

IBERIFIER — Iberian Digital Media Observatory

Assessment of VLOP and VLOSE Signatory Reports for the Strengthened Code of Practice on Disinformation. Report No. 4

Report

Deliverable — D5.4



Co-funded by
the European Union

Iberifier – Iberian Digital Media Observatory has received funding from the European Commission under the Call DIGITAL-2023-DEPLOY-04, European Digital Media observatory (EDMO) — National and multinational hubs, with the reference IBERIFIER Plus - 101158511

Programme: Digital Europe Programme (DIGITAL)
Call: Accelerating the best use of technologies (DIGITAL-2023-DEPLOY-04)
Work program topic: European Digital Media Observatory (EDMO) – National and multinational hubs
Grant Agreement: IBERIFIER Plus - 101158511
Project coordinator: Ramón Salaverría (University of Navarra)
Term: May 1, 2024 – Oct 31, 2026 (30 months)
Report submission date: June 30, 2026
Dissemination level: Public
Task 5.4: Assessment of VLOPs and VLOSEs Signatory Reports for the Strengthened Code of Practice on Disinformation. Report No. 4
Deliverable Task Leader(s): Carolina Moreno-Castro (Universitat de València) and Miguel Paisana (OberCom)

Cite this Deliverable 5.4 as:

Moreno-Castro, C., Paisana, M., Serra-Perales, A., Azurmendi, A., Vázquez-Herrero, J., Calvo-Miguel, D., Falcó-Moreno, N., Cano-Orón, L., Negreira-Rey, M.-C., Quian, A., Vizoso-García, Á., Cardoso, G.; Cabrera García-Ochoa, Y., Irazo-Cabrera, M., Sánchez-Hualde, A., & Salaverría, R. (2025). IBERIFIER - Assessment of VLOPs (Very Large Online Platforms) and VLOSEs (Very Large Online Search Engines). Signatory Reports for the Strengthened Code of Practice on Disinformation. Report No. 4. IBERIFIER. <https://doi.org/10.5281/zenodo.20846772>

Access more reports and publications at iberifier.eu

Coordinator



Partners



Associate



List of participants

Participant	Partner Institution
University of Valencia team (UEVG)	
Moreno-Castro, Carolina	UEVG
Serra-Perales Ana	UEVG
Calvo, Dafne	UEVG
Cano-Orón, Lorena	UEVG
Falcó Moreno, Núria	UEVG
Cabrera García-Ochoa, Yolanda	UEVG
Iranzo Cabrera, María	UEVG
University of Santiago de Compostela team (USC)	
Vázquez-Herrero, Jorge	USC
Negreira-Rey, María-Cruz	USC
Quián, Alberto	USC
Vizoso-García, Ángel	USC
Iscte - Instituto Universitário de Lisboa	
Cardoso, Gustavo	Iscte
OberCom - Observatório da Comunicação	
Paisana, Miguel	OberCom
Universidad de Navarra (UNAV)	
Azurmendi, Ana	UNAV
Salaverría, Ramón	UNAV
Sánchez-Hualde, Anais	UNAV

Table of contents

1. Introduction	6
1.1. Objectives	9
1.2. Methodology	9
2. Deep assessment for the signatory Google Search	11
2.1. Media context: How did outlets in Portugal and Spain cover Google's commitment under Article 17 of the EU Code of Practice on Disinformation?	11
2.2. Quality of response assessment	16
2.3. Verification of reported information	16
2.4. Evaluation of actions taken	17
2.5. Recommendations/need for improvement	17
3. Deep assessment for the signatory YouTube	18
3.1. Media context: How did outlets in Portugal and Spain cover YouTube's commitment under Article 17 of the EU Code of Practice on Disinformation?	18
3.2. Quality of response assessment	23
3.3. Verification of reported information	23
3.4. Evaluation of actions taken	23
3.5. Recommendations/need for improvement	24
4. Deep assessment for the signatory TikTok	25
4.1. Media context: How did outlets in Portugal and Spain cover TikTok's commitment under Article 17 of the EU Code of Practice on Disinformation?	25
4.2. Quality of response assessment	30
4.3. Verification of reported information	31
4.4. Evaluation of actions taken	31
4.5. Recommendations/need for improvement	32
5. Deep assessment for the signatory Facebook	33
5.1. Media Context: How did outlets in Portugal and Spain cover Facebook's commitment under Article 17 of the EU Code of Practice on Disinformation?	33
5.2. Quality of response assessment	38
5.3. Verification of reported information	39
5.4. Evaluation of actions taken	39
5.5. Recommendations/need for improvement	40
6. Deep assessment for the signatory Instagram	41
6.1. Media context: How did outlets in Portugal and Spain cover Instagram's commitment under Article 17 of the EU Code of Practice on Disinformation?	41
6.2. Quality of response assessment	47
6.3. Verification of reported information	48
6.4. Evaluation of actions taken	48

6.5. Recommendations/need for improvement	49
7. Deep assessment for the signatory on LinkedIn	50
7.1. Media context: How did outlets in Portugal and Spain cover LinkedIn's commitment under Article 17 of the EU Code of Practice on Disinformation?	50
7.2. Statements from LinkedIn for not endorsing Commitment 17	52
9. Conclusions	53
10. Bibliography	57

1. Introduction

In this Deliverable 5.4, **usability is understood as both a technical attribute of platform design and a democratic condition for meaningful user empowerment.** Citizens can only benefit from media literacy and transparency measures if they can find them, understand them, use them, contest platform decisions through them, and receive contextual information that supports autonomous judgment. This approach aligns with the evolution of the European regulatory framework, particularly the Digital Services Act, which links platform accountability to users' rights, transparency obligations, complaint mechanisms, recommender-system transparency, and systemic risk mitigation duties (European Parliament and Council, 2022).

During the present reporting period, the assessment of platform compliance with Commitment 17 of the Code of Practice on Disinformation must therefore be situated in a substantially different regulatory and operational context from that examined in Deliverable 5.3. The Code has continued to consolidate its role within the DSA framework. At the same time, the European Commission has intensified its scrutiny of the design, accessibility and effectiveness of platform interfaces from the perspective of citizens' rights, user empowerment and democratic resilience (European Commission, 2025a, 2025b, 2026b). This transition is particularly relevant because the Code of Practice on Disinformation was recognised as a Code of Conduct under Article 45 of the DSA, with effect from 1 July 2025, thereby transforming platform commitments on disinformation, media literacy and user empowerment into a meaningful benchmark for assessing the compliance efforts of VLOPs and VLOSEs adhering to the Code (European Commission, 2025b; European Commission, 2026b).

A key development for this deliverable is the publication, in late March 2026, of the latest reports by the Code's signatories in the Transparency Centre. These reports cover the period from 1 July to 31 December 2025 and constitute the first reporting cycle submitted following the Code's new status under the DSA framework (European Commission, 2026a). This timing is significant for IBERIFIER's monitoring work, as it allows the present deliverable to assess whether the commitments reported by platforms after July 2025 show substantive progress in media literacy, user empowerment, and the practical usability of transparency-related tools.

This new phase is particularly relevant for the analysis of platform usability for citizens. Under Commitment 17, **signatories are expected not only to support media literacy initiatives, but also to design tools that help users understand, contextualise and critically assess the information they encounter online.** The European Commission's media literacy framework stresses that platforms should provide effective measures and tools, raise awareness of them, and empower users by offering context for visible content (European Commission, 2026e). In this sense, usability becomes a substantive dimension of compliance: tools that are difficult to locate, poorly explained, insufficiently localised, weakly promoted or disconnected from everyday platform use cannot be considered fully effective media literacy measures.

The March 2026 reporting round shows that several signatories continue to report actions under Commitment 17, including media literacy campaigns, partnerships with fact-checkers, election-related awareness initiatives and educational resources targeting young people or vulnerable groups (European Commission, 2026a). TikTok, for instance, reported ongoing media literacy and critical thinking campaigns in several EU/EEA countries, including Spain and Portugal, as well as temporary election integrity campaigns in cooperation with media literacy and fact-checking partners. Meta announced the publication of its first Media Literacy Annual Plan in July 2025, covering tools and initiatives on Facebook and Instagram, such as digital citizenship programmes and election literacy activities. Google also reported support for youth-focused media literacy and election disinformation initiatives, including funding for ThinkYoung hackathons across Europe. However, the analysis of these reports must distinguish between the existence of initiatives and their practical accessibility, visibility, measurability and usefulness for citizens in Spain and Portugal, as already highlighted in previous IBERIFIER assessments of platform reporting under the Code (Moreno Castro et al., 2025a; Moreno Castro et al., 2025b; Moreno Castro et al., 2025c).

At the same time, developments between December 2025 and June 2026 reinforce the need to evaluate platforms not only by their declared commitments, but also by the usability and effectiveness of the interfaces through which citizens exercise their rights. The DSA requires platforms to provide mechanisms that allow users to flag illegal content, obtain clear explanations for moderation decisions, and access meaningful transparency and redress mechanisms (European Parliament and Council, 2022). The Commission has also highlighted that the DSA is intended to enhance online transparency through stronger reporting obligations, researcher access to data, risk assessment procedures, and scrutiny of systemic risks (European Commission, 2025c; European Commission, 2026c). Nevertheless, **recent enforcement actions indicate that interface design remains a central area of concern**, especially where reporting flows, recommender systems, age assurance mechanisms or appeal procedures are not sufficiently clear, accessible or effective.

This is particularly important given the Commission's enforcement activity in the first half of 2026. On 26 January 2026, the Commission opened a new formal investigation into X concerning Grok and extended its ongoing investigation into X's recommender systems under the DSA (European Commission, 2026f). On the same date, WhatsApp Channels was designated as a Very Large Online Platform, bringing its dissemination-oriented functionality within the enhanced DSA supervision framework (European Commission, 2026g). In February 2026, the Commission preliminarily found TikTok in breach of the DSA in relation to addictive design, citing features such as infinite scroll, autoplay, push notifications and highly personalised recommender systems (European Commission, 2026h). In April 2026, the Commission preliminarily found Meta's Instagram and Facebook in breach of the DSA for failing to adequately identify, assess, and mitigate the risks posed by minors under 13 accessing their services (European Commission, 2026i). These cases underscore that user **empowerment** cannot be reduced to information provision alone; it also depends on

whether platform architecture enables users to make informed, autonomous and safe choices.

Further EU initiatives adopted or advanced during this period also point to a broader policy shift from declaratory transparency towards citizen-centred effectiveness. In May 2026, the Commission published draft guidelines on trusted flaggers under the DSA, seeking to clarify how this mechanism should operate and how platforms should prioritise notices from expert organisations (European Commission, 2026d). In the field of digital and media literacy, the 2026 update of the Commission's guidelines for teachers introduced new sections on generative AI, influencers, pre-bunking, updated lesson plans and recommendations for school leaders and policymakers (European Commission, 2026e). These developments are highly relevant to Commitment 17, as they expand the understanding of media literacy beyond generic awareness-raising and towards practical resilience against manipulative content, AI-generated disinformation and platform-mediated influence.

In this context, **the present deliverable provides an updated assessment of the actions reported by VLOP and VLOSE signatories under Commitment 17**, with particular attention to their relevance for users in Spain and Portugal. Building on Deliverables 5.1, 5.2 and 5.3, this report moves further from a purely descriptive account of declared platform initiatives towards an assessment of their usability for citizens, including visibility, accessibility, clarity, linguistic and territorial coverage, integration into platform interfaces, and the availability of measurable outcomes (Moreno-Castro, Paisana, et al., 2024; Moreno Castro et al., 2025a; Moreno Castro et al., 2025b). This approach is also consistent with IBERIFIER's broader monitoring activities on the implementation of the Code of Practice on Disinformation, which have stressed the importance of evaluating not only the presence of platform commitments, but also the quality, specificity and accountability of the evidence provided by signatories (Moreno Castro et al., 2025c).

The analysis, therefore, examines whether the actions reported in the March 2026 Transparency Centre submissions provide sufficient evidence that platforms are strengthening media literacy and critical thinking in ways that are meaningful to citizens. It also considers whether these measures address the regulatory priorities that have emerged since December 2025, including safer interface design, protection of minors, transparency in recommender systems, effective reporting channels, trusted flagger procedures, researcher data access, and the growing risks associated with generative AI and manipulated content (European Commission, 2025c; European Commission, 2026c; European Commission, 2026d; European Commission, 2026e). In doing so, the deliverable contributes to IBERIFIER's continued monitoring of the Code's implementation in the Iberian context. It offers updated recommendations to improve the public value, accountability and usability of platform-based media literacy measures.

1.1. Objectives

1. To evaluate a sample of the **reports submitted on 31 March 2026 by selected VLOPs and VLOSEs** on their good practices under the Strengthened Code of Practice on Disinformation, focusing on their continued compliance and engagement with measures aimed at empowering users through Commitment 17.
2. To identify and recommend **concrete and specific measures to improve the implementation of the Code of Practice on Disinformation**, with particular attention to the effectiveness of Commitment 17 on media literacy and to its localised implementation in Portugal and Spain, based on the good practices reported by the selected platforms in the reports submitted on 31 March 2026.
3. To examine **media coverage in Portugal and Spain** for the period from January 1st to April 30th 2026, to assess whether news was published concerning the selected signatories of the Code of Practice on Disinformation, particularly in relation to Commitment 17 and the good practices described in the reports submitted by the platforms on 31 March 2026.

1.2. Methodology

In line with the methodology of previous IBERIFIER Plus reports (Deliverable D5.2 and Deliverable 5.3) released in June 2025 and December 2025, the assessment presented in this report (Deliverable D5.4) follows a standardised and collaborative methodology developed in coordination with the methodological subgroup, which includes representatives from each European Digital Media Observatory (EDMO) Hub. This subgroup met several times to discuss and harmonise the assessment approach, ensuring a consistent, transparent, and fair evaluation process across all participating hubs.

A key feature of the methodology is the use of **double coding for each assessment**. This means that two independent reviewers evaluated each signatory's report, compared their interpretations, and resolved discrepancies through discussion at the first level within the same partner team. The second level was intercode among different partners. At the end, 8 coders agree on the evaluation score. This approach enhances the objectivity of the findings and reduces the risk of individual bias. The combination of evidence-based and qualitative assessments ensures that the report documents compliance in formal terms and critically reflects on the effectiveness and strategic relevance of the signatories' actions in combatting disinformation and fostering a more informed public. **The assessment focused on three core dimensions:**

1. **Quality of the report:** This criterion considers how clearly and systematically the signatories present their activities. It includes the organisation of the report, the clarity of language, the use of structured formats, and the overall

accessibility of the information for external reviewers. A three-point Likert scale (1 = Poor, 2 = Partial, 3 = Good, 4 = Not applicable) was used, following the EDMO standardisation and guidelines (Cesarini et al., 2024; EDMO, 2024; EFCSN, 2024; Botan & Meyer, 2025).

