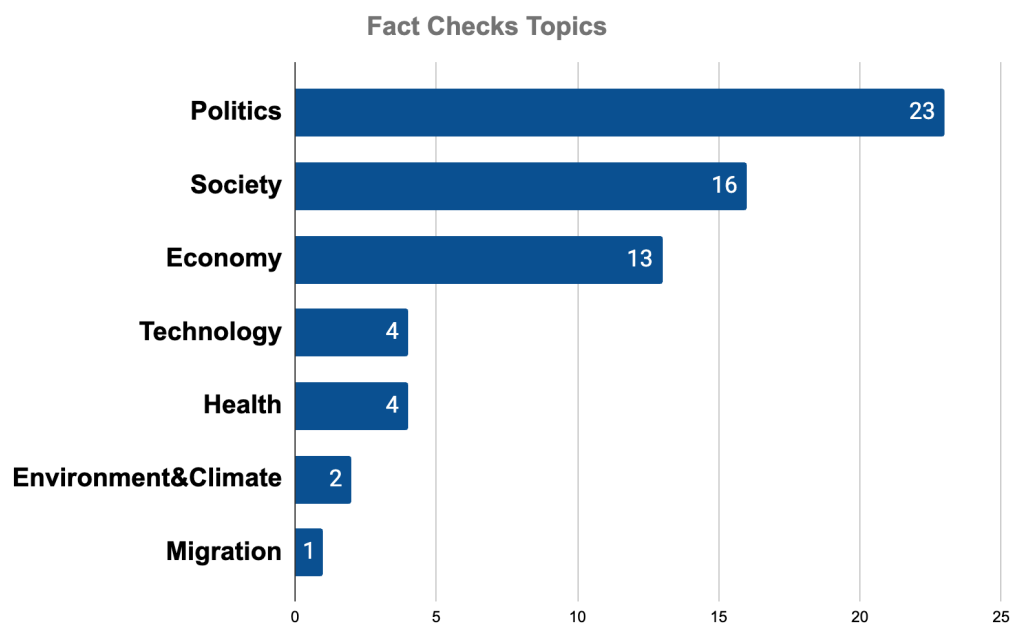

Figure 13: Fact-Check Topics for Malta for the period of this analysis

False Claims circulating in Malta and Social Media Platforms:

Fact-checks often include references—either in the body of the text or in accompanying visuals—indicating where the false claims were initially identified. Figure 14 presents the number of times each platform was mentioned in Malta-based fact-checks as a source of disinformation. Facebook is the most frequently mentioned platform, appearing in 33.33% of the fact-checks.

There are also mentions of Instagram (6.06%), YouTube (3.03%), and TikTok (3.03%), though these are much less frequent. A significant portion of the fact-checks—66.67%—fall under the "Other" category. In the case of Malta, this includes a broad range of sources such as political speeches (e.g., by the Prime Minister or Opposition leader), WhatsApp conversations, online news outlets, public statements from organisations, or, in some cases, no specific source being identified in our analysis of the fact-checks.

