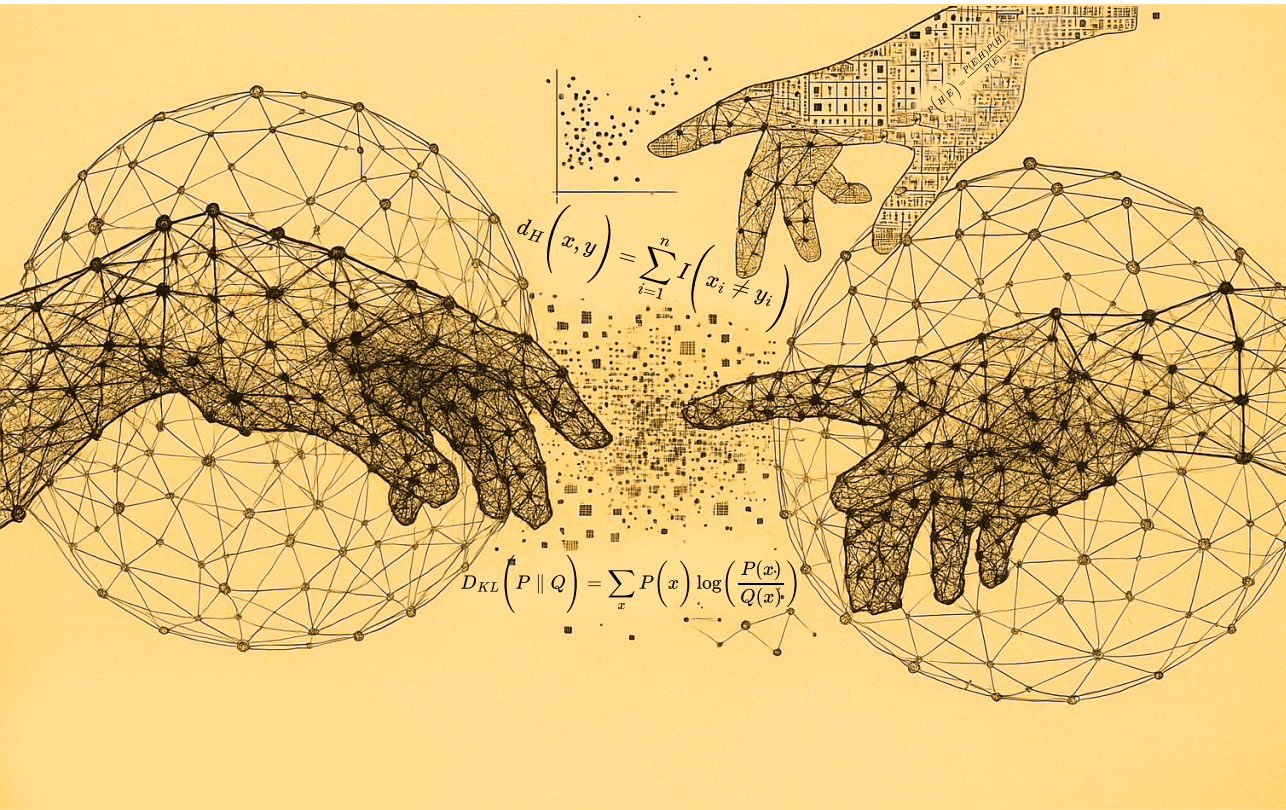
FORO CONTRA LAS CAMPAÑAS DE DESINFORMACIÓN

INICIATIVAS 2025: CONCLUSIONES Y RECOMENDACIONES DE LOS EXPERTOS

LA DESINFORMACIÓN EN LA NORMATIVA EUROPEA. EL REGLAMENTO DE SERVICIOS DIGITALES (DSA)Y SU CÓDIGO DE CONDUCTA



LA DESINFORMACIÓN EN LA NORMATIVA EUROPEA. EL REGLAMENTO DE SERVICIOS DIGITALES (DSA) Y SU CÓDIGO DE CONDUCTA

Todos los expertos participantes en los Grupos de Trabajo, tanto del sector público como del privado, lo son a título personal y no a título institucional. Por lo tanto, sus opiniones y recomendaciones no representan ni comprometen a las instituciones a las que pertenecen.

El resultado de los trabajos es producto de un ejercicio de reflexión colectivo, si bien, no tiene por qué representar la opinión individual de todos los participantes ni de las organizaciones o entidades públicas y privadas representadas, quienes no necesariamente comparten todas las conclusiones o propuestas.

ÍNDICE

LA DESINFORMACIÓN EN LA NORMATIVA EUROPEA. EL REGLAMENTO DE SERVICIOS DIGITALES (DSA) Y SU CÓDIGO DE CONDUCTA

COORDINADORES

Coordinadores:

Alejandro Perales

Expertos de la Comisión Nacional de los Mercados y la Competencia (CNMC)

Autores y Colaboradores:

Agustín Yanel

José Domingo Gómez Castallo

Juan Carlos Suárez

Raquel Vinader

Expertos del Centro de Inteligencia contra el Terrorismo y el Crimen Organizado, Ministerio del Interior


FAKE
NEWS

PRESENTACIÓN

La "desinformación", tal y como se define canónicamente por la Comisión Europea, es "información verificablemente falsa o engañosa que se crea, presenta y divulga con fines lucrativos o para engañar deliberadamente a la población y que puede causar un perjuicio público".

De acuerdo con esa definición, la desinformación como fenómeno (*disinformation*) se deslinda de la mera información errónea (*misinformation*), fruto del desconocimiento o de la falta de diligencia en la procura de la veracidad. Pero también cabe diferenciar la mera difusión, singular o reiterada, de noticias falsas, de las operaciones de influencia que buscan, si no imponer un relato "alternativo" al que se deriva de los propios hechos, al menos, de generar la suficiente confusión como para inocular la idea de que no existe (o no es cognoscible) la realidad y lo que importa es el relato. Dicho de otro modo, instalarnos en la postverdad.

Existen sectores de la opinión pública, ya sea en el ámbito político o económico, interesados en promover la desinformación, o en negar o desacreditar la realidad desde el marco de la posverdad (el relato frente al dato) para beneficiar a sus propios intereses. Las posibilidades que ofrecen internet y las tecnologías de la comunicación, a la hora de propagar mensajes, maximizar la interactividad y segmentar a los receptores, han permitido el desarrollo exponencial del fenómeno de la desinformación.

La lucha contra la desinformación en internet requiere de una respuesta normativa, global y armonizada a nivel internacional, y fruto de las iniciativas desarrolladas en estos años en la UE por la Comisión, el Consejo, el Parlamento y los organismos consultivos es el Reglamento de Servicios Digitales, pero también los referidos a la Libertad de Medios y a la Inteligencia Artificial, o el anunciado sobre las prácticas desleales en internet (fairness o equidad digital).

El presente capítulo recoge los principales aspectos del trabajo realizado por el Grupo de DSA y Desinformación, integrado en el Foro Contra las Campañas de Desinformación del Departamento de Seguridad Nacional del Ministerio de Presidencia. Entronca con el

llevado a cabo en 2023 y en esta ocasión hemos analizado las novedades aportadas por la implementación del Código de Conducta en Materia de Desinformación, continuador del Código de Buenas Prácticas en este ámbito e integrado ya en el marco del Reglamento de Servicios Digitales. Dicho análisis se complementa con la evaluación de los informes de rendición de cuentas presentados por las grandes plataformas y buscadores firmantes del Código.

ANTECEDENTES: EL CÓDIGO DE BUENAS PRÁCTICAS EN MATERIA DE DESINFORMACIÓN

Tal y como ya señalábamos en nuestro trabajo de 2023[1], las instituciones europeas han promovido en los últimos años diferentes iniciativas en la lucha contra la desinformación, tres de ellas en 2018: el Informe del Grupo de Alto Nivel *Un enfoque multidimensional de la desinformación*[2]; la Comunicación de la Comisión *La lucha contra la desinformación en línea: un enfoque europeo*[3] y el *Código de Buenas Prácticas en Materia de Desinformación*[4], suscrito por las plataformas de internet Facebook, Google, Twitter y Mozilla, así como por anunciantes y otros actores de la industria publicitaria. Microsoft se unió en mayo de 2019, y TikTok firmó el Código en junio de 2020.

Ámbitos

El Código de Buenas Prácticas planteaba actuar contra la desinformación mediante una serie de compromisos en diferentes ámbitos:

Publicidad

Desarrollar políticas y procesos para impedir los ingresos de las cuentas y sitios web con ese tipo de contenidos, con el objetivo de reducir los ingresos de los proveedores de desinformación en línea atacando a su capacidad de monetización. Ello implica no colocar publicidad en dichas páginas web ni promocionarlas.