2. **Evidence provided by the signatories:** This dimension evaluates whether the described actions are supported by concrete documentation or data. Evaluators assessed the credibility, relevance, and sufficiency of the evidence presented, including links to relevant resources, examples of implemented tools, training materials, and statistical indicators. A three-point Likert scale (1 = Poor, 2 = Partial, 3 = Good, 4 = Not applicable) was used, following the EDMO standardisation and guidelines (Cesarini et al., 2024; EDMO, 2024; EFCSN, 2024; Botan & Meyer, 2025).
3. **Impact of the actions implemented:** The final dimension assesses the extent to which the actions undertaken by the signatories align with the goals of Commitment 17, which focuses on enhancing media literacy. This includes developing digital skills, deploying educational tools, targeting vulnerable populations, and implementing broader awareness-raising strategies. The assessment also considers how well these actions are adapted to regional or national contexts, which is particularly relevant in the Iberian context of Spain and Portugal, where media literacy needs and disinformation dynamics may differ from those in other parts of Europe. A five-point Likert scale (1 = Very Poor, 2 = Poor, 3 = Fair, 4 = Good, 5 = Excellent) was used, following the EDMO standardisation and guidelines (Cesarini et al., 2024; EDMO, 2024; EFCSN, 2024; Botan & Meyer, 2025).
4. **Media context:** Analysis of how news outlets in Portugal and Spain reported on the signatories' actions and engagement under Commitment 17 of the EU Code of Practice on Disinformation, focusing on the tone, frequency, and framing of coverage during the reporting period. The methodology used for this descriptive analysis involved a targeted media search in Portugal and Spain, focusing on identifying coverage of companies that have signed the *Code of Practice on Disinformation*, particularly in connection with Commitment 17, during the reporting period for the signatories (from 1 January to 1 April 2026). For the Spanish and Portugal media, searches were conducted using the Factive News platform. Both research teams employed Boolean operators designed to capture relevant content. The search criteria included: a) News stories referring to the different signatories analysed in this report (Google Search, YouTube, LinkedIn, TikTok, Instagram, Facebook, Bing); The time frame between January 1st and April 30th 2026; and c) the exact text: “código de buenas prácticas” in Spanish and “Código de boas-práticas” in Portuguese.

2. Deep assessment for the signatory Google Search

2.1. Media context: How did outlets in Portugal and Spain cover Google's commitment under Article 17 of the EU Code of Practice on Disinformation?

During the first months of 2026, Spanish-language media coverage presents a more complex picture of Google's performance under Article 17 of the EU Code of Practice on Disinformation. Unlike the previous period, where Google was mainly portrayed as retreating from EU commitments, the 2026 coverage combines criticism of Google Search and Gemini with references to concrete media literacy initiatives supported by Google, particularly in Spain through Fad Juventud and Google.org.

Overall, the press frames Google as a central actor in the information ecosystem: on the one hand, it supports educational and fact-checking initiatives aimed at strengthening users' resilience against misinformation; on the other hand, its search engine, AI Overviews and Gemini are repeatedly presented as sources of new risks, especially in relation to health information, electoral content, AI-generated images and the amplification of false or poorly verified material.

The evidence is therefore mainly based on Spanish outlets and Spanish-language international coverage, with several articles from Latin American media that address Google's tools and policies in broader Spanish-speaking contexts.

The most relevant headlines on Google Search in Spanish-language dailies and agencies

- “Buscar síntomas en Google: cinco errores que pueden empeorar tu salud y cómo evitarlos” (*Searching for symptoms on Google: five mistakes that can worsen your health and how to avoid them*). Source: *Infobae*, April 29, 2026. The news story warns that Google has become the first source of medical information for millions of users. Still, it stresses that search results may be shaped by popularity, SEO, advertising or unverified testimonials rather than clinical accuracy. It recommends prioritising official and specialised sources, checking authorship and publication dates, and treating Google as an orientation tool rather than a diagnostic authority.
- “EH Verifica lanza ClaimCheck, nuestra herramienta de IA para combatir la mentira política en Honduras” (*EH Verifica launches ClaimCheck, its AI tool to combat political lies in Honduras*). Source: *El Heraldo*, April 24, 2026. The piece reports on ClaimCheck, an AI tool presented to the Inter-American Press Association and the Google News Initiative. The tool transcribes speeches, interviews and YouTube videos, identifies verifiable claims and helps fact-checkers prioritise political statements. The article presents the Google

News Initiative as a partner in strengthening newsroom capacity to combat disinformation, while also emphasising that human editorial judgement remains essential.

- “La plataforma de Google Trends para las elecciones presidenciales en Colombia” (*Google Trends platform for the presidential elections in Colombia*). Source: *Tecnogus*, April 20, 2026. The article describes Google’s launch of a public election-search-trends platform in Colombia. It presents Google Trends as a tool to democratise access to search data, identify citizens’ electoral concerns, and connect users with official information, including information on where to vote. The article frames Google Search as part of the civic information infrastructure during elections.
- “Elecciones 2026: cómo usar la IA contra la desinformación y lograr un voto informado” (*2026 elections: how to use AI against disinformation and achieve an informed vote*). Source: *El Comercio*, March 25, 2026. The article presents Google tools such as NotebookLM, Gemini, Google News, Google Lens, and SynthID as resources for comparing electoral programmes, accessing official sources, verifying visual content, and detecting AI-generated material. The framing is positive and associates Google’s AI ecosystem with informed voting and digital verification.
- “Las plataformas combaten menos la desinformación tras la vuelta de Trump, según un informe” (*Platforms are doing less to combat disinformation after Trump’s return, according to a report*). Source: *Agencia EFE*, March 3, 2026. The article cites the European Fact-Checking Standards Network report, *The Great Retreat*, which argues that Google Search, YouTube, Bing, and LinkedIn have moved away from key fact-checking commitments under the EU Code of Practice on Disinformation. It also warns that generative AI and AI-powered search summaries increase the industrial-scale production and amplification of misleading content.
- “Las plataformas defienden sus medidas —IA y humanas— para desterrar los contenidos de odio” (*Platforms defend their AI and human measures to remove hateful content*). Source: *Agencia EFE*, March 11, 2026. The article reports on a forum in Madrid where Google, TikTok and Meta defended their moderation systems against hate speech and disinformation. Google’s representative stressed the role of content removal, recommendations and reduced visibility for lower-quality content, while acknowledging that these systems are not infallible.
- “Engañan a Gemini y ChatGPT para que den información falsa en sus resultados, simplemente escribiendo un blog con esos datos” (*Gemini and ChatGPT are tricked into giving false information simply by writing a blog with those data*). Source: *CE Noticias Financieras*, February 23, 2026. The article

describes an experiment showing that Google's AI tools and ChatGPT can reproduce fabricated information when it is published online in a credible-looking format. The piece highlights the vulnerability of AI-powered search and chatbot systems to source manipulation and recommends that users verify sources and compare multiple answers.

- “Google elimina algunos resultados de salud de AI Overviews después de identificar desinformación peligrosa” (*Google removes some health results from AI Overviews after identifying dangerous misinformation*). Source: *CE Noticias Financieras*, January 12, 2026. The article reports that Google removed or restricted some health-related AI Overviews after problematic summaries were found to contain potentially dangerous misinformation. This reinforces media concerns about the reliability of AI-generated search answers in high-risk domains such as health.
- “La mitad de las menciones a inmigración en X contienen desinformación, según un estudio de Fad Juventud” (*Half of the mentions about immigration on X contain disinformation, according to a Fad Juventud study*). Sources: *El Diario Vasco / Europa Press*, February 6, 2026; *El País*, February 6, 2026. The articles report on the study *Entre el ruido y los datos*, produced by Fad Juventud with support from Google.org. The study finds that almost two out of ten analysed posts on X contain disinformation, rising to half of the posts related to immigration. It identifies algorithmic incentives, polarisation, and a lack of critical skills as key drivers of disinformation, directly linking the issue to media literacy.
- “La clave no es prohibir Internet, sino educar con él: la estrategia de Fad Juventud y Google para usar la red de forma segura” (*The key is not to ban the Internet, but to educate with it: Fad Juventud and Google's strategy for safer internet use*). Source: *La Razón*, February 5, 2026. The article presents the initiative “De la incertidumbre a la confianza: educar para entornos digitales seguros,” which is supported by Google and Fad Juventud. It highlights the long-running programme *Surfear la Red*, active since 2019, which combines classroom workshops, gamified activities, teacher training, family resources, online materials, digital mentoring and awareness campaigns. More than 40,000 children, teenagers and young people are reported to have participated in training actions.
- “Cuando Google es el dietista: la desinformación sobre nutrición es un peligro” (*When Google is the dietitian: nutrition misinformation is dangerous*). Source: *El Periódico / Verificat*, January 3, 2026. The article warns that internet searches for diets and nutrition increase sharply in January, according to Google Trends. Still, it stresses that many users rely on unofficial sources, influencers, and unverified advice. The piece underlines the health risks of misinformation and recommends professional and official sources.

Main findings on Google Search and Article 17

1. Partial evidence of media literacy initiatives

Unlike the previous reporting period, the 2026 coverage does identify media literacy initiatives involving Google, particularly through Google.org's support for Fad Juventud. The *Surfear la Red* programme and the event "De la incertidumbre a la confianza" are directly aligned with Article 17, as they promote critical thinking, digital resilience, safer internet use and the ability to detect and respond to misinformation.

However, these initiatives are framed as educational projects mainly focused on young people and digital safety. The coverage does not show a broader, systematic Google Search-led strategy in Spain or Portugal to address media literacy among the general population.

2. Google Search remains a gateway to risky information

Several articles portray Google Search as the first point of access for sensitive information, especially in health, nutrition and electoral contexts. This makes Google central to citizens' information practices but also creates risks of responsibility. Coverage on health searches and nutrition misinformation stresses that search ranking, SEO, advertising and non-expert sources can expose users to misleading or harmful content.

This framing suggests that Google's role under Article 17 cannot be limited to providing access to information; it must also include stronger user guidance, clearer source evaluation and more visible cues for reliability.

3. AI Overviews and Gemini introduce new verification risks

The attached news sample repeatedly highlights vulnerabilities in Google's AI ecosystem. Articles on AI Overviews, Gemini and manipulated blog content show that Google's AI systems may reproduce false claims, amplify single-source misinformation or provide misleading summaries in high-risk areas.

This undermines the objectives of Article 17 because AI-generated answers can reduce users' incentive to compare sources. When misinformation appears as a concise answer at the top of the search experience, users may perceive it as more authoritative than ordinary search results.

4. Google's tools are also presented as part of the solution

The coverage is not exclusively critical. Google Trends, Google Lens, SynthID, NotebookLM, Gemini, Google News and the Google News Initiative are presented in several articles as tools that may support civic information, electoral awareness, visual verification and fact-checking capacity.

This creates a dual framing: Google is both a risk amplifier and a resilience provider. Its compliance with Article 17, therefore, appears partial and uneven rather than absent.

5. Human verification remains essential

Across the articles, especially those on ClaimCheck, platform moderation and AI-generated misinformation, the media emphasise that AI tools cannot replace human verification. Journalistic judgement, fact-checkers, professional sources and official institutions remain central to combating disinformation.

This is relevant to Article 17 because media literacy is not only about access to technological tools but also about teaching citizens how to evaluate evidence, understand source credibility, and avoid overreliance on automated systems.

6. Regulatory pressure and platform accountability remain unresolved

The EFE article on *The Great Retreat* presents Google Search and YouTube as stepping back from important fact-checking commitments under the EU Code of Practice on Disinformation. This contrasts with the more positive coverage of Google-supported educational initiatives and suggests a gap between project-level media literacy actions and structural compliance with the broader EU disinformation framework.

Overall assessment

The 2026 media coverage portrays Google's performance under Article 17 as mixed and incomplete. On the positive side, Google is linked to concrete educational and media literacy initiatives in Spain, especially through Fad Juventud and Google.org, and its tools are presented as useful for electoral information, visual verification and newsroom fact-checking.

At the same time, the press identifies persistent and growing risks associated with Google Search, Gemini and AI Overviews. These include the amplification of health misinformation, the reproduction of fabricated claims, the generation or circulation of misleading AI images and the difficulty users face in distinguishing verified information from plausible but false content.

In this sense, **Google appears to be contributing to media literacy in specific contexts, but has not yet offered a sufficiently comprehensive, transparent, or systematic response to the challenges identified under Article 17.** The company's role remains ambivalent: it supports public resilience against misinformation while simultaneously operating search and AI systems that can intensify the very problems Article 17 seeks to address.

Table 1. Analysis of Assessment Commitment 17: Google Search

Signatory: Google Search	
Section	Score
1. Quality and Completeness Documentation of media literacy actions	2- adequate
2. Evidence Provision of supporting data for listed actions	2- adequate
3. Impact Impact of media literacy actions	2- adequate
4. Rating Effectiveness of media literacy actions	3- good

*Scale of scores for sections 1 and 2: 1- poor, 2- adequate, 3- good.

*Scale of scores for sections 3 and 4: 1- very poor, 2- poor, 3- fair, 4- good, 5- excellent.

2.2. Quality of response assessment

Google Search provides metrics on the impact of content advisories by country and on some activities included in the Super Searchers action. **Although this data seems adequate, it would be useful to provide some metrics related to other reported actions** (f.e, people engaging with the Super Searchers programme or how people use the "about this result" option). There is also a lack of detail on the countries where the initiatives took place.

2.3. Verification of reported information

Regarding the impact of the actions, once again, **more context on the provided metrics is needed**. For example, although metrics regarding other features offered by Google, such as content advisory notices, views of the "more about this page" tool, and impressions of rapidly changing results, are enough, they don't seem useful to assess how users interact with the tools, as they are not accompanied by information related to how users interact with those features.

Regarding, for example, the initiative aimed at librarians, the impact seems very low (16 librarians trained for six months). Anyway, it is difficult to assess its impact, as there are no other details (for example, where the initiatives took place, how librarians will teach users or apply the knowledge acquired, etc.). On the other hand, the "Teach Kids Online Safety" lacks information on how the \$7 million has been invested in each European country.

2.4. Evaluation of actions taken

Although there are no new initiatives, **Google Search has updated the informational program "Super Searchers"**, in particular with the improvement of the SIFT method (Stop, Investigate the source, Find better coverage, and Trace claims to the original context), which, thanks to AI tools, **helps users to find reliable information and to understand it in its context**. However, it lacks sufficient granularity to assess the initiative's impact (for example, how many users followed the programme, where they were located, etc.). Google provides metrics and data; however, in general, more context is needed to assess the impact properly.

Regarding features such as Content Advisory Notices and More About this Result (Result, Image, Source), they seem relevant given their placement in Google's navigation. It's an improvement in the searching activity. However, **more context is needed to assess the impact**, as it's difficult to determine whether these figures indicate an improvement in media literacy in Spain. At the same time, some actions seem to have low relevance (for example, impression proportions seem low; it is necessary to understand how users interact with this feature).

2.5. Recommendations/need for improvement

Google Search generally provides comprehensive, detailed reports, including clear explanations of updates that enable better tracking of newly implemented actions. In this case, **the report highlights improvements in the Super Searchers initiative**. Overall, the actions related to media literacy are adequately documented.