Platform Mentions in Fact-Checks Related to False Claims

Facebook (11) — 33.33%

Other (22) — 66.67%

Instagram (2) — 6.06%

YouTube (1) — 3.03%

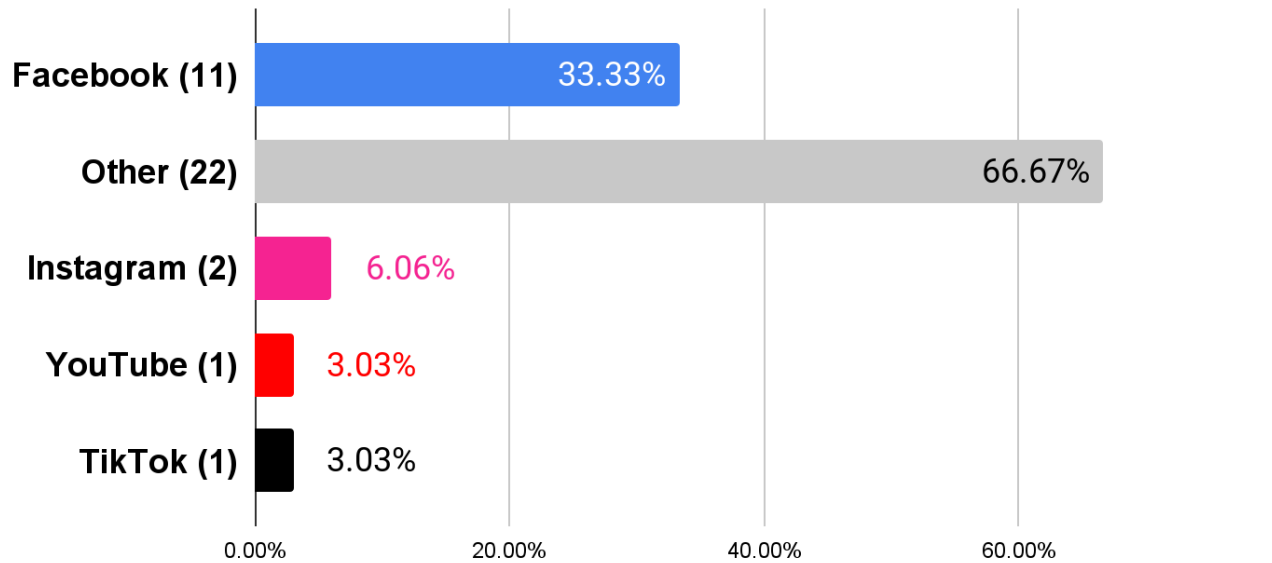**TikTok (1)** — 3.03%

0.00%   20.00%   40.00%   60.00%

Figure 14: Platforms mentioned in Malta's Fact-Checks

In the Malta fact-checks included in our dataset for the specific period, direct links to online content such as social media posts or news articles were generally absent (see Figure 15). As a result, most platform mentions in the fact-checks did not include direct URLs to the original instances of disinformation, making it difficult to determine whether the content had been removed, labelled, or linked to a fact-checking article by the platforms themselves.

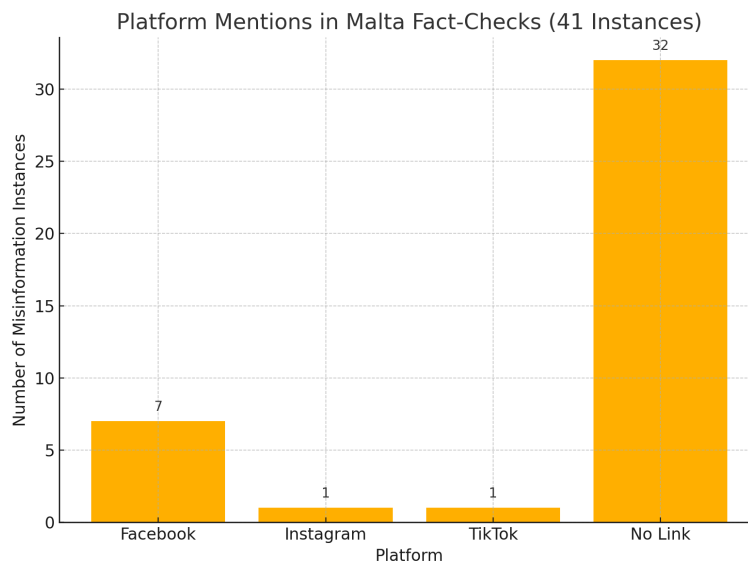Platform Mentions in Malta Fact-Checks (41 Instances)

Figure 15: Number of instances containing false claims per platform found in Malta's Fact-Checks

Out of 41 distinct instances of disinformation, seven included Facebook links, while one was linked to Instagram and one to TikTok. Among these, just one post was flagged with a disinformation label and linked to a

fact-checking article, published by Open.online **[Open Online 2024],** a fact-checking organisation based outside of Malta. This article addressed a recurring false claim that had circulated across multiple countries. Similarly, the instances of false claims identified on TikTok and Instagram did not receive any form of platform labelling, removal, or links to fact-checking resources.