Se indica que estas políticas y procesos deberían llevarse a cabo en colaboración con organizaciones de verificación de datos, a la hora de identificar páginas y contenidos con desinformación. Y que se requieren asimismo de herramientas que permitan a los anunciantes valorar las estrategias de compra en medios y los riesgos para su reputación, ofreciéndoles el acceso necesario a cuentas específicas de clientes para ayudarles a supervisar la colocación de anuncios publicitarios y tomar decisiones respecto a su ubicación.

---

[1] https://www.dsn.gob.es/es/publicaciones/otras-publicaciones/Foro-desinformacion-ambitoSN-trabajos2023

[2] https://op.europa.eu/es/publication-detail/-/publication/6ef4df8b-4cea-11e8-be1d-01aa75ed71a1

[3] https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52018DC0236

[4] https://digital-strategy.ec.europa.eu/es/library/2018-code-practice-disinformation

Usuarios

El empoderamiento o capacitación de los usuarios es otro de los objetivos básicos del Código, proporcionando a éstos herramientas que permitan la identificación y notificación de los casos de desinformación, y facilitando el acceso a fuentes de información diferentes y plurales y a puntos de vista alternativos en temas de interés público.

Es el caso, por ejemplo, del establecimiento de reglas claras para identificar el uso indebido de ordenadores zombis y garantizar que sus actividades no puedan confundirse con las interacciones humanas. O de políticas sobre qué constituye un abuso no permisible de sistemas automatizados, publicándolas en sus plataformas de forma que los usuarios de la UE puedan acceder a ellas. O del compromiso de no permitir el uso de cuentas y servicios de forma anónima o mediante seudónimos.

Al mismo tiempo, los firmantes se comprometían a diluir la visibilidad de la desinformación mediante la mejora de la capacidad de los usuarios para encontrar contenido fiable mediante herramientas que les permitan una experiencia en línea personalizada e interactiva.

Ello significa invertir en medios tecnológicos para:

- Ayudar a las personas a tomar decisiones con conocimiento de causa cuando encuentran noticias en línea que pueden ser falsas, promoviendo iniciativas para desarrollar y aplicar indicadores de fiabilidad eficaces en colaboración con el ecosistema informativo y las asociaciones de medios.

- Dar prioridad a información pertinente, auténtica, diversa y autorizada en canales de búsqueda automatizada de información, con indicadores sobre la fiabilidad de las fuentes, información sobre la identidad de los propietarios de los medios y transparencia respecto a los motivos por los cuales los usuarios son receptores del contenido difundido.

- Promover, junto con la sociedad civil, los gobiernos, las instituciones educativas y otras partes interesadas, iniciativas de mejora del pensamiento crítico y la alfabetización mediática e informacional (AMI) en el ámbito de los medios digitales.

Verificación de datos

Los firmantes del Código reconocían la importancia de los indicadores de fiabilidad de la información y de sus fuentes, basados en criterios objetivos e independientes, como los proporcionados por la red de verificadores de datos promovida por la Comisión Europea, a la hora de aportar datos adicionales sobre proveedores de desinformación.

En este sentido, se comprometían a colaborar con los verificadores de datos y sus organizaciones en el desarrollo de las políticas y procesos ya mencionados para interrumpir la publicidad y los incentivos a la monetización de los contenidos relacionados con la desinformación. Ello implicaba el uso de los mencionados indicadores de fiabilidad.

Comunidad investigadora

Los firmantes del Código reconocían la importancia de fomentar la investigación sobre la desinformación (difundida, por ejemplo, a través de la propaganda política), comprometiéndose a facilitar datos pertinentes sobre el funcionamiento de sus servicios para la realización de estudios independientes por parte de expertos académicos y de las organizaciones de la sociedad civil. Esto incluye información general sobre el uso de algoritmos y el intercambio de conjuntos de datos, siempre protegiendo la intimidad de las personas.

Los firmantes se comprometían, asimismo:

- A adoptar medidas de seguridad contra las declaraciones falsas antes de comercializar nuevos servicios, y a revisar en lo posible los servicios existentes para garantizar que también se aplican dichas medidas de seguridad.
- A convocar un acto anual para fomentar los debates entre el mundo académico, la comunidad de verificación de datos y los diferentes agentes de la cadena de valor.

Criterios de aplicación

El Código de Buenas Prácticas, como toda herramienta voluntaria de autorregulación, sólo obligaba a los firmantes en cuanto a su cumplimiento, los cuales además podían retirarse del mismo, o de determinados compromisos contemplados en el mismo, en cualquier momento, mediante notificación a la Comisión Europea y a los demás signatarios. Debe tenerse en cuenta que la naturaleza heterogénea de los firmantes hacía que, estos asumieran ya desde el inicio únicamente los compromisos relacionados con su oferta de bienes y servicios y con su posición en la cadena de valor, así como con su capacidad técnica, su realidad tecnológica, su tipología de usuarios y sus regímenes de responsabilidad en el marco de la UE y del Espacio Económico Europeo.

Una vez vigente el Código, cualquier modificación del mismo debía ser acordada por todos los firmantes. Cada firmante podía informar a los demás en cualquier momento si consideraba que otro estaba incumpliendo los compromisos asumidos, e indicar los motivos para tal sospecha. En tal caso, los firmantes podían acordar la celebración de una reunión plenaria para escuchar las alegaciones del signatario afectado y concluir, en su caso, invitar al infractor a retirarse del Código, informando a la Comisión Europea.

Los firmantes podían hacer pública, en sus sitios web o en comunicaciones comerciales o de otro tipo, su adhesión al Código, así como adoptar todas las medidas oportunas para informar a sus contactos comerciales.

Las asociaciones profesionales firmantes del Código no contraían obligación alguna en nombre de sus miembros, pero sí se comprometían a promover su conocimiento, y a animar a otros adherirse al mismo. El Código se refiere de modo específico a la Federación Mundial de Anunciantes, la Asociación Europea de Agencias de Comunicación (EACA) e IAB Europe, que han de ofrecer informes agregados para seguir e identificar las distintas actividades y políticas de seguridad de marca.

El Código de Buenas Prácticas, en su versión inicial, incluía una referencia expresa de reconocimiento al trabajo legislativo, entonces en curso, enfoca en la regulación de plataformas y otras empresas en el mercado digital, así como al trabajo del Grupo de Expertos en inteligencia artificial de la UE y al acervo comunitario en materia de consumo

A efectos de su seguimiento y evaluación, se establecía un período de doce meses durante los cuales debían realizarse reuniones periódicas para analizar su funcionamiento, el progreso en su aplicación y la eficacia del mismo en relación con cada uno de los compromisos establecidos anteriormente. Transcurrido ese plazo, los firmantes se comprometían a seleccionar una organización independiente para revisar de forma objetiva los informes

anuales de autoevaluación correspondientes y evaluar el nivel de progreso realizado en relación con los compromisos adquiridos (rendición de cuentas).

A partir de este balance, podrían proponerse nuevas medidas de seguimiento y posibles cambios en los compromisos a adoptar.

Las plataformas y entidades firmantes del Código presentaron en enero de 2019 un informe en el que se exponía la situación de las medidas adoptadas. Por su parte, la Comisión Europea llevó a cabo entre enero y mayo de 2019 un seguimiento específico de la aplicación de los compromisos de tres de los firmantes: Facebook, Google y Twitter, con especial hincapié en la repercusión de su actividad en las elecciones al Parlamento Europeo.

En octubre de ese mismo año se publicó el informe de autoevaluación de los firmantes, recogiendo sus iniciativas para cumplir con los compromisos establecidos en el Código, y la Comisión dio a conocer su informe de evaluación en septiembre de 20205.

El informe de la Comisión puso de manifiesto los avances que el Código había proporcionado en materia de desinformación, facilitando un diálogo con las plataformas en línea; garantizando una mayor transparencia de sus políticas, y propiciando acciones concretas y cambios políticos por parte de las partes interesadas pertinentes para ayudar a contrarrestar la desinformación. Sin embargo, también puso de manifiesto una serie de lagunas y deficiencias importantes, como su aplicación incoherente e incompleta y la falta de definiciones comúnmente compartidas; de claridad en los procedimientos; de precisión en los compromisos, y de indicadores clave de rendimiento (KPI) significativos y transparentes.