- The technological features implemented by Google Search (i.e., "About this result" and "About this image") appear to have a limited impact when compared to the scale of Google Search usage. Their effectiveness would likely increase if accompanied by parallel actions to raise user awareness and promote their use in a similar way to how AI tools are promoted.
- Most of the metrics provided require additional context to properly assess how users interact with the tools designed to promote user empowerment.
- The report relies on the same metrics used in previous submissions, particularly those related to "About this result" and "About this image." While these data points are consistent, they offer limited additional insight compared to past reports.
- Google also includes specific features to detect images developed with AI, which is very valuable. It would be advisable to promote it to users further and include actions related to other AI-generated content.

3. Deep assessment for the signatory YouTube

3.1. Media context: How did outlets in Portugal and Spain cover YouTube's commitment under Article 17 of the EU Code of Practice on Disinformation?

During the first months of 2026, Spanish media coverage of YouTube focused less on its formal relationship with the European Commission and more on the platform's concrete risks as a space for audiovisual disinformation, especially in relation to AI-generated videos, health misinformation, deepfakes, political manipulation and algorithmic amplification. Compared with the first half of 2025, the tone became more problem-oriented: media outlets portrayed YouTube not only as a platform under regulatory pressure, but also as a central channel through which synthetic, misleading or low-quality content can reach vulnerable audiences.

The coverage also showed some positive examples of civic and educational responses to misinformation, particularly campaigns designed for young users and social media environments. However, these initiatives were mainly led by medical, journalistic or civil society actors rather than by YouTube itself. As a result, the media framing suggests that YouTube's contribution to Article 17 remains limited, particularly in proactive media literacy, user education, and cooperation with independent fact-checking and public-interest organisations.

The evidence is therefore mainly based on Spanish outlets and Spanish-language international coverage.

The most relevant headlines on YouTube in Portuguese and Spanish dailies

- “YouTube permite crear deepfakes con IA y reabre el debate sobre identidad digital” (*YouTube allows users to create AI deepfakes and reopens the debate on digital identity*). Source: DiarioBitcoin, April 9th 2026. The article reports that YouTube has expanded its generative AI tools, allowing users to create synthetic videos within its ecosystem. While this is framed as an innovation for audiovisual creativity, the article highlights serious concerns about identity misuse, deception, consent, disinformation and the difficulty of moderating realistic synthetic content at scale. The piece stresses the need for transparency, labelling, traceability and rapid response mechanisms.
- “Falsos médicos creados con IA inundan YouTube con desinformación” (*Fake AI-generated doctors flood YouTube with disinformation*). Source: La Vanguardia / Verificat, March 23rd 2026. The article presents a Verificat investigation into 14 YouTube channels that appear to have been created in Spain and use AI-generated doctors to target older adults with health advice. The study analysed 962 videos, many of which contained non-evidence-based

claims, miracle food recommendations, advice against medical treatments, or misleading health guidance. The investigation found that only 26% of the videos carried YouTube's "altered or synthetic content" label, even though platform policies require creators to disclose realistic AI-generated content.

- "Falsos médicos hechos con IA inundan YouTube con desinformación dirigida a personas mayores" (*Fake AI doctors flood YouTube with disinformation aimed at older people*). Source: La Vanguardia.com / Verificat, March 20th 2026. This shorter version of the Verificat investigation emphasises the scale of the problem: the analysed channels accumulated around 54 million views. The article underlines the vulnerability of elderly audiences and the potential breach of both YouTube's own policies and EU platform obligations regarding systemic risks to public health.
- "La inteligencia artificial genera médicos falsos que inundan YouTube con desinformación dirigida a adultos mayores" (*Artificial intelligence generates fake doctors that flood YouTube with disinformation aimed at older adults*). Source: CE Noticias Financieras, March 20th 2026. The article summarises Verificat's findings and stresses that most videos analysed did not disclose their synthetic nature. It also notes that Verificat alerted YouTube, but that, at the time of publication, no significant action had been taken. This reinforces concerns about enforcement gaps and insufficient platform responsiveness.
- "Los expertos en desarrollo infantil coinciden: los vídeos educativos de IA para niños en YouTube no son contenido neutro sino 'desinformación para bebés a escala industrial'" (*Child development experts agree: AI educational videos for children on YouTube are not neutral content but 'disinformation for babies at industrial scale'*). Source: El HuffPost, March 19th 2026. The article warns about the proliferation of AI-generated children's videos on YouTube that appear educational but include basic errors, unsafe behaviours and misleading concepts. Experts argue that low-quality AI content may affect children's cognitive development, attention and learning. The article also notes that YouTube requires AI labelling mainly when content appears realistic, leaving most children's animations outside the disclosure requirement.
- "Las plataformas combaten menos la desinformación tras la vuelta de Trump, según un informe" (*Platforms are doing less to combat disinformation after Trump's return, according to a report*). Source: Agencia EFE, March 3rd 2026. The article cites the European Fact-Checking Standards Network report *The Great Retreat*, which warns that major technology platforms have begun to step back from their commitments to combat disinformation. It explicitly mentions Google Search, Bing, YouTube and LinkedIn as having distanced themselves from important fact-checking commitments under the EU Code of Practice on Disinformation. This is the clearest reference in the sample to YouTube's potential non-compliance with the Code.

- “Los médicos alertan a los jóvenes sobre los peligros de la desinformación en redes sociales: ‘Desconfía de soluciones mágicas’”(Doctors warn young people about the dangers of misinformation on social media: ‘Be suspicious of magic solutions’). Source: El Diario Vasco, April 14th 2026. The article reports on the Spanish Medical Association’s campaign #noteenREDES, designed to alert young people to the risks of health misinformation on social media. The campaign includes four audiovisual pieces adapted to Instagram, TikTok and YouTube, aimed at debunking viral content about miracle products, mental health pseudotherapies, unsupported diets and anti-vaccine narratives.
- “OMC lanza la campaña ‘#noteenREDES, habla con tu médico’ para alertar de la desinformación sanitaria en redes sociales” (The Spanish Medical Association launches the campaign ‘#noteenREDES, talk to your doctor’ to warn about health misinformation on social media). Source: Infobae, April 7th 2026. This article also covers the #noteenREDES campaign and highlights its focus on the language and formats of Instagram, TikTok and YouTube. The campaign’s message is directly aligned with media literacy goals: users should distrust viral videos, consult official sources, contrast information and speak to qualified professionals.
- “‘No te dejes engañar por vídeos virales’: los médicos alertan sobre los peligros para la salud de las redes sociales” (‘Do not be fooled by viral videos’: doctors warn about the health dangers of social media). Source: ABC, April 7th 2026. The article reinforces the role of viral video platforms in spreading health misinformation and presents the OMC campaign as a public-awareness response. Although not led by YouTube, the initiative addresses precisely the type of audiovisual misinformation that Article 17 seeks to counter through user empowerment and media literacy.
- “Rusia habría entrenado a 1.000 influencers en América Latina para difundir desinformación” (Russia allegedly trained 1,000 influencers in Latin America to spread disinformation). Source: CE Noticias Financieras, April 3rd 2026. The article reports on a network of pro-Russian communicators and influencers active across Latin America. It notes that accounts associated with RT en Español and Sputnik Mundo have more than 6 million followers on YouTube. The article portrays YouTube as part of a transnational infrastructure for influence operations, narrative laundering and political polarisation.
- “Informe denuncia que Colombia figura en una red de más de 1.000 influenciadores que habrían sido entrenados por Rusia para difundir desinformación” (Report denounces that Colombia appears in a network of more than 1,000 influencers allegedly trained by Russia to spread disinformation). Source: El Tiempo / Agencia EFE, April 4th 2026. The article builds on the same investigation and identifies YouTube as one of the relevant channels for Russian state-aligned media to reach Spanish-speaking

audiences. It highlights tactics such as emotional manipulation, false equivalences, conspiracy injection and amplification of extremes.

- “El doctor Pedro de Pablos crea ‘Prostalia’, un canal para visibilizar la realidad del cáncer de próstata” (*Dr. Pedro de Pablos crea ‘Prostalia’, un canal para visibilizar la realidad del cáncer de próstata*). Source: La Provincia, March 29th 2026. This article offers a more positive counterpoint. It presents a YouTube channel created by a urologist that provides rigorous, accessible information about prostate cancer and combats misinformation. The article notes that patients increasingly turn to online content and AI tools for medical questions, creating a need for trustworthy professional content on platforms such as YouTube.

Alerts on code violations

1. Weak enforcement of synthetic-content labelling
The Verificat investigation shows that a large majority of AI-generated health videos analysed on YouTube did not carry the platform’s altered or synthetic content label. This indicates a possible gap between YouTube’s stated policies and their practical enforcement, especially in high-risk areas such as public health.
2. Health misinformation targeting vulnerable audiences
Several articles highlight YouTube’s exposure of older adults, young people and the vulnerable to harmful misinformation despite existing moderation policies.
3. Insufficient YouTube-led media literacy initiatives
The #noteenREDES campaign and other awareness efforts are clearly aligned with Article 17, but they are led by external actors, such as the Spanish Medical Association, rather than by YouTube. The coverage does not identify major YouTube-led initiatives in Spain or Portugal aimed at improving audiovisual verification skills, critical understanding of AI-generated content or resilience against health and political misinformation.
4. Retreat from fact-checking commitments under the EU Code
The EFE report on *The Great Retreat* explicitly mentions YouTube as one of the platforms that has moved away from important fact-checking commitments under the EU Code of Practice on Disinformation. This reinforces the concern already identified in the 5.3 period: YouTube’s cooperation with independent verification standards remains contested.
5. Deepfake and AI-video risks intensify Article 17 responsibilities.
YouTube’s expansion of AI-based video creation tools increases the importance of transparency, consent, labelling and user education. Media coverage stresses that deepfakes are no longer marginal or experimental, but are entering mainstream content creation. This makes media literacy around synthetic video a central requirement for Article 17 compliance.
6. Algorithmic amplification and monetisation remain unresolved.
Several articles point to the commercial and algorithmic incentives behind

disinformation: mass-produced AI videos, influencer networks, viral health claims and children's content generated at scale. These dynamics suggest that YouTube's responsibility under the Code cannot be limited to content removal; it must also address visibility, recommendation and monetisation systems.

Overall assessment

During the first months of 2026, Spanish-language media coverage portrays YouTube's performance under Article 17 as insufficient and increasingly challenged by AI-generated audiovisual disinformation. The platform is repeatedly associated with risks involving fake doctors, synthetic videos, deepfakes, low-quality children's content, health misinformation and transnational influence operations.

Although the coverage includes examples of positive countermeasures, such as the OMC's #noteenREDES campaign and professional medical channels like Prostalia, these initiatives are not presented as YouTube-led actions. Rather, they suggest that civil society, medical professionals and fact-checkers are trying to compensate for the lack of stronger platform-level education, labelling and enforcement.

The most serious concern relates to the gap between YouTube's policies and their implementation. Media reports show that synthetic or AI-generated content can circulate without adequate labelling, accumulate millions of views and target vulnerable groups before meaningful action is taken. This weakens the platform's contribution to public resilience against misinformation.

Overall, the press frames YouTube as a platform with heightened responsibility under Article 17, but with limited evidence of proactive compliance. Its role in media literacy appears mostly indirect, even as the risks associated with AI-generated video, health disinformation, and algorithmic amplification become more visible and urgent.

Table 2. Analysis of Assessment Commitment 17: YouTube

Signatory: YouTube	
Section	Score
1. Quality and Completeness Documentation of media literacy actions	1- poor
2. Evidence Provision of supporting data for listed actions	1- poor
3. Impact Impact of media literacy actions	2- poor
4. Rating Effectiveness of media literacy actions	2- poor

*Scale of scores for sections 1 and 2: 1- poor, 2- adequate, 3- good.

*Scale of scores for sections 3 and 4: 1- very poor, 2- poor, 3- fair, 4- good, 5- excellent.

3.2. Quality of response assessment

The VLOP describes several initiatives but does not provide consistent details on the impact of existing initiatives (for example, information panels and the Hit Pause initiative), nor on the expected impact of the grants and projects it supports in the context of Commitment 17. As stated by the platform, no new implementation measures were employed in terms of services, tools, or policies; in fact, the list of implementations is a copy of the previous report. There are no plans for further measures in the next six months to improve the commitment, and therefore, they provide no information on such measures. It should be noted that the same report does show changes to Commitments other than Commitment 17, which are not considered here because they are out of the scope of this report.

3.3. Verification of reported information

Data is insufficient; the reporting is technically the same as before, and the same questions persist. The only significant trend is the evolution of the impressions, which we cannot really attribute to any tangible factor, i.e., changes in the signatory's behaviour. In this sense, some questions arise: for example, is this rise due to the campaign organically gaining traction, or to a change in content typology?

3.4. Evaluation of actions taken

Although the metrics indicate increased impact, the data quality remains poor, making it difficult to assess the initiatives' real impact (once again, the metrics lack context).

Regarding the reported initiatives, some questions arise that would be advisable to address to improve the report's quality. For instance, the information panels have a significantly lower reach in Portugal than in Spain. Does YouTube know why and which measures are going to be carried out in order to improve its impact in Portugal? Or is there a way to strengthen the tool's reach? Also, it would be useful to know how YouTube can guarantee that panels are coherently associated with relevant videos or which systems are in place to ensure that panels are shown in all relevant videos. Other indicators are also needed, such as click-through rates and clicks. Lastly, it would be advisable to provide metrics beyond impressions, as YouTube is very effective at identifying which content attracts the most attention and motivates action.

Regarding the Hit Pause campaign, the channel has published only one new video in the past six months, and it was released in several languages. The video relates to the issue of synthetic content originating from AI, a very important concern that the VLOP might want to explore, given its major role as a content indexator and provider. Other than that, we find the same issues as with the Information panels in Measure 17.1.

Even though the impressions appear to be similar in Spain and Portugal, in proportion to the internet user base in both countries, deeper insights are needed, as impressions are a poor KPI by themselves. Other indicators are needed, such as click-through rates and clicks. The absence of appropriate metrics here is less understandable: the Hit Pause campaign is a YouTube channel, and YouTube provides state-of-the-art metrics for its content producers per channel.

The difficulties in assessing the effectiveness of the campaign remain in place, and as such, the assessment and evaluation remain unchanged. The Initiatives appear benign and interesting in nature, but their impact cannot be measured.

- Information Panels - 1 Very poor - for the reasons listed above (poor reporting and metrics), it is not possible to assess the effectiveness of the initiative. Also, no relevant information regarding the specific impact in Spain and Portugal. Impressions are an insufficient KPI in this instance.
- Hit Pause - 1 Very poor - for the reasons listed above (poor reporting and metrics), it is not possible to assess the effectiveness of the initiative. Also, no relevant information regarding the specific impact in Spain and Portugal. Impressions are an insufficient KPI in this instance.

3.5. Recommendations/need for improvement

The rating has been lowered to very poor in some areas because the reporting is the same as in the last one, and there is no consideration of context. Furthermore, there is no mention of any AI-related initiatives, although the VLOP is known to be interested in the technology and its impacts.