Despite the limited platform intervention, the false claims observed in Malta exhibited high levels of user engagement, including large numbers of likes, comments, shares, and views. This underscores the broad reach of such content and highlights the urgent need for social media platforms to actively label and contextualise disinformation to help users identify false narratives.

Within the MedDMO framework, fact-checking activities are carried out by the Times of Malta and the University of Malta, focusing on debunking false claims circulating locally. However, these organisations are not part of any third-party fact-checking programmes, such as those operated by Meta or TikTok. This means that platforms do not incorporate their findings into official disinformation response mechanisms, leaving content targeting the Maltese public largely unmoderated by platform-based fact-checking systems.
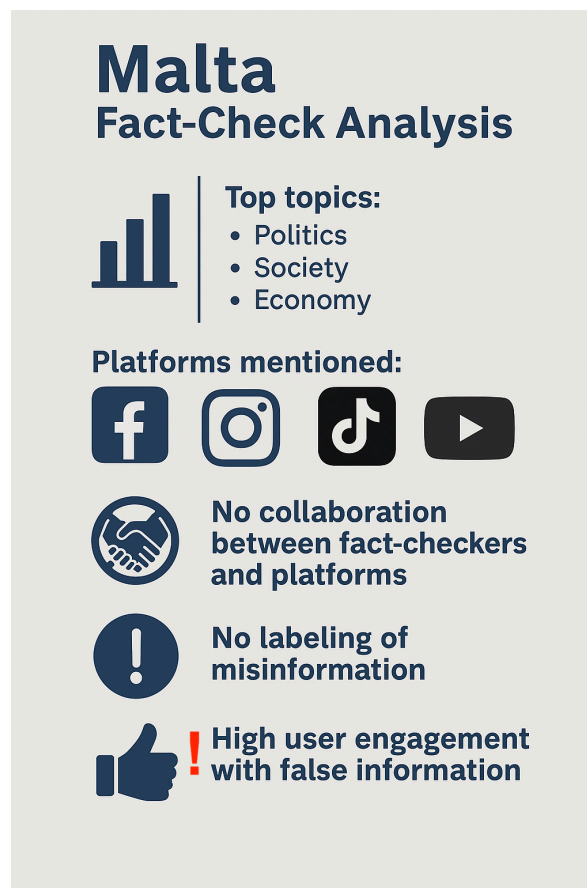

Figure 16: Findings summary of Malta's Fact-Checks analysis

The analysis of fact-checks in Malta reveals a limited but telling snapshot of the disinformation landscape. Most fact-checks focused on political, societal, and economic issues, with Facebook being the most frequently referenced platform, followed by Instagram, TikTok, and YouTube. However, the majority of false claims lacked

direct links, making it difficult to assess platform responses. Crucially, no systematic labelling or moderation of false content was observed, even in posts that achieved high levels of user engagement, highlighting a gap in platform accountability. Additionally, AI-generated content and online scams were part of the disinformation ecosystem.

Despite the efforts of local actors such as the Times of Malta and the University of Malta, there is no formal collaboration between these organisations and major platforms like Meta or TikTok. This absence significantly limits the visibility and impact of local fact-checking in platform-level content moderation.

These findings underline the importance of greater platform transparency, enhanced labelling practices, and inclusion of local fact-checkers in global disinformation response frameworks.

5 Conclusions

This report examines how over 1,200 online content that carried false claims identified in 150 fact-checks from the MedDMO's fact-checking partners between January and June 2024 were treated by social platforms and online news outlets, and we have the following findings.