La Comisión hacía un especial hincapié, además, en la falta de acceso a datos que permitieran una evaluación independiente de las tendencias emergentes y las amenazas que plantea la desinformación en línea, acceso muy supeditado a la voluntad de los firmantes para compartir la información necesaria. Ello hacía difícil evaluar con una mínima precisión el impacto de sus acciones de los firmantes, por lo que se demandaba un modelo más estructurado de cooperación con la comunidad investigadora.

Planteaba establecer un procedimiento adecuado de seguimiento y evaluación, que permitiera una mejor garantía de rendición de cuentas, así como una extensión del Código a otras partes interesadas pertinentes, en particular del sector de la publicidad.

Y finalmente, anticipaba que el Código deberá reformularse y enmarcarse en el futuro Reglamento de Servicios Digitales, una vez fuera éste aprobado.

---

[5] https://digital-strategy.ec.europa.eu/es/node/700

El Código de Buenas Prácticas en Materia de Desinformación Reforzado

Las objeciones de la Comisión Europea ante el Código de Buenas Prácticas quedaron plasmadas en la *Guía de Orientaciones para reforzar el Código de buenas prácticas en materia de desinformación* de mayo de 2021. Pero es necesario mencionar, como paso previo, la Comunicación de la Comisión *Plan de Acción para la Democracia Europea*, aprobado en diciembre de 2020.

Básicamente, en dicho Plan se destaca que el rápido crecimiento de las plataformas en línea ha evidenciado nuevas vulnerabilidades y ha hecho más difícil mantener la integridad de las elecciones; garantizar unos medios libres y plurales, esenciales para el proceso democrático, y proteger a la ciudadanía frente a la desinformación y otros tipos de manipulación.

En línea con otros documentos ya mencionados, esta Comunicación define la desinformación como "un contenido falso o engañoso que se difunde con intención de engañar o de obtener una ganancia económica y política y que puede causar un perjuicio público" (*disinformation*), diferenciándola de la información engañosa o falsa compartida sin intención de perjudicar, aunque sus efectos puedan ser nocivos (*misinformation*). Y distinguiendo también entre las "operaciones de influencia" por un lado, fruto del esfuerzo coordinado de actores nacionales, y la "injerencia extranjera" por otro, que a menudo forma parte de una operación híbrida más amplia.

De cara al futuro, la Comunicación apostaba por revisar el Código de Buenas Prácticas de acuerdo con lo señalado por la propia CE, pero también por su integración en la (entonces) próxima regulación de los servicios digitales, como marco horizontal que vendría a garantizar una mayor responsabilidad de las plataformas a la hora de rendir cuentas sobre la forma en la que moderan sus contenidos y aplican sus algoritmos.

Volviendo a la Guía, ésta orientaba sobre cómo podría reforzarse el Código de Buenas Prácticas en diferentes ámbitos, con el fin de convertirlo en una herramienta más eficaz para contrarrestar la desinformación garantizando una aplicación completa y coherente entre las partes interesadas y los países de la UE:

- Mayor participación en los compromisos adoptados.
- Mejor desmonetización de la desinformación.
- Garantizar la integridad de los servicios.
- Mejorar la capacitación de los usuarios.

- Aumentar la cobertura de la verificación de datos.
- Proporcionar a los investigadores un mayor acceso a los datos.
- Crear un marco de seguimiento más sólido.

También reclamaba un centro en el que pudiera recogerse información sobre las políticas adoptadas para aplicar los compromisos del Código y su grado de cumplimiento.

En ese contexto, los firmantes del Código y los posibles nuevos firmantes se reunieron el 8 de julio de 2021, poniendo en marcha una convocatoria de manifestaciones de interés dirigida a una amplia gama de partes interesadas, en la que se invitaba a éstas a convertirse en firmantes y a participar en la actualización y reforzamiento del Código. La convocatoria incluía a los proveedores de servicios en línea que difunden contenidos al público (plataformas y redes sociales), a los servicios de búsqueda, a las aplicaciones de mensajería privada, al sector de la publicidad y a otros agentes interesados en contribuir a la lucha contra la difusión de desinformación a través del desarrollo de herramientas y del "trabajo filantrópico", o por contar con conocimientos especializados.

La Asamblea de firmantes procedió a revisar Código de 2018 ofreciendo su versión reforzada en junio de 2022[6].

Esta nueva versión reunía un elenco de firmantes más extenso que el de 2018, entre las grandes plataformas, buscadores y redes sociales Google, Meta, Microsoft, TikTok, Twich y Twitter. Contenía 44 compromisos específicos que podían asumirse total o parcialmente, en general proponiendo, con poca concreción, medidas algo más estrictas en los ámbitos básicos establecidos en 2018:

- Desmonetización de los contenidos desinformativos.
- Identificabilidad (etiquetado) y transparencia de la propaganda política; que requiere de repositorios (ad libraries) eficientes y con capacidad de búsqueda.
- Medidas contra las prácticas de manipulación como cuentas falsas y bots para amplificar su diseminación de los mensajes; suplantación de identidad, *malware*, etc.

[6] https://digital-strategy.ec.europa.eu/es/policies/code-practice-disinformation

- Empoderamiento de los verificadores, garantizando para ellos recursos financieros justos, proporcionándoles un mayor acceso a la información y extendiendo su trabajo, con el compromiso de utilizar sus datos de forma coherente y eficiente.
- Empoderamiento de los investigadores, brindándoles un mayor acceso (automatizado y de escrutinio) a los contenidos difundidos y almacenados por las plataformas en línea y las redes sociales.
- Empoderamiento de los usuarios, proporcionándoles mayor protección desde las propias plataformas y los dispositivos ("desde el diseño y por defecto"); mayor transparencia de los sistemas de recomendación, primando las fuentes fiables, y posibilitando participación activa mediante la alfabetización y las herramientas de identificación, contraste, señalamiento de la desinformación para procurar su eliminación.

El Código reforzado preveía la creación de un Centro de Transparencia, accesible a todos los ciudadanos, que permita conocer las acciones desarrolladas por los firmantes para la implementación de las medidas previstas, con actualizaciones periódicas de los datos relevantes.

También contemplaba la labor de un Grupo de Trabajo permanente, dedicado al seguimiento del cumplimiento del Código y a su evolución en el marco de los objetivos perseguidos, con el establecimiento de un foro que revise precisamente la vigencia de los compromisos contraídos y su posible actualización a la vista de los desarrollos tecnológicos, sociales, de mercado y legislativos.

Este Grupo de Trabajo estaba compuesto por representantes de los firmantes, del Grupo de Reguladores Europeos de Servicios de Medios Audiovisuales (ERGA), del Observatorio Europeo de Medios Digitales (EDMO) y el Servicio Europeo de Acción Exterior, presidido por la Comisión.

X abandona el Código en 2022, y antes de finalizar ese año se aprueba el de Servicios Digitales.

EL REGLAMENTO DE SERVICIOS DIGITALES

El Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales, conocido como Reglamento de Servicios Digitales o, más comúnmente, como DSA (Digital Services Act)[7] es el marco de referencia regulatorio del entorno digital (junto el Reglamento 2022/1925/CE, de 8 de junio, de Mercados Digitales).

La DSA, que entró en vigor el 17 de febrero de 2024[8], es una legislación pionera en la UE que regula determinados servicios online, incluyendo los mercados en línea, las redes sociales o las tiendas de aplicaciones.

Tiene como objetivo regular el entorno digital para hacerlo más seguro, predecible y fiable, y para que los ciudadanos de la Unión puedan ejercer los derechos garantizados por la Carta de los Derechos Fundamentales de la UE.

El Reglamento de Servicios Digitales establece un nuevo marco normativo para abordar retos como la circulación de contenido ilícito en la red, la falta de transparencia en la moderación de contenidos, la protección de los menores online frente a contenido ilícito o dañino o la salvaguarda de los derechos de los consumidores.

Para los ciudadanos supone una mayor protección de los derechos fundamentales; mayor facilidad para denunciar contenidos ilícitos; mayor protección de los menores (por ejemplo, la prohibición de publicidad a ellos dirigida); menor exposición a contenidos ilícitos; una mayor transparencia sobre las decisiones de moderación[9], y, en general, mayor control democrático y supervisión del entorno digital.

---

[7] BOE.es - DOUE-L-2022-81573 Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE (Reglamento de Servicios Digitales).

[8] Para las VLOPS y VLPSE el Reglamento previó aplicación de determinadas previsiones con carácter previo al 17 de febrero de 2024.