- There is no ongoing direct collaboration between YT and Iberifier. YT should increase the number of organisations and experts it consults with, as there is a wide range of organisations specialising in different facets of media and literacy skills. There are plenty of partners to consider in Spain and Portugal, and research has shown that specific issues arise in the Iberian Peninsula that could be actively mitigated by engaging with experts knowledgeable about the realities at hand.
- There appears to be some reach for said initiatives, but there is little to no granularity in reported data, between and within countries. Based on the impressions numbers, even when they are lower in some cases, we estimate there has been reach in the labels and campaigns implemented by the VLOP. But the data is limited and does not provide all the context needed to interpret the data. It would be necessary to include more detailed metrics

4. Deep assessment for the signatory TikTok

4.1. Media context: How did outlets in Portugal and Spain cover TikTok's commitment under Article 17 of the EU Code of Practice on Disinformation?

During the first months of 2026, Spanish-language media coverage of TikTok continued to portray the platform as a central actor in the spread of disinformation, particularly among young users and vulnerable audiences. Compared with the first half of 2025, the coverage became more diversified: alongside warnings about political manipulation, viral health misinformation, AI-generated content and migration-related falsehoods, several news pieces also reported on TikTok's cooperation with public institutions to provide official electoral information and promote safer information practices.

Overall, the press portrays TikTok as a platform with a dual role. On the one hand, it is repeatedly identified as a channel through which misleading content, viral rumours, synthetic videos, and emotionally persuasive narratives circulate rapidly. On the other hand, some articles indicate that TikTok is developing institutional partnerships and informational guides, particularly in electoral contexts, which partially align with the objectives of Article 17 regarding media literacy, user awareness, and resilience against disinformation.

The evidence is therefore mainly based on Spanish-language outlets and agency coverage, including Spanish and Latin American media.

The most relevant headlines on TikTok in Portuguese and Spanish dailies

- “Migrar por TikTok: el guía equivocado” (*Migrating through TikTok: the wrong guide*). Source: Agencia EFE – Servicio Internacional, April 30th 2026. The article tells the story of Alejandra Ramírez, a Venezuelan migrant who relied on TikTok videos that presented the route through the Darién jungle as quick and manageable. The reality was far more dangerous: hunger, rivers, deaths, kidnappings, scams and the failure of migration pathways. The article frames TikTok as a misleading source of practical information for vulnerable migrant. It connects the case to broader efforts, such as UNHCR's “Confía en El Tucán” campaign, to combat migration-related misinformation.
- “Una guía incorrecta sobre la migración a través de TikTok” (*An incorrect guide to migration through TikTok*). Source: CE Noticias Financieras, April 30th 2026. This shorter version of the EFE story reinforces the idea that TikTok can serve as a harmful guide to information when users rely on viral content rather than official or verified sources. The piece highlights the gap between idealised migration narratives on social media and the real risks migrants face.

- “ONPE y TikTok presentan guía para las Elecciones Generales de 2026 en el Perú” (*ONPE and TikTok present a guide for Peru’s 2026 General Elections*). Source: Andina / El Peruano, April 8th 2026. The article reports that TikTok partnered with Peru’s National Office of Electoral Processes to provide users with reliable and official electoral information. The guide includes voting dates, polling information, voting procedures, elected offices, and additional information to help you vote informed. Users can access the guide via TikTok searches for the election. The article also notes that TikTok provides educational videos through its safety account, including practical advice on identifying and combating disinformation.
- “La ONPE y TikTok lanzan una guía para las elecciones generales de Perú en 2026” (*ONPE and TikTok launch a guide for Peru’s 2026 general elections*). Source: CE Noticias Financieras, April 8th 2026. This article summarises the same initiative and presents it as a measure to provide accurate information, combat disinformation and support informed democratic participation. It is one of the clearest examples in the sample of TikTok cooperating with a public institution in a manner consistent with the objectives of Article 17.
- “Así funciona el escudo digital de TikTok para las elecciones presidenciales 2026 en Colombia” (*How TikTok’s digital shield works for Colombia’s 2026 presidential elections*). Source: Infobae, April 6th 2026. The article presents TikTok as an influential platform among young Colombians and describes its measures to protect the integrity of the 2026 presidential election. The piece frames TikTok not only as an entertainment platform but also as a space where political information, news and campaign content circulate. It therefore stresses the importance of ensuring that electoral content is reliable, accessible and safe.
- “Los médicos alertan a los jóvenes sobre los peligros de la desinformación en redes sociales: ‘Desconfía de soluciones mágicas’” (*Doctors warn young people about the dangers of misinformation on social media: ‘Be suspicious of magic solutions’*). Source: El Diario Vasco, April 14th 2026. The article reports on the Spanish Medical Association’s campaign #noteenREDES, designed to warn young people about health misinformation on social media. The campaign consists of four audiovisual pieces adapted to Instagram, TikTok and YouTube, addressing viral content such as miracle products, mental health pseudotherapies, unsupported diets and anti-vaccine narratives. Although not led by TikTok, the campaign directly targets the type of viral misinformation that circulates on the platform.
- “OMC lanza la campaña ‘#noteenREDES, habla con tu médico’ para alertar de la desinformación sanitaria en redes sociales” (*The Spanish Medical Association launches the campaign ‘#noteenREDES, talk to your doctor’ to warn about health misinformation on social media*). Source: Infobae, April 7th

2026. The article presents the same campaign as an audiovisual initiative adapted to TikTok's formats and language. It encourages users to distrust viral videos, consult official sources, contrast information and speak to qualified health professionals. This contributes to media literacy objectives, although the initiative is driven by the medical profession rather than by TikTok itself.

- “Bienestar o tendencia viral: ¿Cómo identificar si los ‘hacks’ de salud en TikTok son verdad o moda?” (*Wellness or viral trend: how to identify whether TikTok health hacks are true or just a fad*). Source: Milenio, April 24th 2026. The article warns that TikTok health trends often mix genuine interest in wellbeing with distorted narratives, marketing strategies and misinformation. Experts recommend verifying scientific evidence in humans, checking costs and demanding health certifications before adopting routines, treatments or supplements promoted online. The article frames TikTok as a space where health literacy is urgently needed.
- “Algunas personas buscan confirmar un diagnóstico que han creado en TikTok con un psicólogo” (*Some people seek confirmation from a psychologist for a diagnosis they created on TikTok*). Source: CE Noticias Financieras, April 21st 2026. The article discusses the spread of simplified mental health labels and self-diagnosis trends on TikTok. It warns about “labelitis” overdiagnosis and the overexposure of users to mental-health misinformation. The article is relevant to Article 17 because it highlights the need for critical skills and contextual understanding when consuming health-related content online.
- “¿Murió Marco Antonio Solís?: la verdad tras el rumor que causó conmoción” (*Did Marco Antonio Solís die?: The truth behind the rumour that caused shock*). Source: Ámbito Financiero Online, April 12th 2026. The article reports that a false rumour about the death of the singer Marco Antonio Solís spread rapidly on TikTok and Facebook. The false claim originated in an AI-generated video that simulated a solemn news report, using an automated voice and fake media-style branding. The article warns that such content is designed not only to generate clicks but also, in some cases, to steal personal data.
- “Presentación de una demanda para eliminar el perfil de inteligencia artificial ‘Doña María’, que critica a Lula y al gobierno” (*A lawsuit is filed to remove the AI profile ‘Doña María’, which criticises Lula and the government*). Source: CE Noticias Financieras, April 25th 2026. The article reports on a legal complaint against an AI-generated political avatar active on platforms including TikTok, Instagram, Facebook, YouTube and X. The complaint alleges early electoral campaigning, the misleading use of artificial intelligence, a lack of proper identification, and the spread of distorted or false information. This case reinforces concerns about AI-generated political content and the need for clear labelling and platform accountability.

- “Patria V: la batalla por el relato propio en un ecosistema digital hostil” (*Patria V: the battle for one’s own narrative in a hostile digital ecosystem*). Source: La República, April 19th 2026. The article presents the digital environment as an asymmetric battlefield in which major platforms, including TikTok, reward polarisation, disinformation, and hate as vectors of visibility. It highlights the role of generative AI in producing hyper-realistic disinformation at an industrial scale and stresses the need for practical tools, fact-checking, digital sovereignty and critical awareness.

TikTok’s main alerts on Code violations

1. Persistent exposure to viral and harmful misinformation
The analysed articles repeatedly identify TikTok as a space where misleading narratives circulate quickly, particularly on migration, health, mental health, celebrity rumours and politics. The migration case reported by EFE shows that viral TikTok videos can influence high-risk real-world decisions, exposing vulnerable users to physical danger, scams and exploitation.
2. Health and mental health misinformation affecting young audiences
Several news pieces focus on health “hacks”, pseudotherapies, miracle products, self-diagnosis and viral mental-health labels. These stories show that TikTok’s short-video environment can amplify simplified or misleading health narratives, making media literacy and source verification especially urgent for young users.
3. AI-generated and synthetic content risks
The cases of the false Marco Antonio Solís death rumour and the AI-generated “Doña María” political avatar show how synthetic content can be used to simulate credible news, manipulate public perception or influence political debate. These cases point to weaknesses in labelling, transparency and user education around AI-generated content.
4. Algorithmic amplification and emotional manipulation
The coverage repeatedly associates TikTok with content dynamics based on virality, emotion and polarisation. Articles on migration, political avatars and digital hostility suggest that the platform’s recommendation environment may amplify sensational or misleading content before users have the tools to verify it.
5. Partial but geographically limited evidence of institutional cooperation
Unlike the 2025 period, the 2026 sample includes positive examples of TikTok collaborating with public institutions, especially the ONPE electoral guide in Peru and the electoral integrity measures reported in Colombia. These initiatives are consistent with the spirit of Article 17 because they provide official information and encourage users to identify misleading content. However, the attached sample does not identify equivalent TikTok-led initiatives in Spain or Portugal.
6. External actors continue to compensate for platform gaps
The #noteenREDES campaign by the Spanish Medical Association is strongly

aligned with media literacy objectives, but it is not a TikTok initiative. This suggests that doctors, public institutions, journalists and civil society actors are taking on part of the responsibility for educating users about misinformation circulating on TikTok.

7. Insufficient evidence of Spain- and Portugal-specific Article 17 implementation
Although TikTok appears more active in electoral information partnerships in Latin America, the reviewed coverage does not show a comparable level of media literacy work in Spain or Portugal during the period analysed. This remains the main weakness for assessing TikTok's compliance with Article 17 in the Iberian context.

Overall assessment

During the first months of 2026, Spanish-language media coverage portrays TikTok's performance under Article 17 as mixed but still insufficient in Spain and Portugal. The platform is repeatedly associated with disinformation risks affecting vulnerable migrants, young people, patients, voters and consumers of viral entertainment content. The main concerns are the speed of viral amplification, the persuasive power of short video, the rise of AI-generated content and the lack of critical skills among users.

At the same time, the coverage is not entirely negative. The ONPE electoral guide in Peru and the electoral integrity measures reported in Colombia show that TikTok can cooperate with public institutions to provide official information, promote informed voting and combat online deception. These initiatives demonstrate partial alignment with Article 17, particularly in electoral contexts.

However, the evidence provided in the Factiva sample does not identify equivalent TikTok-led media literacy initiatives in Spain or Portugal. The most relevant Spanish example, the OMC's #noteenREDES campaign, is driven by medical institutions rather than by the platform. As a result, TikTok's compliance with Article 17 appears uneven: proactive in some Latin American electoral contexts, but weakly visible in the Iberian media landscape.

Overall, the press frames TikTok as a platform that has begun to develop some institutional tools against disinformation, but whose core risks remain unresolved. Its contribution to media literacy remains fragmented, geographically uneven, and insufficiently embedded in Spain and Portugal. After examining how Spanish- and Portuguese-language outlets framed Commitment 17, the table below outlines TikTok's performance in meeting the CoP's requirements.

Table 3. Analysis of Assessment Commitment 17: TikTok

Signatory: TikTok	
Section	Score
1. Quality and Completeness Documentation of media literacy actions	3- good
2. Evidence Provision of supporting data for listed actions	1- poor
3. Impact Impact of media literacy actions	4- good
4. Rating Effectiveness of media literacy actions	4- good

*Scale of scores for sections 1 and 2: 1- poor, 2- adequate, 3- good.

*Scale of scores for sections 3 and 4: 1- very poor, 2- poor, 3- fair, 4- good, 5- excellent.

4.2. Quality of response assessment

Quantitative data are predominant and do not fully capture the activities they report. Some of the listed actions have been supported by the number of impressions and clicks on search interventions by country. Beyond the metrics provided, it would be important to indicate the number of participants in the activities organised with partner organisations.

For the media literacy and critical thinking campaigns (A), no details are provided about their development, beyond mentioning the countries affected and the partner organisations in those related to the war in Ukraine. Regarding their impact, they include the number of impressions of the H5 Page, as well as impressions, clicks, and click-through rate for search interventions. Spain (with Maldita): H5 Page impressions (21,922); search intervention impressions (17,986,294); click-through rate (0.24%). Portugal (with Polígrafo): H5 Page impressions (10,369); search intervention impressions (5,426,533); click-through rate (0.42%).

The inclusion of links and screenshots enables a better assessment of election-related measures. For example, in the context of the local and presidential elections in Portugal (B), screenshots from the Election Centre are provided, including videos created in collaboration with Polígrafo.

Regarding the in-app guides (C) in the context of natural disasters and other rapidly evolving events, no specific data is provided (the link is broken).

Data on impressions, clicks, and click-through rate are included for specific in-app interventions (D), related to Holocaust misinformation/denial and elections. In the previous report, these were detailed across more thematic areas.

No detailed data is provided on the collaboration with the WHO's Fides network (E)—a linked profile contains only a single post from February 23, 2024—nor on the AI Literacy Fund (F)—the link provides limited information on this initiative, referring to collaboration with Mtoto News, Africa Check, and the Centre for Journalism Innovation and Development.

4.3. Verification of reported information

Overall, we must acknowledge limitations in properly assessing the effectiveness of the actions with the available data. Specifically, the report lacks detailed data for Spain and Portugal in most of the actions.

Although the number of impressions is high, it is difficult to assess it without contextual data on platform usage; for example, those related to Holocaust misinformation/denial reached 200 million impressions and 520k clicks, but only a 0.26% click-through rate. Regarding elections, H5 Page impressions totalled 21,992 in Spain and 10,369 in Portugal, indicating low impact.

The collaboration between TikTok and fact-checking and media literacy partners in temporary media literacy and election integrity campaigns represents an excellent combination of resources to combat disinformation and misinformation. Thanks to this collaboration, users receive rapid verification of information directly on the same screen where they search for news about regional elections. A similar effect can be observed in the partnership with the WHO's Fides network, which, in this case, provides health information. The AI Literacy Fund initiative makes it easier and more affordable for users to access basic training on artificial intelligence. Additionally, the Transparency Centre Global Elections Hub has the potential to become a reliable space for verifying and contrasting information.