In the contexts of Cyprus and Greece, Meta indeed applies disinformation labels and provides direct links to fact-checking articles, but we observed significant inconsistency in their deployment. Identical posts carrying the same false claim sometimes bore a warning banner and reference link, while others did not, suggesting a flaw or gap in Meta's labelling system (and raising the question of whether this inconsistency is specific to Greek-language content). On X, Community Notes were used to flag only a small fraction of misleading posts, leaving most false claims unannotated. Instagram, YouTube, and TikTok each yielded too few flagged instances in our dataset to support firm conclusions about their disinformation interventions. Finally, among the leading online news outlets in Greece and Cyprus where these false claims appeared, fewer than one in five articles were ever corrected, "labelled" as fake, or taken down—meaning that debunked content remains live without context.

In Malta, although our fact-check sample was small, false claims recurred across Facebook, Instagram, YouTube, and TikTok. The majority of the checks focused on politics, society, or the economy, often tracing falsehoods back to speeches by the Prime Minister or Opposition leader, forwarded WhatsApp messages, online news outlets, official statements from organisations, or—in some cases—no identifiable origin. We observed only a single instance of a Meta-warning label—on a claim already debunked internationally by a non-local fact-checking organisation. There are no systematic removals or labels on Maltese-language content across the platforms. This can be justified since there are no known collaborations between Malta's fact-checking organisations and online social media platforms. Fact-checking in Malta was recently established, driven by the Times of Malta and later the University of Malta under MedDMO, and remains disconnected from platform moderation channels, so local debunkings rarely feed into in-platform remediation or labelling. These few traced posts or online content nonetheless drew high user engagement, underscoring the absence of any in-platform context tools for Maltese audiences.

Together, these findings reveal critical gaps in both platform moderation and publisher accountability. Inconsistent labelling, sparse takedowns, and near absence of post-publication corrections allow debunked content to persist unchallenged. Strengthening collaboration between MedDMO fact-checkers, national

authorities, and Very Large Online Platforms—and advocating for uniform labelling, timely removals, and broader language coverage—will be central to achieving the objectives of MedDMO Task 5.1 and promoting digital information resilience in Cyprus, Greece, and Malta. Moreover, this methodology leverages the MedDMO fact-check archive as a "ground truth": recent fact-checked entries can be regularly sampled to verify whether each was labelled, contextualised, demoted, or removed on the platform where it appeared, and to track over time any improvements—or backsliding—in inconsistent labelling or delayed takedowns. By combining automated data collection with targeted manual audits, Task 5.1 can deliver robust, nuanced compliance assessments that drive continuous platform accountability. Coupled with reinforced regulatory oversight—especially through alignment with the EU's Digital Services Act—these efforts can foster more consistent platform compliance and meaningful publisher corrections. Ultimately, by institutionalising these evidence-based practices within national media frameworks, Cyprus, Greece, and Malta can safeguard democratic discourse and restore public trust in the online information environment.

6 References

[European Commission 2022] European Commission (2022). Strengthened Code of Practice on Disinformation. https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation

[European Commission 2018] European Commission (2018). Code of Practice on Disinformation. https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation

[Leonidou et al. 2023] Pantelitsa Leonidou, Nikos Salamanos, Michael Sirivianos, Loukia Taxitari, Socrates Ioakeim, Andreas Phylactou (2023), 1st Report of platform practices and national authorities support 2023 (for Cyprus, Greece & Malta). https://meddmo.eu/wp-content/uploads/2024/05/FINAL_1stReport_Monitoring_the_platform_practices_compressed.pdf

[Kaplan 2025] Kaplan, J. (2025, January 7). More speech and fewer mistakes. https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/

[BBC 2025] BBC News. (2025, May 4). TikTok adds 'unverified content' warning to videos. https://www.bbc.com/news/articles/cly74mpy8klo

[Open Online 2024] Fact-checking Team. (2024, January 5). Attenzione alla truffa dei fantomatici bagagli smarriti in aeroporto "venduti a due euro". Open Online. https://www.open.online/2024/01/05/truffa-bagagli-smarriti-aeroporto-venduti-due-euro-fc/