[9] https://transparency.dsa.ec.europa.eu/?lang=es
https://transparency.dsa.ec.europa.eu/page/data-retention-policy?lang=es

La normativa se aplica a todos los servicios de intermediación de la sociedad de la información[10] —por ejemplo, las plataformas en línea o los motores de búsqueda, los *marketplaces*, las redes sociales, o las tiendas de aplicaciones— ofrecidos a destinatarios del servicio que estén situados en la Unión Europea, con independencia de donde los prestadores de dichos servicios intermediarios tengan su lugar de establecimiento.

Es importante destacar que el nuevo Reglamento no establece qué está o no permitido ofrecer o publicar en las plataformas. Las prohibiciones se rigen por las normas que regulan sus respectivos ámbitos legales (normativa de protección de datos, la de consumo, normativa penal, entre otras).

Los prestadores de servicios en línea sí que pueden ser considerados responsables por la forma en que gestionan el contenido en sus plataformas; por ejemplo, por no atender órdenes de retirada, no poner a disposición de los usuarios sistemas de denuncia de contenido o no establecer medidas para proteger a los usuarios de riesgos específicos.

Por ello, deben disponer de mecanismos de reclamación y de resolución de conflictos y de procedimientos para notificar y retirar contenidos ilícitos, y en general, de protección de los usuarios frente a los riesgos que afrontan en el mundo digital. En definitiva, se establecen obligaciones de diligencia debida, transparencia e información por parte de los proveedores.

Las plataformas deben ser transparentes sobre cómo funcionan sus algoritmos, especialmente aquellos que recomiendan contenido o personalizan servicios, y deben ofrecer a los usuarios la opción de un sistema de recomendación que no se base en perfiles personales. En el ámbito de la publicidad, se prohíben los anuncios dirigidos a menores o basados en datos sensibles, y se exige que los usuarios sean informados de manera clara sobre por qué se les muestra un anuncio y quién lo financia.

Además, el Reglamento fortalece la protección de los usuarios al exigir a las plataformas que establezcan mecanismos claros y accesibles para que éstos puedan denunciar contenidos ilegales y apelar las decisiones de moderación.

En el caso de las **plataformas y los buscadores que tienen más de 45 millones de usuarios**[11] al mes en la UE, se establece una regulación más estricta por el impacto que tienen en los usuarios y el ecosistema digital.

---

[10] i) Servicio de mera transmisión, ii) servicio de memoria caché, iii) servicio de alojamiento de datos

[11] https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses

Estas plataformas se denominan plataformas en línea de muy gran tamaño y motores de búsqueda de muy gran tamaño *Very Large online Platforms -VLOPs- y Very Large Online Search Engines -VLOPSE-*, por sus siglas en inglés).

Las VLOP y VLOSE tienen obligaciones adicionales al resto de prestadores, como la de realizar evaluaciones anuales de riesgos sistémicos, (entre los que se incluye, la difusión de contenido ilícito o cualquier efecto negativo real o previsible sobre el discurso cívico y los procesos electorales) , y de implementar medidas para reducir dichos riesgos.

La supervisión del Reglamento se comparte entre la Comisión Europea, con competencias sobre las VLOP y VLOPSE, y los denominados Coordinadores de Servicios Digitales de cada Estado miembro, a los que luego nos referiremos, con competencias para supervisar y hacer cumplir el Reglamento por parte de los prestadores de servicios intermediarios establecidos en dicho Estado miembro-.

El Reglamento atribuye a la Comisión Europea y a los coordinadores de servicios digitales amplias competencias para poder supervisar el cumplimiento del mismo, incluida la potestad sancionadora con multas de hasta el 6% del volumen de negocios anual.

El Reglamento de Servicios Digitales (DSA) contribuye a la consolidación de la equidad digital en la UE, junto con normas como la Directiva sobre los Requisitos de Accesibilidad de los Productos y Servicios[12] (Ley Europea de Accesibilidad, EAA), el Reglamento de Mercados Digitales (DMA) y la futura normativa sobre prácticas desleales en el entorno online (digital fairness act). Si la EAA y la DMA se centran, respectivamente, en la inclusión y la competencia, la DSA tiene como eje principal la protección de los usuarios y la transparencia. Todas ellas forman un paquete regulatorio que aborda los desafíos del entorno digital desde múltiples ángulos y se cimentan, su vez en los valores democráticos establecidos en la Declaración Europea sobre los Derechos y Principios Digitales, que sitúa a las personas en el centro de la transformación digital. Y hay que mencionar también el papel que el Reglamento Europeo de Libertad de Medios (EMFA) y la Directiva de Servicios de Medios Audiovisuales (AVMSD) pueden desempeñar en la lucha contra la desinformación.

Como se ha señalado anteriormente, una característica distintiva de la DSA es su enfoque de la regulación asimétrica, que impone obligaciones diferenciadas según el tamaño y la influencia del servicio de intermediación. Todos los prestadores de servicios intermediarios tienen obligaciones básicas, pero las VLOP y los VLOSE se enfrentan a responsabilidades más estrictas.

---

[12] https://www.boe.es/doue/2019/151/L00070-00115.pdf

Esta regulación asimétrica de la DSA, si bien es una solución sensata, requiere una vigilancia continua para garantizar que las obligaciones se apliquen de manera efectiva pero sin generar una carga indebida al as empresas más pequeñas.

Cabe destacar que, en el marco de la DSA, en abril de 2023 la Comisión Europea puso en marcha el Centro Europeo para la trasparencia algorítmica (ECAT). Este Centro aporta experiencia científica y técnica al papel exclusivo de supervisión y aplicación de la Comisión Europea en relación con las VLOP y VLOPSE.

Son sus actividades:

- Llevar a cabo ensayos técnicos de sistemas algorítmicos para comprender su funcionamiento.
- Analizar los informes de transparencia, las evaluaciones de riesgos y las auditorías independientes.
- Apoyar las investigaciones e inspecciones.
- Identificar los riesgos emergentes asociados con el uso de VLOP/VLOSE;
- Actuar como un centro de conocimiento para la investigación realizada gracias al acceso a los datos contemplado en el Reglamento.

Uno de los aspectos clave de la DSA son los agentes y modelos contemplados para su aplicación y gestión: coordinadores de servicios digitales, alertadores fiables, resolución extrajudicial de conflictos y códigos de conducta.

Los coordinadores de servicios digitales

Como se ha señalado anteriormente, la DSA establece un sistema de gobernanza compartida entre la Comisión Europea y los coordinadores de servicios digitales.

El Reglamento de Servicios Digitales establece en su artículo 49 la obligación de los Estados miembros de designar una o varias autoridades competentes responsables de la supervisión de los prestadores de servicios intermediarios y de la ejecución del Reglamento, y, adicionalmente, de designar a una de estas autoridades competentes como su coordinador de servicios digitales.

El coordinador de servicios digitales, además de ser responsable de la supervisión y garantía del cumplimiento del Reglamento en el Estado miembro en cuestión, debe garantizar la coordinación en el ámbito nacional respecto a las materias objeto del Reglamento.

El 24 de enero de 2024, en cumplimiento de su obligación establecida en el artículo 49.3 del Reglamento de Servicios Digitales, el Ministerio para la Transformación Digital y de la Función Pública designó a la Comisión Nacional de los Mercados y la Competencia (CNMC) como coordinador de servicios digitales nacional. Adicionalmente, se designó a la Agencia Española de Protección de Datos autoridad competente en materia de supervisión del cumplimiento de la normativa de protección de datos, con plena cooperación entre ambos organismos.

Es necesario mencionar que, a pesar de que se produjo la designación del coordinador de servicios digitales, la habilitación legal para atribuir las plenas competencias a la CNMC y a la AEPD que establece el Reglamento no se ha producido aún. Esto impide que el coordinador de servicios digitales pueda ejercer todas las competencias encomendadas por el Reglamento[13] (entre ellas, la potestad para certificar alertadores fiable o el ejercicio de la potestad de inspección y sanción, entre otras).

[13] El 16 de diciembre de 2024 la Comisión Europea remitió a España un Dictamen motivado por la falta de habilitación legal de la CNMC. El 7 de mayo de 2025 la Comisión Europea inició el procedimiento de infracción contra España ante el Tribunal de Justicia de la UE por la falta de implementación efectiva del Reglamento de Servicios Digitales: Commission decides to refer Czechia, Spain, Cyprus, Poland and Portugal to the Court of Justice of the European Union due to lack of effective implementation of the Digital Services Act | Shaping Europe's digital future

Actualmente, se encuentra en tramitación en el Congreso de los Diputados[14] el proyecto de Ley por el que se modifican diversas disposiciones legales para la mejora de la gobernanza democrática en servicios digitales y ordenación de los medios de comunicación en el que se ha incluido la habilitación a la CNMC como coordinador de servicios digitales.