Collaborations with fact-checkers and other organisations are described, and in relation to Iberifier Plus, Maldita (for the media literacy campaign in Spain) and Polígrafo (for local and presidential elections, as well as the media literacy campaign in Portugal) are specifically mentioned. Although there are no specific actions linked to the hub under Commitment 17, these partner organisations are participating.

4.4. Evaluation of actions taken

Although they provide some quantitative data suggesting a certain level of impact, there are also activities for which no impact data is available. Moreover, we reiterate

that assessing effectiveness based solely on these metrics remains limited in scope. The report refers to media literacy campaigns for which there is insufficient information—for example, the number of participants in off-platform activities, the volume of content produced by partner organisations in each country, and the actual impact of this content on users.

It would also be important to reinforce the value of collaboration with external stakeholders, such as local entities, civil society organisations, fact-checkers, and networks like the WHO's Fides initiative, by providing clearer, more consistent impact data. While these partnerships are presented as a key pillar of the actions, the report offers very limited evidence of their concrete outcomes, reach, or effectiveness at the local level.

4.5. Recommendations/need for improvement

- Reinforce collaboration with external stakeholders.
- Provide more contextual metrics or data to assess the impact of the actions described properly.
- It would be advisable to develop more initiatives that empower users in navigating the risks posed by AI-generated content.

5. Deep assessment for the signatory Facebook

5.1. Media Context: How did outlets in Portugal and Spain cover Facebook's commitment under Article 17 of the EU Code of Practice on Disinformation?

During the first months of 2026, Spanish-language media coverage of Facebook continued to portray the platform as a central infrastructure for the circulation of disinformation, political manipulation, fraud, public alarm and algorithmic amplification. Compared with the first half of 2025, when coverage focused heavily on Meta's decision to remove independent fact-checkers and weaken external verification, the 2026 news sample shows a broader set of risks: AI-generated political avatars, false rumours, migration and social-policy disinformation, climate and energy misinformation, foreign influence operations and Facebook's role in targeted political advertising.

Overall, Facebook is framed as part of a wider Meta ecosystem, alongside Instagram and WhatsApp, whose design facilitates the rapid diffusion of misleading content and whose safeguards remain insufficient. The coverage does not identify major Facebook-led media literacy initiatives in Spain or Portugal during the period analysed. Instead, media outlets highlight actions by fact-checkers, public authorities, civil society organisations, or sectoral actors to counter misinformation circulating on Facebook.

The evidence is therefore primarily drawn from Spanish-language outlets and agency coverage.

The most relevant headlines on Facebook in Portuguese and Spanish dailies

- “Presentación de una demanda para eliminar el perfil de inteligencia artificial ‘Doña María’, que critica a Lula y al gobierno” (*A lawsuit is filed to remove the AI profile ‘Doña María’, which criticises Lula and the government*). Source: CE Noticias Financieras, April 25th 2026. The article reports on a legal complaint filed by the PT, PV and PCdoB coalition against an AI-generated profile allegedly involved in early electoral campaigning and the spread of disinformation. The complaint requests the removal of profiles linked to the avatar on platforms including Facebook, Instagram, TikTok, YouTube, and X. The profiles allegedly simulate an elderly Black woman and publish political content against Lula, left-wing figures, and the PT. The case highlights risks linked to synthetic identities, political manipulation, undisclosed AI use and monetisation of misleading content.
- “No tengo por qué disculparme: hermana de Julio César Jasso habla tras ataque en Teotihuacán” (*“I do not have to apologise”: Julio César Jasso’s sister*

speaks after the Teotihuacán attack). Source: SDP Noticias, April 24th 2026. The article describes how the sister of the perpetrator of the Teotihuacán shooting used Facebook to denounce what she described as false and sensationalist information circulating about her family. She claimed that “90%” of the information shared was false. The article illustrates Facebook’s role as both a channel for direct public response and a space where misinformation around high-impact events can spread rapidly, generating harassment and reputational damage.

- “El certificado de vulnerabilidad solo sirve para la regularización, no da ayudas sociales” (*The vulnerability certificate is only valid for regularisation; it does not grant social benefits*). Source: Agencia EFE / EFEAgro, April 23rd 2026. The article debunks viral messages circulating on Facebook, Instagram and X claiming that immigrants who obtain a vulnerability certificate in Spain can automatically access social benefits such as the Minimum Living Income. EFE clarifies that the certificate applies only to the regularisation process established by Royal Decree 316/2026 and does not confer additional rights to social benefits. The article is relevant because it shows Facebook’s role in spreading misleading narratives about migration, welfare and public spending.
- “Las redes sociales europeas publican 44 noticias falsas sobre energía eólica cada día” (*European social networks publish 44 false stories about wind energy every day*). Source: Energías Renovables, April 21st 2026. The article reports on a study by WindEurope and CASM Technology that identifies a large, organised ecosystem of anti-wind disinformation in Europe. The study found more than 42,000 posts during the research period, with around 30,000 classified as disinformation or misinformation. Facebook and X are described as the most used networks for spreading false claims. The article links the problem to delays and cancellations of renewable energy projects and calls for media and digital literacy to build long-term resilience.
- “Patria V: la batalla por el relato propio en un ecosistema digital hostil” (*Patria V: the battle for one’s own narrative in a hostile digital ecosystem*). Source: La República, April 19th 2026. The article presents the digital space as an asymmetric battlefield where major platforms reward polarisation, disinformation and hate as vectors of visibility. It identifies Facebook, X and TikTok as channels through which campaigns of confrontation, discrimination and stigmatisation circulate. The piece also stresses the role of generative AI in producing hyper-realistic disinformation at an industrial scale and underlines the need for fact-checking, digital sovereignty and critical awareness.
- “La FGJES niega informes de un robo violento a una persona mayor en Guaymas y pide al público no difundir información falsa” (*The Sonora Prosecutor’s Office denies reports of a violent robbery against an elderly person in Guaymas and asks the public not to spread false information*). Source: CE

Noticias Financieras, April 18th 2026. The article reports that the Sonora Prosecutor's Office denied Facebook reports about an alleged violent robbery against an older adult. After the investigation, authorities found no evidence of the crime and urged the public to verify information with official sources before sharing it. The case shows Facebook's capacity to generate local alarm through unverified crime-related content.

- “De la Guerra Fría a Facebook: un libro analiza cómo se construye el relato público” (*From the Cold War to Facebook: a book analyses how the public narrative is constructed*). Source: Clarín, April 14th 2026. The article reviews Cristina Martín Jiménez's book *La tiranía de la mentira. Desinformación, censura y caos*. It describes Facebook as one of the central platforms in the contemporary information ecosystem, where algorithms shape what users see, amplify or hide messages, and influence trends, debates and perceptions. Although the article is not a direct report on Meta's compliance with the Code, it contributes to the broader framing of Facebook as an actor in the struggle over public narratives.
- “¿Murió Marco Antonio Solís?: la verdad tras el rumor que causó conmoción” (*Did Marco Antonio Solís die?: the truth behind the rumour that caused shock*). Source: Ámbito Financiero Online, April 12th 2026. The article debunks a viral rumour spread mainly on TikTok and Facebook claiming that Mexican singer Marco Antonio Solís had died. The false story originated in an AI-generated video that simulated a solemn news report, using an automated voice and fake media-style branding. The case illustrates how Facebook can amplify emotional disinformation and how synthetic content can be used to generate clicks, confusion and possible data-theft risks.
- “La verdad detrás del rumor sobre la supuesta muerte de Marco Antonio Solís” (*The truth behind the rumour about Marco Antonio Solís's alleged death*). Source: CE Noticias Financieras, April 12th 2026. This article summarises the same disinformation case, stressing that Facebook and TikTok posts misled thousands of users with manipulated AI content. It reinforces the need for users to consult reliable sources and avoid sharing unverified high-impact claims.
- “El arma secreta” (*The secret weapon*). Source: eldiario.es, April 10th 2026. The opinion article presents Facebook, Instagram, WhatsApp and X as tools of political disinformation and electoral warfare. It revisits the Cambridge Analytica scandal, Facebook's role in microtargeted political propaganda, and the later weakening of fact-checking systems. The piece argues that Meta's response has been ambivalent. While Facebook showed it could reduce the spread of false content during the 2020 US elections, Meta later shifted course, ending external fact-checking in the United States and replacing it with a Community Notes-style model.

- “The Cube: Cómo la propaganda rusa interviene en las elecciones húngaras” (*The Cube: how Russian propaganda interferes in the Hungarian elections*). Source: Euronews, April 10th 2026. The article describes a pro-Russian disinformation campaign targeting the Hungarian elections. It reports that false claims about opposition leader Péter Magyar were promoted through Facebook ads, including posts falsely suggesting he supported the return of compulsory military service. The article notes that Meta allows advertisers to target users by location or age and that the ads reached more than 20,000 people in Hungary before the page was removed. This case highlights risks linked to political advertising, microtargeting and foreign influence operations.
- “Rusia habría entrenado a 1.000 influencers en América Latina para difundir desinformación” (*Russia allegedly trained 1,000 influencers in Latin America to spread disinformation*). Source: CE Noticias Financieras, April 3rd 2026. The article reports on a network of Russian-trained communicators and influencers operating in Latin America. It notes that accounts associated with RT en Español and Sputnik Mundo have more than 18 million followers on Facebook and more than 6 million on YouTube. The report describes tactics such as narrative laundering, emotional manipulation, and the amplification of extremes to polarise societies and weaken trust in institutions.
- “Una red de desinformación busca desestabilizar a los países del Sahel”(A *disinformation network seeks to destabilise Sahel countries*). Source: Información, April 4th 2026. The article describes a network of fake accounts on platforms including TikTok, Facebook and X that impersonate media outlets and seek to destabilise countries in the Sahel. The report is relevant to Spain, particularly the Canary Islands, because of the region’s proximity to the Sahel and its connection to migration flows. It frames Facebook as part of the infrastructure used for geopolitical destabilisation and institutional distrust.

Main alerts on code violations

1. Insufficient control over AI-generated political manipulation
The “Doña María” case shows how AI-generated avatars can simulate real identities, circulate political content, monetise influence and potentially mislead voters. The fact that the complaint includes Facebook among the platforms on which the profile was active raises concerns about Meta’s ability to detect, label, and prevent synthetic political manipulation.
2. Persistent spread of false public-interest information
The analysed coverage includes false or misleading claims about migration regularisation, social benefits, violent crime, celebrity deaths, energy policy and international politics. These cases show Facebook’s role in amplifying information that can generate public alarm, social polarisation, reputational harm or distrust in institutions.

3. Risks linked to political advertising and foreign influence

The Euronews article on Hungary shows that Facebook ads can be used to disseminate false political narratives to targeted audiences. The reports on Russian influence networks in Latin America and the Sahel also show Facebook as part of a broader geopolitical disinformation infrastructure. These cases raise concerns about transparency, ad targeting and coordinated manipulation.

4. Algorithmic amplification of polarising content

Several articles describe Facebook and other social platforms as environments where algorithms reward attention, emotion, polarisation and confrontation. This dynamic runs counter to the spirit of Article 17, which requires platforms to help users understand the processes of disinformation and build resilience against manipulation.

5. Weak evidence of Facebook-led media literacy initiatives

The reviewed Factiva sample does not identify significant Facebook- or Meta-led media literacy campaigns in Spain or Portugal during the first months of 2026. Media literacy responses are instead attributed to external actors: fact-checkers, public authorities, researchers, WindEurope, CASM Technology and other civil society or sectoral organisations.

6. Climate and energy disinformation as a systemic risk

The WindEurope/CASM report identifies Facebook as one of the main networks used to spread false claims about wind energy. The article explicitly calls for digital and media literacy to strengthen long-term resilience. This connects directly with Article 17, as climate and energy misinformation can affect democratic debate, public policy and Europe's energy security.

7. Fact-checking deficit and external dependence

Many of the reviewed stories rely on external verification by EFE, Euronews, sectoral researchers or public authorities. This suggests that Facebook's own mechanisms are not perceived as sufficient by media outlets, especially after Meta's previously reported retreat from independent fact-checking.

Overall assessment

During the first months of 2026, Spanish-language media coverage portrays Facebook's performance under Article 17 as weak, opaque and largely reactive. The platform continues to appear as a major channel for false claims, manipulated narratives, synthetic political content, targeted propaganda, local rumours and foreign influence operations.

Unlike the 2025 coverage, where the main concern was Meta's removal of independent fact-checking mechanisms, the 2026 sample shows the consequences of a broader ecosystemic problem: misinformation on Facebook affects migration policy, public safety, climate and energy policy, electoral integrity, celebrity news, geopolitical conflict and public trust. The platform is not only a host for disinformation,

but also part of the algorithmic and advertising infrastructure that helps misleading content reach specific audiences.

The evidence of compliance with Article 17 remains limited. The analysed media coverage does not show major Facebook-led initiatives in Spain or Portugal aimed at promoting media literacy, educating users about AI-generated content, improving resilience to political manipulation or helping citizens identify misleading posts. Instead, public authorities, fact-checkers and civil society actors appear to carry most of the burden.

Overall, the press frames Facebook as a platform whose social and informational responsibilities remain insufficiently fulfilled. Its contribution to media literacy under Article 17 appears minimal in the Iberian media sample, while the risks associated with disinformation, AI manipulation, algorithmic amplification and targeted political advertising continue to grow. After examining how Spanish- and Portuguese-language outlets framed Commitment 17, the table below outlines Facebook's performance in meeting the requirements of the CoP.

Table 4. Analysis of Assessment Commitment 17: Facebook

<i>Signatory: Facebook</i>	
Section	Score
1. Quality and Completeness Documentation of media literacy actions	<i>1- poor</i>
2. Evidence Provision of supporting data for listed actions	<i>1- poor</i>
3. Impact Impact of media literacy actions	<i>1- very poor</i>
4. Rating Effectiveness of media literacy actions	<i>1- very poor</i>

**Scale of scores for sections 1 and 2: 1- poor, 2- adequate, 3- good.*

**Scale of scores for sections 3 and 4: 1- very poor, 2- poor, 3- fair, 4- good, 5- excellent.*

5.2. Quality of response assessment

The Meta report for Facebook provides less information and data than its predecessor, which had already fallen short of expectations. This report merely repeats, verbatim, most of the measures and implementations mentioned in the last report, using the same wording and making no updates.

The only addition refers to a "youth-focused campaign across eight EU countries, France, Italy, Belgium, Denmark, Germany, Spain, Ireland, and Greece, running from

late September through late November, to highlight support for parental approval for teens accessing online services", but fails to provide any data to evaluate its impact or to explain why this action empowers users.

The previous report described the actions taken to ensure the integrity of elections in several European countries within measure 17.2, but this report omits those actions and forwards the reader to the Election Report section, which is missing in Commitment 17. Therefore, the evaluation of this commitment lacks that information. Plus, however, it should be noted that Measure 17.2 pertains to "develop, promote and/or support or continue to run activities to improve media literacy and critical thinking such as campaigns to raise awareness about Disinformation, as well as the TTPs that are being used by malicious actors, among the general public across the European Union, also considering the involvement of vulnerable communities." That measure extends far beyond elections and would require attention beyond elections, which is the case. Finally, much of the report relies on broad statements of intent (e.g., "Our goal is to help people feel confident about the content and accounts that they interact with") rather than providing concrete, verifiable descriptions of actions. This limits the ability to properly assess the completeness of the reporting.

5.3. Verification of reported information

No data of any kind is supplied to support the few measures that were implemented and mentioned anew in this report. META says, "We were not able to deliver this SLI for this report." There is a lack of supporting evidence across the reported actions. The technological features (e.g., warning screens, verified badges, notification screens) are described in a very general, descriptive manner, without providing quantitative metrics (e.g., impressions, engagement, usage rates), qualitative assessments of effectiveness, or external references or links to further documentation. Similarly, although the report mentions collaborations with various organisations and experts, it does not explain how these partnerships operate in practice, nor does it provide any measurable outputs or results. The lack of data makes it difficult to assess the actual reach or effectiveness of the reported measures.

5.4. Evaluation of actions taken

The information provided is too limited to assess the impact. The report lacks both quantitative indicators and qualitative evidence necessary to evaluate whether the described tools, activities, and partnerships effectively improve media literacy or critical thinking among users. Without clearer evidence of outcomes, behavioural change, or user engagement, the actual impact of these measures remains unclear. Meta's campaigns, especially the Fraud & Scams and the upcoming Youth campaign, may help raise awareness and provide users with context. However, a lack of data on

user engagement and limited visibility on outcomes specific to Spain and Portugal undermines effectiveness.

Meta has deployed a wide range of tools that support media literacy by informing users about the credibility of content (e.g., verified badges and warnings for outdated articles), campaigns focused on fraud, and digital literacy have reached a broad audience. However, the lack of localised or interactive interventions limits their transformative potential in Iberian contexts. Also, tools like outdated-article notifications or verified badges support media literacy passively but are not integrated with impact measurement. For some initiatives, Facebook provides impact metrics; however, without more context, they seem insufficient.

5.5. Recommendations/need for improvement

Reporting remains extremely generic. The technologies described are largely part of Facebook's core platform architecture, rather than dedicated tools specifically designed to address disinformation or to empower users' media literacy skills. The report appears to selectively highlight existing features to justify compliance, rather than demonstrating a coherent, targeted media literacy strategy.

As explained before, there is no evidence provided on how these features improve users' critical thinking, whether users understand or engage with them, or what measurable outcomes they generate. As a result, their contribution to media literacy objectives appears minimal.

- The campaigns appear centralised and not adapted to local media literacy needs. It would be advisable to develop specific initiatives for each country.
- The HUB does not collaborate with META. META collaborates with actors in a dispersed fashion. For instance, they fund the Portuguese fact-checker *Polígrafo* through their Fact-checking network but do not mention this in the reporting.

6. Deep assessment for the signatory Instagram

6.1. Media context: How did outlets in Portugal and Spain cover Instagram's commitment under Article 17 of the EU Code of Practice on Disinformation?

During the first months of 2026, Spanish-language media coverage of Instagram continued to frame the platform as a relevant channel for visual, political, health-related and algorithmically amplified disinformation. Compared with the first half of 2025, when the main focus was Meta's withdrawal from independent fact-checking and the introduction of a Community Notes-style model, the 2026 coverage shows a broader set of risks: AI-generated political avatars, manipulated images and videos, health misinformation, gender disinformation, migration-related falsehoods and electoral manipulation.

The press portrays Instagram as part of the wider Meta ecosystem, alongside Facebook and WhatsApp, where disinformation circulates through visual formats, influencers, short videos, memes, and algorithmic feeds. Although some articles refer to public awareness and media literacy initiatives designed for Instagram's formats, these initiatives are generally led by external actors, such as medical organisations, fact-checkers, or civil society groups, rather than by Instagram itself.

No significant Portugal-specific coverage focused exclusively on Instagram was identified in the database Factiva. However, one article refers to disinformation in the Portuguese presidential elections and identifies Facebook and Instagram as the most used Meta platforms in that context. The evidence is therefore primarily drawn from Spanish-language outlets and agency coverage.

The most relevant headlines on Instagram in Portugal and Spanish dailies

- “Presentación de una demanda para eliminar el perfil de inteligencia artificial ‘Doña María’, que critica a Lula y al gobierno” (*A lawsuit is filed to remove the AI profile ‘Doña María’, which criticises Lula and the government*). Source: CE Noticias Financieras, April 25th 2026. The article reports on a legal complaint filed by the PT, PV and PCdoB coalition against an AI-generated profile accused of early electoral campaigning and spreading disinformation. The complaint requests the removal of profiles linked to the avatar on platforms including Instagram, TikTok, Facebook, YouTube, and X. The profiles allegedly simulate an elderly Black woman and publish political content against Lula, left-wing figures, and the PT. The case highlights the risks associated with synthetic identities, undisclosed AI use, political manipulation and monetised disinformation.

- “El certificado de vulnerabilidad solo sirve para la regularización, no da ayudas sociales” (*The vulnerability certificate is only valid for regularisation; it does not grant social benefits*). Source: Agencia EFE / EFEAgro, April 23rd 2026. The article debunks viral messages circulating on Facebook, Instagram and X claiming that immigrants who obtain a vulnerability certificate in Spain can automatically access social benefits. EFE clarifies that the certificate applies only to the regularisation process established by Royal Decree 316/2026 and does not grant access to any additional welfare benefits. The article shows how Instagram is used to disseminate misleading narratives about migration, welfare and public spending.
- “¿Saldrá de la cárcel? Pareja de Epa Colombia rompe el silencio tras rumores de su supuesta libertad” (*Will she be released from prison? Epa Colombia’s partner breaks her silence after rumours of her alleged release*). Source: Semana, April 17th 2026. The article reports on a wave of disinformation on social media claiming that influencer Daneidy Barrera Rojas, known as Epa Colombia, had been released from prison. Her partner used her official Instagram account to deny the rumours, explaining that old or manipulated images had been taken out of context. The case illustrates Instagram’s dual role as both a vector of rumours and a tool for direct correction by public figures.
- “Los médicos alertan a los jóvenes sobre los peligros de la desinformación en redes sociales: ‘Desconfía de soluciones mágicas’” (*Doctors warn young people about the dangers of misinformation on social media: ‘Be suspicious of magic solutions’*). Source: El Diario Vasco, April 14th 2026. The article reports on the Spanish Medical Association’s campaign #noteenREDES, aimed at warning young people about health misinformation on social media. The campaign includes four audiovisual pieces adapted to Instagram, TikTok and YouTube, designed to debunk viral content about miracle products, mental health pseudotherapies, unsupported diets and anti-vaccine narratives. This is one of the clearest examples of an Article 17-aligned media literacy action in the sample, although Instagram or Meta does not lead it.
- “OMC lanza la campaña ‘#noteenREDES, habla con tu médico’ para alertar de la desinformación sanitaria en redes sociales” (*The Spanish Medical Association launches the campaign ‘#noteenREDES, talk to your doctor’ to warn about health misinformation on social media*). Source: Infobae, April 7th 2026. The article presents the same OMC campaign as an audiovisual initiative adapted to Instagram’s language and formats. It encourages users to distrust viral videos, consult official sources, contrast information and speak to qualified medical professionals. The story directly connects social media literacy with public health protection.

- “No te dejes engañar por vídeos virales”: los médicos alertan sobre los peligros para la salud de las redes sociales (*‘Do not be fooled by viral videos’: doctors warn about the health dangers of social media*). Source: ABC, April 7th 2026. The article also covers the #noteenREDES campaign and highlights the risks posed by viral videos promoting miracle treatments, unsupported diets and pseudotherapies. Instagram is included among the platforms for which the campaign’s audiovisual messages were adapted. The article reinforces the idea that media literacy efforts are being driven by professional organisations rather than by Meta.
- “El arma secreta” (*The secret weapon*). Source: eldiario.es, April 10th 2026. This opinion article describes Facebook, Instagram, WhatsApp and X as tools of political disinformation and electoral warfare. It revisits Cambridge Analytica, microtargeted political propaganda and the use of social networks to exploit emotions and prejudices. The piece argues that, by 2026, disinformation has spread across platforms including YouTube, Instagram and TikTok, intensified by generative AI, deepfakes and weakened fact-checking systems.
- “La foto del embajador iraní con grupos del 23 de Enero es auténtica” (*The photo of the Iranian ambassador with groups from 23 de Enero is authentic*). Source: CE Noticias Financieras, April 5th 2026. The article reports on an investigation by Cazadores de Fake News that found several Instagram accounts belonging to a network called “La Fábrica de Desinformación” impersonated the fact-checking outlet to spread a debunked claim. These accounts claimed that real photos of the Iranian ambassador in Venezuela were AI-generated, using a fact-checker’s visual identity to mislead users. The case is particularly relevant because it shows disinformation actors exploiting the aesthetics of verification itself.
- “A foto do embaixador iraniano com grupos do 23 de Enero é autêntica” (*The photo of the Iranian ambassador with groups from 23 de Gener is authentic*). Source: CE Noticias Financieras Portuguese, April 5th 2026. The Portuguese-language version of the same report also identifies Instagram accounts as part of the disinformation network. It reinforces the concern that Instagram can be used not only to circulate false content, but also to impersonate legitimate verification actors.
- “El panameño se informa con el celular, el algoritmo y, a veces, con la verdad” (*Panamanians get informed through their phones, algorithms and, sometimes, the truth*). Source: La Prensa, April 2nd 2026. The article states that Instagram and WhatsApp are among the main digital information channels in Panama. It warns that algorithmic feeds reinforce biases, reduce exposure to contrasting views and make users vulnerable to manipulated videos and AI-generated electoral disinformation. Although the case is not Iberian, it is relevant to Instagram’s Article 17 responsibilities because it directly links the platform to

algorithmic information consumption and the need for verification habits.

- “Sora, la IA que escaló la desinformación, anuncia su cierre” (*Sora, the AI that escalated disinformation, announces its shutdown*). Source: El Heraldó, April 1st 2026. The article argues that Sora’s realistic AI-generated videos created major risks for information integrity on social media, especially Facebook, Instagram and TikTok. It describes cases where deepfake-style videos were interpreted as real events before they could be verified. The article calls for mandatory labelling, digital literacy, automated verification tools and clear regulatory frameworks.
- “Redes, al banquillo” (*Social networks in the dock*). Source: El Tiempo, April 1st 2026. The editorial states that platforms such as Facebook, Instagram and X have democratised communication but also exacerbated polarisation and disinformation. It also discusses court rulings in the United States that place greater emphasis on platform responsibility for the effects of their design on users, especially minors. The article frames social media regulation as necessary to prevent harmful effects from outweighing the benefits of technological innovation.
- “Towanda Rebels: ‘La desinformación de género es una estrategia de descrédito hacia las mujeres que cuestionan el orden establecido’” (*Towanda Rebels: ‘Gender disinformation is a strategy to discredit women who question the established order’*). Source: El Comercio, March 11th 2026. The article reports on a conference in Gijón about equality, disinformation and hoaxes. The speakers warn that young people increasingly obtain political and social information through Instagram, described as the main source for 47% of young people in Spain. They argue that Instagram’s compressed formats make it difficult to understand issues in depth and facilitate the spread of gender disinformation aimed at discrediting women in public life.
- “Expertos indican que se propagan imágenes falsas y vídeos manipulados tras la captura de El Mencho, generando miedo y confusión” (*Experts say false images and manipulated videos spread after El Mencho’s capture, generating fear and confusion*). Source: CE Noticias Financieras, February 23rd 2026. The article reports that false images and manipulated videos circulated on Instagram and TikTok after the death of Nemesio Oseguera Cervantes, “El Mencho”. Experts estimated between three and five million potential exposures to false or unverified content in the first hours after the event. The article stresses the need for media literacy, verification and prudence in high-sensitivity news scenarios.
- “La desinformación sobre las elecciones presidenciales alcanza niveles altos” (*Disinformation about the presidential elections reaches high levels*). Source: CE Noticias Financieras, January 20th 2026. The article reports that during the

Portuguese presidential elections, more than 8.3 million views of disinformation were detected on social media during the first round. The majority of cases were identified on Meta platforms, with Facebook and Instagram as the most used. This is the most directly Portugal-related item in the sample and raises concerns about Instagram's role in electoral disinformation.

Main alerts regarding code violations

1. **AI-generated political manipulation and synthetic identities**
The “Doña María” case shows how AI-generated avatars can simulate real social identities, circulate political messages and potentially mislead audiences. Instagram's inclusion among the platforms targeted by the legal complaint raises concerns about Meta's capacity to detect, label and limit synthetic political manipulation.
2. **Spread of migration and welfare misinformation**
The EFE fact-check on the vulnerability certificate shows that Instagram was used to circulate misleading claims about immigration regularisation and access to social benefits in Spain. This type of content can feed xenophobic narratives and institutional distrust, making user education and source verification essential.
3. **Health misinformation and viral pseudoscience**
Several articles on the #noteenREDES campaign identify Instagram as one of the platforms where young users encounter viral misinformation about health, diet, mental health and pseudotherapies. Although the campaign aligns with the goals of Article 17, it is led by the Spanish Medical Association, not by Instagram.
4. **Manipulated images and deepfake risks**
The articles on Sora, El Mencho and the false claims about the Iranian ambassador show that Instagram remains vulnerable to manipulated visual content, deepfakes and false AI-detection narratives. The use of fake verification aesthetics by “La Fábrica de Desinformación” is especially problematic because it can erode trust in legitimate fact-checkers.
5. **Gender disinformation and harm to public participation**
The El Comercio article on Towanda Rebels highlights Instagram as a key source of political and social information among young people in Spain. It identifies gender disinformation as a strategy to discredit women who challenge established power structures. This connects Instagram's Article 17 obligations with equality, democratic participation and digital safety.
6. **Electoral disinformation and algorithmic influence**
The Portuguese presidential-election item and the Panama-focused article both frame Instagram as a central information channel in electoral contexts. The concern is not only the existence of false content, but also the algorithmic ranking systems that reinforce biases, reduce exposure to alternative viewpoints and accelerate the spread of emotionally charged narratives.

7. Absence of Instagram-led media literacy initiatives in Spain and Portugal

The reviewed sample does not identify any major Instagram- or Meta-led media literacy campaign in Spain or Portugal during the first months of 2026. Media literacy actions are instead attributed to external organisations, including the OMC, fact-checkers, journalists and public institutions. This weakens evidence of substantive compliance with Article 17.

8. Continued consequences of weakening independent fact-checking

The 2026 coverage continues to reflect the consequences of Meta's earlier retreat from external verification. Several articles rely on fact-checkers or public authorities to correct misinformation circulating on Instagram, suggesting that platform-level safeguards are not perceived as sufficient.

Overall assessment

During the first months of 2026, Spanish- and Portuguese-language media coverage portrays Instagram's performance under Article 17 as insufficient, fragmented and largely dependent on external actors. The platform is repeatedly associated with visual disinformation, AI-generated manipulation, political rumours, health misinformation, gender disinformation and misleading narratives about migration and public policy.