Funciones del coordinador de servicios digitales

El Reglamento de Servicios Digitales no contiene una definición legal de desinformación, aunque identifica la lucha contra este fenómeno como uno de los objetivos del Reglamento[15], y la desinformación se aborda en varios considerandos[16].

Las obligaciones relativas a desinformación que contiene el Reglamento, y que van a ser determinantes del papel del coordinador de servicios digitales, se pueden clasificar en dos grupos:

(i) Obligaciones específicas dirigidas a las VLOPS y VLOPSE respecto a los riesgos sistémicos derivados del diseño y funcionamiento de su servicio y de los sistemas relacionados con este, incluidos los sistemas algorítmicos, o del uso que se haga de sus servicios (artículos 34 y 35).

A este respecto, las funciones de la CNMC como coordinador de servicios digitales son limitadas, ya que no hay ninguna VLOPSE establecida en España a esta fecha, y, además, la supervisión de estas obligaciones del Reglamento es competencia exclusiva de la Comisión Europea.

En todo caso, si el coordinador de servicios digitales tuviera conocimiento de actuaciones de las VLOPS y VLOPSE que pudieran implicar infracción de estas obligaciones -especialmente fruto de las reclamaciones recibidas en virtud del artículo 53, o de las órdenes de actuación y órdenes de información emitidas por autoridades competentes, y recibidas en virtud de los artículos 9 y 10- dará traslado al coordinador de servicios digitales del estado miembro donde el prestador esté establecido, y a la Comisión Europea, como consecuencia de la obligación de asistencia mutua.

---

[14] https://www.congreso.es/es/proyectos-de-ley?p_p_id=iniciativas&p_p_lifecycle=0&p_p_state=-normal&p_p_mode=view&_iniciativas_mode=mostrarDetalle&_iniciativas_legislatura=XV&_iniciativas_id=121/000068

[15] Considerando 9.

[16] Considerandos 2, 9, 69, 83, 84, 88, 95, 104, 106 y 108.

(ii) Obligaciones contra prácticas inapropiadas que pueden ser relevantes en materia de desinformación, que son obligatorias para los VLOPS y VLOPSE, pero también, ciertas de ellas afectan a otro tipo de intermediarios.

a. transparencia de la moderación de contenido:

i. transparencia y aplicación diligente de las reglas de moderación de contenido (artículo 14) y declaración motivada de las restricciones impuestas (artículo 17);

ii. informes anuales (artículos 15 y 24).

b. protección de los usuarios contra prácticas engañosas y perfilado engañoso de sus datos para uso publicitario o de los sistemas de recomendación:

i. prohibición de los patrones oscuros "dark patterns" (artículo 25);

ii. transparencia publicitaria y prohibición de elaboración de perfiles basados en datos protegidos para enviar publicidad dirigida (artículo 26). Se señala expresamente que el uso de técnicas de manipulación puede contribuir a campañas de desinformación o discriminando a determinados grupos (considerando 69);

iii. transparencia de los sistemas de recomendación (artículo 27);

iv. medidas para reforzar la protección de menores y prohibición de la elaboración de perfil si el destinatario es un menor (artículo 28).

c. obligaciones adicionales relativas a la publicidad y a los sistemas de recomendación:

i. el Reglamento exige que los VLPS y VLOPSE mantengan un repositorio público de publicidad (artículo 39). Estos repositorios ayudarán a los investigadores a supervisar y estudiar los riesgos emergentes, como las campañas de desinformación que afectan negativamente a la salud pública, la seguridad, el discurso civil, la participación política o la igualdad (considerando 95);

ii. ofrecer al menos una opción para que cada uno de los sistemas de recomendación de las VLOPS y VLOPSE no se base en la elaboración de perfiles (artículo 38).

En relación con este grupo de obligaciones, la CNMC, como coordinador de servicios digitales de España, tiene competencias de supervisión respecto a los prestadores de servicios intermediarios establecidos en España y respecto a aquellos prestadores sin establecimiento en la UE que nombren representante legal en España.

El artículo 51 del Reglamento establece las facultades específicas de los coordinadores de servicios digitales para llevar a cabo esta supervisión. En concreto, los coordinadores dispondrán de facultades de investigación (incluida la facultad de llevar a cabo inspecciones), de ejecución (incluidas las facultades de aceptar compromisos, de imponer multas, o de adoptar medidas cautelares).

Dentro de las funciones de los coordinadores destaca también la de certificación en cada Estado miembro de los alertadores fiables y de los organismos de resolución extrajudicial de conflictos, que hemos mencionado y a los que nos referiremos a continuación.

Como el resto de coordinadores de servicios digitales, la CNMC forma parte de la Junta Europea de Servicios Digitales (en adelante, la Junta) creado por la DSA como un grupo consultivo independiente integrado por los citados coordinadores para la supervisión de los prestadores de servicios intermediarios.

La Junta tiene como principal función asesorar a los coordinadores de servicios digitales y a la Comisión para el cumplimiento de los siguientes objetivos:

- Contribuir a la aplicación coherente del Reglamento y a la cooperación efectiva de los coordinadores de servicios digitales y la Comisión.
- Asistir a la Comisión Europea y a los coordinadores de servicios digitales en sus tareas de supervisión de las VLOP y VLOSE.
- Coordinar y contribuir a las directrices y los análisis de la Comisión, los coordinadores de servicios digitales y otras autoridades competentes sobre problemas emergentes en el mercado interior relacionados con el Reglamento de Servicios Digitales.

En el seno de la Junta se han constituido ocho grupos de trabajo[17] que asisten y reportan a la misma. Cada uno de ellos está dedicado a un aspecto específico de la DSA.

[17] https://digital-strategy.ec.europa.eu/en/policies/dsa-board-working-groups

El grupo de trabajo 4, denominado "Integridad del espacio europeo" se encarga de aspectos relacionados con la desinformación, además de procesos electorales[18], manipulación e interferencia extranjera en la información y otras cuestiones relacionadas con el discurso cívico.

Los miembros de los grupos de trabajo son los miembros de la Junta, a saber, los coordinadores de servicios digitales que desean participar en el grupo de trabajo pertinente. Cada grupo de trabajo está presidido por un representante técnico de la Comisión Europea, junto con un coordinador de servicios digitales para que actúe como vicepresidente.

Los alertadores fiables

Los alertadores fiables previstos en el Reglamento de Servicios Digitales son entidades responsables de detectar contenidos potencialmente ilícitos en línea, tales como productos falsificados, discursos de odio, desinformación, publicidad ilícita la incitación al odio o los contenidos terroristas, y en notificarlos a las plataformas.

Designados por los coordinadores nacionales de servicios digitales del Estado miembro donde el solicitante se encuentre establecido, sus notificaciones sobre contenidos ilícitos deben tratarse con prioridad por las plataformas, ya que por su experiencia, especialización y dedicación se espera que sean más precisos que un usuario medio. Los proveedores tienen la responsabilidad exclusiva de decidir sobre los avisos y, cuando esté justificado, eliminar el contenido.

La DSA establece los siguientes requisitos que han de reunir los solicitantes para poder ser certificados como alertadores fiables: (i) poseer conocimiento y experiencia en la detección y notificación de contenidos ilícitos; (ii) ser independientes de las plataformas; (iii) realizar sus actividades de manera diligente, precisa y objetiva.

Por su parte, los alertadores fiables deben publicar informes anuales detallados sobre las notificaciones enviadas, los tipos de contenidos ilícitos denunciados, y las medidas adoptadas por las plataformas.

Los beneficios que aportan los alertadores fiables son los siguientes: en primer lugar, una mayor eficiencia en la lucha contra contenidos ilícitos, en la medida en que agilizan el proceso de denuncia y retirada de productos y contenidos ilegales, lo que contribuye

[18] A destacar DSA Election Toolkit como guía para los coordinadores de servicios digitales en los procesos electorales: https://digital-strategy.ec.europa.eu/en/library/dsa-elections-toolkit-digital-services-coordinators

a un entorno digital más seguro. En segundo lugar, una mayor transparencia, al dar cumplimiento a su obligación de publicar informes anuales sobre las notificaciones realizadas y resultado de las mismas. Y, en tercer lugar, un refuerzo de la protección de los derechos de los usuarios, al facilitar y contribuir a la eliminación de contenidos ilícitos y dañinos.

Los órganos de resolución extrajudicial de litigios

Los órganos de resolución extrajudicial de litigios previstos también en el Reglamento son órganos independientes que ayudan a resolver conflictos entre usuarios y plataformas en línea, sin necesidad de acudir a los tribunales.

Para su certificación, por parte de los coordinadores de los servicios digitales del Estado miembro donde se encuentre establecido el solicitante, y por un período de hasta cinco años renovable, es necesario que acrediten lo siguiente: primero, que son imparciales e independientes; segundo, que cuentan con conocimientos especializados; que garantizan accesibilidad electrónica; tercero, que son capaces de resolver disputas de forma rápida, eficiente, eficaz y económica; y, cuarto, que aplican normas de procedimientos claras y justas, fácilmente accesibles al público y conformes al Derecho. Además, en quinto lugar, sus miembros no pueden ser remunerados en función del resultado del procedimiento, y deben operar al menos en una lengua oficial de la UE.

Los destinatarios del servicio pueden impugnar las decisiones de las plataformas en línea si no están de acuerdo con la misma, por ejemplo, cuando deseen impugnar una decisión de moderación de contenidos adoptada por dicha plataforma, como pueda ser la eliminación de una publicación; la suspensión o supresión de una cuenta; o la supresión, cesación o restricción de la capacidad de monetizar la información proporcionada por los destinatarios. También cuando la plataforma haya rechazado una denuncia de contenido ilícito o contenido incompatible con sus condiciones generales. Según la DSA, Las plataformas en línea están obligadas a disponer de un sistema interno de gestión de reclamaciones que permita la impugnación de las decisiones adoptadas.

La DSA adicionalmente, reconoce el derecho de los destinatarios del servicio a elegir cualquier órgano de resolución extrajudicial de litigios certificado, para resolver litigios relativos a esas decisiones.

Los órganos extrajudiciales de resolución de litigios adoptan una decisión rápida tras analizar el caso, teniendo en cuenta las normas de la plataforma y las leyes aplicables; si bien no tienen competencia para imponer a las partes una resolución vinculante del litigio. La imposición de las tasas del proceso dependerá de su resultado; si bien, para los destinatarios del servicio, la resolución de litigios debe ser gratuita o simbólica. En todo caso, los destinatarios del servicio pueden en cualquier momento recurrir a los tribunales nacionales.

El recurso a estos órganos extrajudiciales aporta significativos beneficios: en primer lugar, refuerza la protección del consumidor al facilitar a los usuarios una vía más accesible para resolver sus conflictos con las plataformas en línea. En segundo lugar, la resolución extrajudicial es más rápida y económica que la vía judicial, lo que beneficia tanto a los usuarios como a las plataformas; a la vez que descongestiona los Tribunales. En tercer lugar, estos órganos contribuyen a la transparencia y a aumentar la confianza de los usuarios en las plataformas en línea.

España se cuenta ya con organismo que podrían desempeñar ese rol, previa certificación por el coordinador de servicios digitales, como son los organismos de resolución alternativa de litigios en materia de consumo acreditados conforme a la Ley 7/2017, de 2 de noviembre, por la que se incorpora al ordenamiento jurídico español la Directiva 2013/11/UE, del Parlamento Europeo y del Consejo, de 21 de mayo de 2013, relativa a la resolución alternativa de litigios en materia de consumo (Ley de ADRs). Entre ellos, AUTOCONTROL, que además de estar acreditado como tal ADR, cumple con lo previsto en la Ley 3/1991, de 10 de enero, de Competencia Desleal para tales órganos encargados de la resolución extrajudicial de reclamaciones en su condición de órgano de control de códigos de conducta.

Los códigos de conducta

El Reglamento de Servicios Digitales otorga un papel relevante a los Códigos de Conducta voluntarios suscritos por los agentes vinculados, destacando que pueden contribuir a:

- Ayudar a las partes interesadas a adoptar compromisos que contribuyan a su aplicación. También pueden incluir compromisos adicionales de presentación de informes complementarios a los obligatorios anuales previstos en el Reglamento, para mejorar la transparencia sobre las medidas adoptadas por los servicios intermediarios.

- Proporcionar un foro de intercambio entre plataformas de diferentes tamaños, organizaciones de la sociedad civil, investigadores y otras partes interesadas, lo que ayuda a los firmantes a rendir cuentas de sus compromisos. Asimismo, los códigos permiten a las plataformas más pequeñas y a las organizaciones de la sociedad civil, entre otros, contribuir activamente al intercambio de buenas prácticas.

De acuerdo con lo señalado por el Reglamento, los firmantes pueden ser VLOP y VLOSE, pero también otras plataformas en línea (con medidas proporcionales a su tamaño y recursos), y otros servicios intermediarios. Se contempla, además, la participación de las autoridades competentes, de las organizaciones de la sociedad civil y de otras partes interesadas.

El artículo 45 del Reglamento contempla en su apartado 1 el papel a desempeñar por las autoridades competentes (Comisión y por la Junta Europea de Servicios Digitales) a la hora de fomentar y facilitar la elaboración de códigos de conducta voluntarios en el ámbito de la Unión, con el fin de contribuir a la debida aplicación de esta norma.

En su apartado 2 se indica que, cuando se genere un riesgo sistémico significativo, tal y como se define en el artículo el artículo 34.1 ("*cualquier efecto negativo real o previsible sobre el discurso cívico y los procesos electorales, así como sobre la seguridad pública*") , y afecte a varias plataformas o motores de búsqueda de muy gran tamaño, la Comisión podrá invitar a estos prestadores afectados, así como a otros prestadores de muy gran tamaño, plataformas en línea y de otros servicios intermediarios (según sea oportuno) a participar en la elaboración de códigos de conducta. En esa elaboración habrán de participar también las autoridades competentes, las organizaciones de la sociedad civil y otras partes interesadas.

Ello implica, de acuerdo con este apartado, establecer compromisos de adopción de medidas específicas de reducción de riesgos, así como un marco de información periódica sobre las medidas que se puedan adoptar y sus resultados.

La Comisión y la Junta (apartado 3) tratarán de asegurarse de que los códigos de conducta expongan claramente sus objetivos específicos, contengan indicadores clave de eficacia para valorar el cumplimiento de dichos objetivos y tengan debidamente en cuenta las necesidades e intereses de todas las partes, y, en particular, de los ciudadanos, en el ámbito de la Unión. También ratarán de asegurarse de que los participantes informen periódicamente a la Comisión y a sus respectivos coordinadores de servicios digitales de establecimiento acerca de las medidas que puedan adoptarse y sus resultados, valoradas con arreglo a los indicadores clave de eficacia que contengan. Los indicadores clave de eficacia y los compromisos de información tendrán en cuenta las diferencias de tamaño y capacidad de los diferentes participantes.

La Comisión y Junta (apartado 4) evaluarán si los códigos de conducta cumplen los fines especificados, y vigilarán y evaluarán periódicamente el cumplimiento de sus objetivos, teniendo en cuenta los indicadores clave de eficacia que puedan contener, publicando sus conclusiones y fomentando y facilitando su revisión y adaptación periódicas.

Señalan, en este sentido, algunos aspectos a los que deben atender los códigos de conducta:

- Definir claramente la naturaleza de los objetivos de interés público que se persiguen.

- Contener mecanismos para la evaluación independiente de la consecución de estos objetivos, que deben ser medibles y estar sujetos a supervisión pública.

- Presentar una definición clara de la función de las autoridades competentes.

- Prestar especial atención a la evitación de efectos negativos en la seguridad, la protección de la privacidad y los datos personales.

- Explorar medidas de reducción de riesgos relativas a tipos concretos de contenidos ilícitos, a través de acuerdos de autorregulación y corregulación.

- Valorar las posibles repercusiones negativas de los riesgos sistémicos para la sociedad y la democracia, como la desinformación, las actividades manipulativas y abusivas, o las específicamente perjudiciales y nocivas para los menores. En el caso concreto de la desinformación, el Reglamento se refiere a las operaciones coordinadas dirigidas a su amplificación, con el uso de *bots* y cuentas falsas para generar información deliberadamente incorrecta o engañosa, a veces con el fin de obtener un beneficio económico.

- Facilitar la accesibilidad de las personas con discapacidad. En particular, garantizando que la información se presente de forma perceptible, funcional, comprensible y sólida y que los formularios estén disponibles de manera fácil de localizar por estas personas.