Compared with the 2025 period, the concern has shifted from Meta's institutional retreat from independent fact-checking to the practical consequences of that weaker verification environment. The analysed articles show misinformation circulating on Instagram in multiple formats: reels, images, short videos, out-of-context old content, synthetic avatars, and fake fact-checking posts.

There are some positive media literacy responses, especially the OMC's #noteenREDES campaign, which is adapted to Instagram's audiovisual language and encourages young users to consult official sources, distrust miracle claims and verify information with professionals. However, this initiative is not led by Instagram or Meta. The same pattern appears in other cases: fact-checkers, medical organisations, journalists and authorities are responsible for correcting or contextualising misinformation after it has circulated.

Overall, the press frames Instagram as a platform with significant Article 17 responsibilities but limited visible compliance in Spain and Portugal. Its contribution to media literacy appears weakly documented, while the risks associated with AI-generated content, algorithmic amplification, and the erosion of trust in verification continue to grow. After examining how Spanish- and Portuguese-language outlets framed Commitment 17, the table below outlines Instagram's performance in meeting the CoP's requirements.

Table 5. Analysis of Assessment Commitment 17: Instagram

Signatory: Instagram	
Section	Score
1. Quality and Completeness Documentation of media literacy actions	1- poor
2. Evidence Provision of supporting data for listed actions	1- poor
3. Impact Impact of media literacy actions	1- very poor
4. Rating Effectiveness of media literacy actions	1- very poor

*Scale of scores for sections 1 and 2: 1- poor, 2- adequate, 3- good.

*Scale of scores for sections 3 and 4: 1- very poor, 2- poor, 3- fair, 4- good, 5- excellent.

6.2. Quality of response assessment

Overall, the report is of low quality, primarily due to insufficient detail in the activities described and a lack of innovation compared to previous reports. For instance, the content on 17.1 and 17.3 is the same as in the previous Meta report, even though the links are not working (September 2026), except that the previous reports included some metrics. Also, it relies heavily on general statements of intent and high-level descriptions rather than providing clear, structured, and verifiable information about the actions implemented. Instagram (Meta) describes a range of media literacy-related actions; however, most of these initiatives appear to be long-standing and were not specifically developed or updated during the current reporting period. This makes it difficult to assess progress over time. In addition, several activities (such as youth-focused campaigns) are only briefly mentioned, with insufficient detail on their scope, implementation, or concrete outcomes. What is more, its contribution to improving users' critical thinking or ability to identify misinformation is unclear.

There are no indicators showing user awareness of these features, behavioural change, or effectiveness in reducing misinformation exposure or engagement. Additionally, given Instagram's strong visual and algorithm-driven environment, the report does not address how these tools are adapted to platform-specific risks.

6.3. Verification of reported information

Instagram does not provide any metrics; in fact, the report explicitly states that no SLI data has been delivered, severely limiting the ability to assess the scale and effectiveness of these tools. There is a significant lack of supporting evidence for the measures described. The tools highlighted (e.g., warning screens, verified badges) are presented descriptively, without quantitative data (e.g., impressions, reach, engagement), nor are there user interaction metrics or comparative data across Member States.

Furthermore, although collaborations with external partners and experts are mentioned, the report does not clarify the nature of these partnerships, their operational role, or any measurable outputs derived from them. This absence of evidence makes the assessment of impact largely speculative.

Regarding, for instance, actions taken during electoral processes, Instagram reports that it “proactively point users to reliable information on the electoral process through in-app ‘Election Day Information’”; however, we couldn’t find any data or metrics to assess the effectiveness of this action. What is more, Instagram directs the reader to the “election chapter” for more information, but we couldn't find it.

6.4. Evaluation of actions taken

Instagram/Meta documents several actions under Commitment 17: its first Media Literacy Annual Plan (21 July 2025), a youth well-being campaign in Ireland, France, Spain, Italy and the Netherlands, an anti-scam awareness campaign, and a youth-focused campaign on parental approval for teens accessing online services in eight EU countries (France, Italy, Belgium, Denmark, Germany, Spain, Ireland and Greece). It also refers to long-standing tools such as warning screens and verified badges, as well as broad partnerships relevant to media literacy. However, against the D5.4 criteria (clarity, specificity, breadth/depth, transparency and partnerships), the documentation remains only partially adequate. The report provides limited operational detail, does not clearly distinguish new measures from long-standing platform features, does not document deployment by Member State in this section, and offers less concrete detail than previous Meta/Instagram reports.

The youth well-being, anti-scam and parental approval campaigns are relevant to vulnerable groups, including young users and families. Still, the report does not provide reach, engagement, participant numbers, user feedback or Spain-specific impact data.

6.5. Recommendations/need for improvement

The overall rating is based on the absence of verifiable outcomes and measurable evidence of impact. The actions described have conceptual relevance, but the lack of evidence on user awareness, behavioural change, reduced exposure to misinformation or expert-led impact measurement means their effectiveness should be rated as poor. The tools described are primarily part of Instagram's standard platform infrastructure, rather than dedicated media literacy tools. The report appears to repurpose existing features to demonstrate compliance with the commitment, without presenting a clear, user-centred empowerment strategy to improve critical thinking or media literacy skills.

- We presume there are specific collaborations with members of Iberifier or other local hubs; however, it is not specified. It would be necessary to include this information in the report.
- It would be necessary to include quantitative data on how users interact with the developed features or campaigns, including contextual information, to assess their impact (for example, the rate of users who interact with warning screens and the rate of users who do not).
- Warning screens and verified badges may contribute to safety, trust, and contextual awareness, but their direct contribution to critical thinking or resilience to misinformation has not been demonstrated. It would be necessary to provide further explanation on how it improves media literacy and some metrics related to its impact.
- The report also gives limited attention to Instagram-specific risks, such as visual misinformation, Reels, Stories and influencer-driven content; some specific initiatives would be advisable.

7. Deep assessment for the signatory on LinkedIn

7.1. Media context: How did outlets in Portugal and Spain cover LinkedIn's commitment under Article 17 of the EU Code of Practice on Disinformation?

During the first months of 2026, Spanish-language media coverage of LinkedIn's commitment under Article 17 of the EU Code of Practice on Disinformation was very limited. Unlike platforms such as YouTube, Google Search or TikTok, LinkedIn did not receive sustained media attention in Spain or Portugal in relation to media literacy, disinformation, or user-awareness initiatives.

The available coverage mainly frames LinkedIn as part of a broader group of large technology platforms that, according to the European Fact-Checking Standards Network, have started to retreat from key fact-checking commitments under the EU Code of Practice on Disinformation. The reporting does not identify any specific LinkedIn-led media literacy campaigns in Spain or Portugal during the period analysed. This suggests that LinkedIn's implementation of Article 17 remains largely invisible in the public media debate.

No significant Portugal-specific LinkedIn coverage was identified in the database Factiva. The evidence is therefore based on Spanish-language agency coverage.

The most relevant headlines on LinkedIn in Portuguese and Spanish dailies

- “Las plataformas combaten menos la desinformación tras la vuelta de Trump, según un informe” (Platforms are doing less to combat disinformation after Trump's return, according to a report). Source: Agencia EFE – Servicio Internacional, March 3rd 2026. The article reports that the European Fact-Checking Standards Network warned that major technology platforms had begun to step back from their commitment to combat disinformation, particularly after Donald Trump's return to the White House. The organisation, of which EFE Verifica is a member, explicitly mentions Google Search, Bing, YouTube and LinkedIn as platforms that have distanced themselves from “important *fact-checking commitments*” included in the EU Code of Practice on Disinformation. The article also links this retreat to the wider political context, noting that it coincides with attacks by the US administration on EU digital rules, which critics present as harmful to American platforms. The report further warns that the rapid integration of generative AI has enabled the industrial-scale automation of misleading content, increasing risks for democratic debate, public health, economic decision-making and the sustainability of quality journalism. For LinkedIn, the article is relevant because it places the platform within the group of services whose commitment to

independent verification and disinformation control is being questioned at the EU level. However, the news does not provide details about LinkedIn-specific practices, professional misinformation, fake profiles, job scams, or media literacy tools.

The lack of results is itself relevant for the assessment of Article 17, because it suggests that LinkedIn's media literacy and public-awareness efforts were either absent, not newsworthy, or not communicated effectively in Spain and Portugal.

The main alerts on possible code violations

1. Retreat from verification commitments

The most important alert comes from the EFE article, which cites the European Fact-Checking Standards Network. LinkedIn is explicitly included among the platforms that have moved away from important fact-checking commitments under the EU Code of Practice on Disinformation. This raises concerns regarding LinkedIn's practical alignment with the Code, particularly in relation to verification cooperation and transparency.

2. Very limited public visibility of Article 17 actions

The reviewed media coverage does not mention any LinkedIn-led media literacy initiative in Spain or Portugal during the first months of 2026. No campaigns, educational resources, public-awareness partnerships, or user training programmes are identified. This suggests weak visibility of Article 17 implementation and makes it difficult to assess whether LinkedIn is taking active steps to strengthen users' resilience against misinformation.

3. Absence of Spain- or Portugal-specific reporting

The Factiva database does not include Portugal-specific coverage and contains only one substantive Spanish-language article that mentions LinkedIn in the context of EU disinformation commitments. This scarcity of coverage contrasts with the broader attention received by platforms such as YouTube or Google Search. It suggests that LinkedIn remains peripheral in the Iberian public debate on disinformation, despite its relevance as a professional information network.

4. Insufficient transparency on professional misinformation

The available article does not provide evidence of LinkedIn-specific transparency measures, moderation criteria, reporting systems or cooperation with fact-checkers. This absence is important because misinformation on LinkedIn may take specific forms, including false professional credentials, fraudulent job offers, misleading corporate claims, investment scams, reputational manipulation and AI-generated professional profiles.

5. Risks associated with generative AI and automated content

Although the EFE article does not focus exclusively on LinkedIn, it warns that generative AI enables the large-scale automation of misleading content. This is particularly relevant for LinkedIn because professional networks may be exposed to AI-generated profiles, synthetic professional identities, automated influence campaigns and misleading business-related content. The lack of media coverage on LinkedIn-specific responses to these risks suggests a potential gap in public accountability.

6. Formal compliance but weak evidence of substantive Article 17 implementation
The analysed coverage does not deny LinkedIn's formal participation in the EU Code of Practice, but it provides no evidence of concrete media literacy actions. From an Article 17 perspective, this indicates a pattern of formal or institutional compliance without visible educational impact in Spain or Portugal.

Overall assessment

During the first months of 2026, Spanish- and Portuguese-language media coverage portrays LinkedIn's performance under Article 17 as largely invisible and insufficiently documented. The only substantive article identified in the database Factiva sample mentions LinkedIn as part of a wider group of platforms that have retreated from important fact-checking commitments under the EU Code of Practice on Disinformation.

Unlike YouTube or Google Search, LinkedIn does not appear in the reviewed press as a platform actively engaged in public media literacy initiatives, awareness campaigns, user education or cooperation with civil society organisations in Spain or Portugal. The absence of such coverage makes it difficult to identify positive evidence of compliance with Article 17.

The main concern is therefore not only possible non-compliance, but opacity. LinkedIn's role in the disinformation ecosystem receives little media scrutiny, despite the specific risks associated with professional misinformation, fake profiles, fraudulent job content, misleading corporate communication and AI-generated identities.

Overall, the press coverage suggests that LinkedIn's Article 17 performance remains weakly visible, weakly evidenced and potentially incomplete. The platform appears formally connected to the EU disinformation framework. Still, the available news does not show concrete, locally visible efforts to promote media literacy or strengthen users' critical capacity in Spain and Portugal.

7.2. Statements from LinkedIn for not endorsing Commitment 17

LinkedIn withdrew from Commitment 17 under the new agreement signed in January 2025. As a result, it is not possible to assess the effectiveness of actions to empower users through media literacy measures.

9. Conclusions

This fourth assessment confirms a persistent and structural weakness in the implementation of Commitment 17 of the Strengthened Code of Practice on Disinformation. Across the VLOPs and VLOSEs analysed, the information provided by platforms remains insufficient to demonstrate that their actions are producing meaningful, measurable and citizen-centred improvements in media literacy, AI literacy or user empowerment. This is not an isolated finding of the present reporting cycle. It is a recurrent pattern already identified in Deliverables 5.1, 5.2 and 5.3 and confirmed again in Deliverable 5.4. Despite the strengthened regulatory context created by the Digital Services Act and the recognition of the Code as a Code of Conduct under Article 45 DSA, platform reporting continues to fall short of what is needed for robust public accountability.

The main conclusion is clear: most platforms still report the existence of initiatives rather than their effectiveness. They describe tools, campaigns, partnerships, labels, information panels, safety hubs or educational resources, but they rarely provide sufficient evidence that these measures are visible to citizens, used by citizens, understood by citizens, or capable of changing citizens' behaviour when confronted with disinformation. In this sense, the reports remain largely declaratory. They show activity, but not impact. They show outputs, but not outcomes. They show availability, but not usability.

This limitation is particularly serious in relation to Commitment 17, whose purpose is not merely to document that platforms have launched media literacy initiatives, but to assess whether users are being effectively empowered to recognise, understand, contextualise and resist disinformation. The assessment in this deliverable shows that the majority of reported actions lack adequate information on reach, targeting, user interaction, behavioural impact, long-term effectiveness, or relevance to specific national contexts, including Spain and Portugal. The absence of this information prevents independent evaluators, regulators and civil society from determining whether these initiatives meaningfully contribute to democratic resilience.

A second major conclusion concerns the quality of the evidence provided. In general, platforms continue to rely on quantitative indicators that are too generic to support a serious evaluation of effectiveness. Metrics such as impressions, views, visits, clicks, the number of resources published, or the number of campaign participants are not sufficient on their own. They do not explain whether users understood the content, whether they changed their verification practices, whether vulnerable groups were reached, whether users were exposed to the initiative at the moment of risk, or whether the initiative reduced susceptibility to misinformation. Engagement metrics, when not contextualised or compared with relevant baselines, provide only a superficial account of implementation.

This is especially problematic because platforms possess far more granular data than what they disclose in their reports. They can technically determine which users are exposed to specific labels, prompts, warnings, or educational resources; how users interact with them; whether they ignore, click, dismiss, or act upon them; and whether these interventions affect subsequent behaviour. However, the reports do not provide this level of detail. As a result, the assessment of Commitment 17 remains dependent on partial, platform-selected evidence that is insufficient for a comprehensive evaluation of real-world effectiveness.

A third conclusion is that many platforms still treat media literacy as an external or peripheral activity rather than a core design obligation. Many initiatives are presented as campaigns, partnerships, or educational resources that lie outside the ordinary user experience. While such actions may be valuable, they cannot replace platform-native interventions embedded in search results, feeds, recommender systems, video interfaces, image-sharing tools, AI-generated responses, advertising libraries and content-sharing flows. Media literacy measures must be present wherever disinformation is encountered, not only in separate educational spaces that users must actively seek out.