La desinformación en el Reglamento de Servicios Digitales

La DSA no menciona la "desinformación" de modo específico en su parte dispositiva, pero sí lo hace en sus considerandos:

Considerando 2. Necesidad de armonizar las obligaciones que deben exigirse a los prestadores de servicios intermediarios de la sociedad de la información (los que permiten alojar, distribuir o localizar contenidos) por lo que se refiere al modo en que debe hacerse frente, entre otros fenómenos no deseados, a la desinformación.

Considerando 9. Esa armonización de las normas aplicables a los servicios intermediarios en el mercado interior (entre otros contenidos ilícitos, a la desinformación) tiene como objetivo garantizar un entorno en línea seguro, predecible y digno de confianza, buscando proteger eficazmente los derechos fundamentales reconocidos en la Carta de la Unión Europea y facilitando la innovación.

Considerando 69. Las técnicas de manipulación, que optimizan la segmentación de los destinatarios de los contenidos apelando a sus vulnerabilidades, pueden tener efectos negativos especialmente graves y contribuir al desarrollo de campañas de desinformación.

Considerandos 80-83. La DSA incluye entre los "riesgos sistémicos" para la sociedad y la democracia que pueden derivarse de la actividad de los prestadores de plataformas y motores de búsqueda de muy gran tamaño las campañas coordinadas de desinformación.

Considerando 84. La amplificación algorítmica de la información puede contribuir a la expansión rápida de esos riesgos sistémicos, y los prestadores deben reflejarlo debidamente en las evaluaciones de su actividad.

Considerando 88. Los prestadores deben sopesar la adopción de medidas de concienciación, en especial cuando los riesgos estén relacionados con campañas de desinformación.

Considerando 95. Deben facilitar también la supervisión y la investigación de los riesgos emergentes generados por la desinformación sobre la salud pública, la seguridad pública, el discurso civil, la participación política y la igualdad.

Considerandos 103 a 106. Finalmente, el Reglamento señala que debe promoverse la elaboración de códigos de conducta específicos, carácter voluntario, para mitigar los riesgos sistémicos, con sus correspondientes medidas de penalización y sanción en caso de incumplimiento, poniendo como ejemplo el Código de Buenas en Materia de Desinformación.

En octubre de 2024, los firmantes solicitaron a la Comisión y a la Junta Europea de Servicios Digitales la evaluación del Código de Buenas Prácticas, de conformidad con el mencionado artículo 45.4 del Reglamento, con el fin de determinar si cumplía los requisitos para ser considerado "código de conducta voluntario" en el marco del mismo.

En febrero de 2025, la Comisión y la Junta hicieron públicas sus conclusiones[19], concluyendo, tras la evaluación realizada, que el texto de Código cumplía los requisitos establecidos en los apartados 1 y 3 del artículo 45. Se aprobaba, por tanto, su integración en el marco del Reglamento como Código de Conducta en materia de Desinformación, solicitándose a los firmantes que continuaran con su desarrollo y que aplicaran los indicadores estructurales previstos. El nuevo Código entró en vigor a partir del 1 de julio de 2025.

---

[19] https://ec.europa.eu/newsroom/dae/redirection/document/112679
https://ec.europa.eu/newsroom/dae/redirection/document/112680

EL CÓDIGO DE CONDUCTA EN MATERIA DE DESINFORMACIÓN

El Código de Conducta en materia de Desinformación[20] (*Code of Conduct on Disinformation*) actualiza los compromisos asumidos por plataformas en línea, actores de la industria publicitaria, verificadores de datos, organizaciones de investigación y de la sociedad civil.

Los firmantes se comprometen a tomar medidas específicas de mitigación de riesgos; en concreto, el Código de Conducta incluye 43 compromisos y diversas medidas destinadas a implementar dichos compromisos, centradas en las áreas a las que ya se refería el Código de Buenas Prácticas:

- Desmonetización: Impedir la financiación de quienes difunden en las plataformas y redes noticias falsas o desarrollan estrategias de desinformación (*follow the money*).

- Trasparencia de la publicidad política (propaganda), mediante un etiquetado eficiente de la misma y la adopción de obligaciones por parte de los prestadores que garanticen dicha transparencia.

- Reducción (sic) de las prácticas de manipulación, mediante un compromiso de los firmantes contra esas prácticas, tanto las actuales y ya consolidadas como las emergentes o futuras.

- Empoderamiento de los usuarios, a través de un mejor acceso a la información confiable y mediante la adopción de más y mejores herramientas para identificar y señalar la desinformación y reaccionar ante ella.

- Accesibilidad de los investigadores, lo que supone un apoyo a esa labor de los expertos independientes y un acceso más amplio y sencillo a los datos de las plataformas y a los algoritmos (transparencia).

- Verificación de datos, reconociendo y utilizando de modo consecuente los resultados de la actividad de los verificadores y garantizando su soporte financiero.

[20] https://ec.europa.eu/newsroom/dae/redirection/document/112678

El cumplimiento por parte de las VLOP y VLOSE de los compromisos asumidos en virtud del Código ha de evaluarse mediante una auditoría independiente anual (prevista en el artículo 37 del Reglamento).

Se contempla la creación de un grupo de trabajo permanente con la participación de los diferentes firmantes. Este "foro de múltiples partes interesadas", presidido por la Comisión, establece un "Sistema de Respuesta Rápida" (SRR) de aplicación en circunstancias especiales, como períodos electorales o situaciones de crisis, así como una serie de "indicadores estructurales" para evaluar el impacto del código.

El Centro de Transparencia creado en el marco del Código pone a disposición pública los informes de los firmantes del Código de Conducta sobre la Desinformación[21], en los que detallan las medidas que están adoptando en virtud del mismo para combatir la propagación de la desinformación en línea.

Hay que recordar que, según el Código, las grandes plataformas en línea firmantes se comprometieron a informar cada seis meses sobre sus acciones, mientras que el resto de firmantes informan una vez al año.

El Código menciona, como instancias y herramientas de evaluación y control:

- El Centro de Transparencia, que ofrece información (resumida) sobre su implementación, accesible al público.

- El Grupo de trabajo Permanente (Task force) creado para desarrollar y adaptar el Código, en el que participan todos los firmantes, los supervisores audiovisuales (ERGA / EBMS, EDMO, SEAE) y presidido por la Comisión.

- El marco de seguimiento y monitorización, con la elaboración informes de carácter cuantitativo y cualitativo sobre la aplicación del Código en el ámbito de la UE y los Estados miembros.

A fecha de febrero de 2025 el código contaba con 42 firmantes, no sólo grandes plataformas en línea y motores de búsqueda[22], sino también plataformas en línea de menor tamaño y/o

[21] https://disinfocode.eu/reports

[22] Google (para Google Advertising, Search y YouTube), Meta (para Facebook, Instagram, Messenger y WhatsApp), LinkedIn, Microsoft Ads, Microsoft Bing, TikTok y la organización comercial DOT Europe.

especializadas[23], industria publicitaria[24], verificadores[25], organizaciones de la sociedad civil o dedicadas a la investigación[26] y empresas tecnológicas[27].

Observaciones generales

En su preámbulo, los firmantes del Código:

- Reconocen la importancia de su papel a la hora de contribuir a la lucha contra la desinformación y las noticias falsas, así como contra las operaciones de influencia y las injerencias extranjeras[28].
- Consideran que la desinformación es un desafío para nuestras sociedades democráticas abiertas, que dependen de debates públicos que permitan ciudadanos bien informados para expresar su voluntad a través de procesos políticos libres y justos.

---

[23] Twitch, Vimeo, Seznam, The Bright App.

[24] Asociación Europea de Agencias de Comunicación (EACA), Interactive Advertising Bureau (IAB Europe), DoubleVerify, Ebiquity.

[25] Demagog, European Fact-Checking Standards Network (EFCSN), Faktograf, Maldita, Newtral, Pagella Politica, Science Feedback.

[26] Alliance4Europe, Avaaz, Globsec, Democracy Reporting International (DRI), Debunk EU, CEE Digital Democracy Watch, FIDU (Federación Italiana de Derechos Humanos), Les Surligneurs, Reporteros sin Fronteras (RSF), VOST Europe, WhoTargetsMe.

[27] ActiveFence, Adobe, AI Forensics, Resolver (antes Crisp), Legitimate, Logically, NewsGuard, Valid (antes Daily Ledger), Global Disinformation Index (GDI).