The analysis also shows that platform reports rarely explain how media literacy tools are promoted to users or how citizens are expected to find them. A tool that exists but is hidden, poorly explained, weakly localised, available only in a limited number of languages, or disconnected from everyday platform use cannot be considered an effective empowerment measure. Usability is therefore not a secondary issue. It is a substantive condition of compliance. If citizens cannot easily find, understand, and use these tools, they do not fulfil the objectives of Commitment 17.

A fourth conclusion concerns artificial intelligence. The reporting period confirms that AI-generated and AI-mediated disinformation has become one of the most urgent challenges for user empowerment. Platforms increasingly provide labels, provenance signals, or detection tools that inform users whether certain content may have been generated or altered by AI. However, these measures remain insufficient. Citizens not only need to know that content may have been produced by AI; they also need to understand the different levels of risk associated with different uses of AI.

For example, using AI to generate an illustrative image for explanatory purposes is not equivalent to using AI to fabricate a public figure's statement, impersonate a real person, produce a synthetic political avatar, manipulate electoral content, generate wrong medical advice, or create a realistic video of an event that never happened. Current platform reporting does not show that users are being systematically educated about these distinctions. This gap is central. AI literacy must become a mandatory component of media literacy under Commitment 17, not an optional or fragmented add-on.

A fifth conclusion is that the platforms' reported initiatives remain insufficiently localised. The reports often refer to EU-wide, global or multi-country initiatives without clearly specifying what was implemented in Spain and Portugal, in which languages, through which partners, with which target groups, and with what measurable results. This weakens the value of the reports for national and regional assessment. Disinformation dynamics are context-dependent. Electoral disinformation, health misinformation, migration narratives, gender disinformation, climate misinformation and AI-generated manipulation do not affect all countries, languages or communities in the same way. Commitment 17 requires implementation that is sensitive to these national and linguistic differences.

The media context analysed in this deliverable reinforces this concern. News coverage in Spain and Portugal continues to identify substantial risks linked to platform environments: AI-generated political manipulation, health misinformation, misleading search results, synthetic video, fake profiles, electoral disinformation, migration-related falsehoods, gender disinformation, scams, and the amplification of polarising or misleading content. In several cases, the most visible corrective actions came from external actors, fact-checkers, medical associations, public authorities, journalists, civil society organisations or educational institutions, rather than from the platforms themselves. This suggests that platforms are still relying too heavily on third parties to compensate for weaknesses in their own systems.

A sixth conclusion is that platforms continue to shift excessive responsibility onto users. Some tools and labels may help users make better decisions, but they cannot be a substitute for platform accountability. The burden of identifying false, manipulated, or harmful content cannot be placed on citizens without providing them with meaningful context, accessible explanations, effective reporting tools, visible corrections, and clear routes for redress. User empowerment must not become a rhetorical justification for reducing professional verification, weakening independent fact-checking, or replacing expert assessment with crowd-based mechanisms of uncertain reliability.

This is particularly relevant in light of the reported retreat from independent verification commitments on some platforms. Where professional fact-checking is reduced, discontinued or replaced by user-based systems, media literacy obligations become even more important. Yet the evidence reviewed in this deliverable does not show that platforms have compensated for such changes with robust, systematic, localised and independently assessable user-empowerment measures. This creates a serious accountability gap.

A seventh conclusion concerns vulnerable groups. Platform reporting does not provide sufficient information on whether media literacy and AI literacy initiatives effectively reach those most exposed to harm: minors, elderly users, migrants, users with lower digital skills, people consuming health information, users exposed to political manipulation, and communities targeted by hate, scams or gendered disinformation.

Without data on targeting criteria, exposure, accessibility and outcomes for these groups, the reported initiatives cannot be considered adequate from a public-interest perspective.

The repeated lack of transparency also affects the assessment of platform design. Several risks identified during the reporting period are not only content-related but architecture-related: recommender systems, autoplay, infinite scroll, algorithmic ranking, advertising microtargeting, search summaries, AI-generated answers, viral short-video formats and closed or semi-closed sharing environments. Yet platform reports under Commitment 17 rarely explain how media literacy measures interact with these design features. This is a significant omission. Citizens cannot be empowered only through educational messages if the surrounding interface continues to reward speed, emotional intensity, polarisation and passive consumption.

Taken together, these findings lead to a firm conclusion: the current level of reporting under Commitment 17 is not adequate for the post-DSA phase of platform accountability. The European Commission, national coordinators and independent evaluators require more than descriptive reporting. They require evidence that is granular, comparable, independently verifiable and relevant to citizens' actual experience on the platforms.

Future reporting should therefore require, at minimum, the following improvements: country-level data for Spain and Portugal; clear descriptions of target audiences; information on how users are selected or exposed to each initiative; visibility and usability indicators; interaction data beyond impressions; evidence of behavioural impact; independent evaluation results; language and accessibility information; specific AI literacy measures; and explanations of how interventions are integrated into high-risk platform areas such as recommender systems, AI-generated outputs, political content, health information, advertising and visual media.

The final assessment of Deliverable 5.4 is therefore unequivocal. While some platforms have reported relevant initiatives and, in some cases, useful tools or partnerships, the overall standard remains far below the level required to demonstrate effective compliance with Commitment 17. The reports continue to provide poor, incomplete and insufficiently contextualised information. They do not allow a reliable assessment of whether citizens in Spain and Portugal are better equipped to understand, detect and resist disinformation as a result of the actions described.

For this reason, IBERIFIER concludes that the implementation of Commitment 17 remains incomplete, uneven and insufficiently accountable. The gap between platform declarations and citizens' real needs remains substantial. At this stage of the Code's implementation, and especially after its integration into the DSA framework, platforms should no longer be assessed on whether they can list media literacy initiatives, but on whether they can prove that these initiatives work.

10. Bibliography

Botan, M., Meyer, T. (June, 2025). *Implementing the EU Code of Practice on Disinformation: An evaluation of VLOPSE compliance and effectiveness (Jan–Jun 2024)*. European Digital Media Observatory (EDMO).<https://edmo.eu/publications/implementing-the-eu-code-of-practice-on-disinformation-an-evaluation-of-vlopse-compliance-and-effectiveness-jan-jun-2024/>

Cardoso, G., Paisana, M., & Pinto-Martinho, A. (2025). *Digital News Report Portugal 2025*. Publicações OberCom. <https://doi.org/10.5281/zenodo.15599071>

Cesarini, P., Ginsborg, L., Gori, P., Maggi, E., & Nijmeijer, R. (May 2024). *Platforms' implementation of the CoP commitments on media literacy, research and fact-checking*. European Digital Media Observatory (EDMO).
<https://edmo.eu/wp-content/uploads/2024/06/EDMO-CoP-Report.pdf>

EDMO (October 2024). *Guidelines for Effective Media Literacy Initiatives*. European Digital Media Observatory (EDMO).
https://edmo.eu/edmo-news/establishment-of-edmo-working-group-media-literacy-standards-and-best-practices/?utm_source=chatgpt.com

European Commission. (2025a). *The Code of Conduct on Disinformation*. Shaping Europe's Digital Future. Brussels: EC.
<https://digital-strategy.ec.europa.eu/en/library/code-conduct-disinformation>

European Commission. (2025b). *DSA: Code of Practice on Disinformation*. Press release, 13 February 2025. Brussels: EC.
https://ec.europa.eu/commission/presscorner/detail/en/ip_25_505

European Commission. (2025c). *Delegated act on data access under the Digital Services Act*. Shaping Europe's Digital Future. Brussels: EC.
<https://digital-strategy.ec.europa.eu/en/library/delegated-act-data-access-under-digital-services-act-dsa>

European Commission. (2026a). *Code of Conduct on Disinformation: Signatories publish their latest reports in the Code's Transparency Centre*. Shaping Europe's Digital Future. Brussels: EC.
<https://digital-strategy.ec.europa.eu/en/news/code-conduct-disinformation-signatories-publish-their-latest-reports-codes-transparency-0>

European Commission. (2026b). *Codes of conduct under the Digital Services Act*. Shaping Europe's Digital Future. Brussels: EC.
<https://digital-strategy.ec.europa.eu/en/policies/dsa-codes-conduct>

European Commission. (2026c). *How the Digital Services Act enhances transparency online*. Shaping Europe's Digital Future. Brussels: EC.
<https://digital-strategy.ec.europa.eu/en/policies/dsa-brings-transparency>

European Commission. (2026d). *Draft Commission guidelines on trusted flaggers*. Shaping Europe's Digital Future. Brussels: EC.
<https://digital-strategy.ec.europa.eu/en/library/draft-commission-guidelines-trusted-flaggers>

European Commission. (2026e). *Guidelines for teachers and educators on tackling disinformation and promote digital literacy through education and training*. European Education Area. Brussels: EC.
<https://education.ec.europa.eu/focus-topics/digital-education/actions/plan/guidelines-for-teachers-to-foster-digital-literacy-and-tackle-disinformation>

European Commission. (2026f). *Commission investigates Grok and X's recommender systems under the Digital Services Act*. Press release, 26 January 2026. Brussels: EC.
https://ec.europa.eu/commission/presscorner/detail/en/ip_26_203

European Commission. (2026g). *Commission designates WhatsApp as Very Large Online Platform under the Digital Services Act*. Digibyte, 26 January 2026. Brussels: EC.
<https://digital-strategy.ec.europa.eu/en/news/commission-designates-whatsapp-very-large-online-platform-under-digital-services-act>

European Commission. (2026h). *Commission preliminarily finds TikTok's addictive design in breach of the Digital Services Act*. Press release, 6 February 2026. Brussels: EC. https://ec.europa.eu/commission/presscorner/detail/en/ip_26_312

European Commission. (2026i). *Commission preliminarily finds Meta in breach of Digital Services Act for failing to prevent minors under 13 from using Instagram and Facebook*. Press release, 29 April 2026. Brussels: EC.
<https://digital-strategy.ec.europa.eu/en/news/commission-preliminarily-finds-meta-breach-digital-services-act-failing-prevent-minors-under-13>

European Fact-Checking Standards Network (EFCSN) (2024). *Fact-checking and related risk-mitigation measures for disinformation in the Very Large Online Platforms and Search Engines*. European Fact-Checking Standards Network.
<https://www.efcsn.eu/publication/vlop-risk-mitigation-report-2024/>

European Parliament and Council. (2022). *Regulation (EU) 2022/2065 on a Single Market for Digital Services and amending Directive 2000/31/EC — Digital Services Act*. Official Journal of the European Union. Brussels: EC.
<https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng>

Moreno-Castro, C., Paisana, M., Serra-Perales, A., Vázquez-Herrero, J., Azurmendi, A., Moreno, J. C., Calvo-Miguel, D., Negreira-Rey, M.C., Quian, A., Pérez-Seijo, S., Rodríguez-Vázquez, A.I., Silva-Rodríguez, A., Sixto-Rodríguez, J., Vizoso-García, Á., Iranzo-Cabrera, M., & Salaverría, R. (2024). *IBERIFIER - Assessment of VLOPs (Very Large Online Platforms) and VLOSEs (Very Large Online Search Engines)*. Signatory

Reports for the Strengthened Code of Practice on Disinformation. Report No 1.
Pamplona. IBERIFIER. <https://doi.org/10.5281/zenodo.15309350>

Moreno-Castro, C., Paisana, M., Falcó-Moreno, N., Vázquez-Herrero, J., Azurmendi, A., Calvo-Miguel, D., Pérez-Seijo, S., Moreno, J. C., Cano-Orón, L., Negreira-Rey, M.-C., Rodríguez-Vázquez, A.-I., Sixto-García, J., Cabrera García-Ochoa, Y., Vizoso-García, Á., Sánchez-Hualde, A., & Salaverría, R. (2025a). *IBERIFIER - Assessment of VLOPs (Very Large Online Platforms) and VLOSEs (Very Large Online Search Engines). Signatory Reports for the Strengthened Code of Practice on Disinformation. Report No. 2.* IBERIFIER. <https://doi.org/10.5281/zenodo.15753700>

Moreno Castro, C., Paisana, M., Serra-Perales, A., Azurmendi, A., Vázquez Herrero, J., Calvo, D., Moreno, J., Cano-Orón, L., Negreira Rey, M. C., Sixto-García, J., Cabrera García-Ochoa, Y., Iranzo-Cabrera, M., Sánchez-Hualde, A., & Salaverría, R. (2025b). *IBERIFIER – Assessment of VLOP and VLOSE Signatory Reports for the Strengthened Code of Practice on Disinformation. Report No. 3.* IBERIFIER. <https://doi.org/10.5281/zenodo.17966406>

Moreno Castro, C., Paisana, M., Vázquez Herrero, J., Falcó-Moreno, N., Calvo, D., Negreira Rey, M. C., Pérez-Seijo, S., Silva-Rodríguez, A., Sixto-García, J., Vizoso, Á., Moreno, J., Sánchez-Hualde, A., & Salaverría, R. (2025c). *Annual report on monitoring activities of the Code of Practice of Disinformation. No. 1.* IBERIFIER. Doi: <https://doi.org/10.5281/zenodo.16044354>

Sierra, A., Labiano, R., Novoa-Jaso, M. F., & Vara-Miguel, A. (2025). *Digital News Report España 2025: Periodismo y democracia: confianza, comunidad y narrativas innovadoras.* Servicio de Publicaciones de la Universidad de Navarra. <https://doi.org/10.15581/019.2025>

IBERIFIER – Iberian Digital Media Observatory

IBERIFIER is a digital media observatory in Spain and Portugal funded by the European Commission, linked to the European Digital Media Observatory (EDMO). It is made up of thirteen universities, five fact-checking organizations and news agencies, and five multidisciplinary research centers.

Its main mission is to analyse the Iberian digital media ecosystem and address misinformation. To do this, it focuses its research on five lines of work:

1. Research on the characteristics and trends of the Iberian digital media ecosystem.
2. Development of computational technologies for the early detection of misinformation.
3. Fact-checking of misinformation in the Iberian territory.
4. Strategic reports on threats of disinformation, both for public knowledge and for the authorities of Spain and Portugal.
5. Promotion of media literacy initiatives, aimed at journalists and informants, young people and society as a whole.

For more information, please visit the project website at iberifier.eu

Website: iberifier.eu

X: [@iberifier](https://twitter.com/iberifier)

Instagram: [@iberifier](https://www.instagram.com/iberifier)

LinkedIn: [IBERIFIER](https://www.linkedin.com/company/iberifier)

Contacts

Report coordinators:
Carolina Moreno-Castro (carolina.moreno@uv.es)
Miguel Paisana (miguel.paisana@obercom.pt)

IBERIFIER coordinator:
Ramón Salaverría (rsalaver@unav.es)



www.iberifier.eu

Coordinator



Partners



Associate



Co-funded by
the European Union

Iberifier – Iberian Digital Media Observatory has received funding from the European Commission under the Call DIGITAL-2023-DEPLOY-04, European Digital Media observatory (EDMO) — National and multinational hubs, with the reference IBERIFIER Plus - 101158511