[28] En su versión original, el Código diferencia entre "disinformation" y misinformation", a partir de las definiciones de la EDAP. Ambas apelan a un contenido falso o engañoso, pero la primera se difundiría con intención de engañar u obtener beneficios económicos o políticos, pudiendo causar daño público, mientras que la segunda se compartiría sin inención de causar daño (e incluso "de buena fe"), generalmente en entornos cercanos, aun cuando sus efectos puedan llegar a ser perjudiciales. En este documento hemos optado por traducir misinformation como "bulo" o "noticia falsa", atendiendo a su carácter más particular, no planificado o estratégico, y no tanto a la intencionalidad última de sus responsables.
Por otra parte, se consideran "operaciones de influencia" las acciones coordinadas de actores nacionales o extranjeros en el ámbito de la comunicación para condicionar a un público objetivo utilizando diversos medios engañosos, incluyendo tanto la desinformación como la supresión de fuentes de información independientes. En cuanto a la "injerencia extrajera", se refiere a las acciones coercitivas y engañosas para perturbar la libre formación y expresión de la voluntad política de las personas por parte de un actor estatal extranjero o de sus agentes, a menudo como parte de una operación híbrida más amplia.
Hay que tener en cuenta que la publicidad engañosa no se considera desinformación, ni el uso de la ironía, la sátira y la parodia. Tampoco los errores documentales, ni las informaciones u opiniones claramente identificadas como partidistas.

- Se manifiestan, "como la propia Comisión Europea", a favor de la libertad de expresión, la libertad de información y la privacidad, buscando el equilibrio ("delicado", señalan) entre la protección de los derechos fundamentales y la adopción de medidas efectivas para limitar la propagación y el impacto de contenidos ilícitos en materia de desinformación, aplicando el principio de proporcionalidad.

El Código se estructura en compromisos, cada uno de ellos con diferentes medidas, y cada medida con sus respectivos indicadores cuantitativos (ILS) y cualitativos (ECC), referidos a los productos, actividades y servicios que los firmantes y sus filiales ofrecen, y establecidos con criterios de relevancia y pertinencia. Si un firmante no suscribe un compromiso o medida por no ser relevante o pertinente para sus servicios, debe justificar los motivos de la no suscripción.

El Código de Conducta plantea una continuidad con el Código de Buenas prácticas suscrito en 2018 y actualizado en 2020, si bien los firmantes reconocen la necesidad de su actualización. En esa línea, acuerdan revisar regularmente los compromisos y medidas contemplados en el Código, bien para incluir en sus servicios medidas adicionales que consideren relevantes, pertinentes y practicables, bien para eliminarlas si dejan de serlo.

Los firmantes se comprometían a implementar las medidas previstas en el Código en un plazo de seis meses desde su firma, y proporcionar al mes informes de referencia siguiente a la Comisión Europea, en los que se detalla cómo se han implementado los compromisos asumidos.

Los firmantes reconocen que el papel a desempeñar por el Código es complementario a las obligaciones establecidas en el Reglamento de Servicios Digitales y en otras normas de aplicación. Ello significa que dichas obligaciones legales deben prevalecer siempre sobre las que puedan establecerse en el Código, pero también que los firmantes acuerdan garantizar que los compromisos voluntarios contemplados en el Código aporten un valor añadido a dichas obligaciones.

Centrándonos en sus áreas clave, el Código de Conducta señala lo siguiente:

Desmonetización publicitaria

Compromiso 1

Colaborar con los anunciantes y agencias que participan en la compra de espacio publicitario; con los editores y plataformas que participan en la venta de espacio publicitario y en la aprobación de campañas publicitarias; con las empresas de tecnología publicitaria, que participan en la segmentación o selección de espacio publicitario y/o del contenido, así como en la elaboración de informes de verificación, y con los organismos de auditoría que participan en la acreditación de servicios que abarcan desde la segmentación hasta la elaboración de informes, para garantizar que la industria publicitaria no emplace sus comunicaciones comerciales en campañas de desinformación, ni participe en la compraventa de espacio publicitario de páginas de desinformación. Ello requiere de una mayor eficiencia y disponibilidad de las herramientas de seguridad de marca.

Desfinanciar la difusión de desinformación, mejorando las políticas y los sistemas que determinan la elegibilidad del contenido para ser monetizado; los controles para la monetización y la colocación de anuncios, y los datos para informar sobre la precisión y eficacia de los controles y servicios en torno a la colocación de anuncios.

Desarrollar una metodología para informar sobre las iniciativas de desmonetización, incluyendo datos relacionados con el volumen de publicidad que sustenta las fuentes de desinformación, y presentando una propuesta al Grupo de Trabajo en el plazo de seis meses. Divulgar y describir las políticas implementadas y aplicadas, cuantificando las iniciativas por servicios, estados miembros o idiomas.

Endurecer los requisitos de selección y los procesos de revisión de contenido para su monetización mediante el reparto de ingresos publicitarios, prohibiendo la participación de actores que sistemáticamente publiquen contenido o participen en comportamientos vinculados con la desinformación.

Describir el proceso seguido en cada caso, informando sobre el número de revisiones de políticas y/o actualizaciones de políticas relevantes para el cumplimiento de esta medida, así como sobre el número de cuentas o dominios a los que se les ha prohibido participar en la publicidad o beneficiarse de su monetización.

Adoptar medidas comercial y técnicamente viables para brindar a los compradores de publicidad transparencia en el emplazamiento de su publicidad, informándoles sobre los controles establecidos.

Recurrir a vendedores de espacio publicitario que hayan tomado medidas efectivas y transparentes para evitar el emplazamiento de anuncios junto a contenido de desinformación o en espacios que publiquen desinformación de forma reiterada.

Proporcionar a los auditores externos independientes acceso a las medidas adoptadas, con el fin de que éstos evalúen su efectividad a través de sus informes[29], informando de las áreas en las que se ha producido esa acreditación independiente.

Promover el uso de herramientas de seguridad de marca, integrando la información y los análisis de evaluadores de fuentes; servicios que proporcionen indicadores de fiabilidad; verificadores de datos; investigadores, u otras partes interesadas, con el fin de ayudar a los compradores de anuncios (anunciantes, agencias, empresas de tecnología publicitaria, plataformas de medios, editores) a tomar decisiones informadas para la planificación, compra y emplazamiento publicitario evitando los contenidos de desinformación.

Describir el proceso de integración de las herramientas de seguridad de marca, indicando qué porcentaje de su inversión en medios está protegida por dichas herramientas. Proporcionar información razonable sobre los criterios de calificación de los sitios web y de su actuación neutral, otorgando a los editores el derecho a réplica antes de la publicación de las calificaciones.

Compromiso 2

Prevenir el uso indebido de los sistemas publicitarios para difundir mensajes de desinformación, desarrollando e implementando políticas en esta área adecuadas y personalizadas. Divulgar y describir dichas políticas, enlazando con páginas de acceso público que contengan esa información en sus centros de ayuda. Cuantificar las medidas adoptadas diferenciadas por Estado miembro o idioma.

Desarrollar herramientas y métodos, o alianzas con otras entidades como las arriba señaladas, para identificar contenidos y fuentes que distribuyan desinformación. describiendo esas herramientas y métodos. Informar sobre los anuncios eliminados o prohibidos en sus servicios, el contenido y los infractores, por Estado miembro.

[29] Acreditación de Seguridad de Marca a Nivel de Contenido del MRC, certificaciones de Seguridad de Marca TAG u otras certificaciones aceptadas por la industria con reconocimiento similar.

Proporcionar información relevante a los anunciantes cuando se rechacen o eliminen anuncios, o se deshabiliten cuentas publicitarias, sobre qué políticas publicitarias se han infringido, aclarando los procedimientos de apelación. Describir cómo proporcionan esa información e informar, por Estado miembro, sobre el número de apelaciones recibidas de los anunciantes por la aplicación de estos procedimientos, así como sobre la proporción de las que llevaron a un cambio en la decisión inicial.

Compromiso 3

Intercambiar mejores prácticas y fortalecer la cooperación con otros actores en la cadena de valor de monetización en línea, como los servicios de pago electrónico, plataformas de comercio electrónico y entidades de financiación colectiva/donación, con el objetivo de aumentar la efectividad de estas medidas. Facilitar la integración y el flujo de información (sin menoscabo del cumplimiento pleno de las normas de protección de datos y de los acuerdos de confidencialidad pertinentes). Intercambiar información sobre tendencias en la desinformación, nuevas amenazas observadas y casos prácticos, con el fin de mejorar las capacidades y las medidas a adoptar, a través del Grupo de Trabajo del Código u otros foros relevantes[30]. Informar sobre los foros en los que han participado, el tipo de información compartida, el aprendizaje obtenido.

Transparencia de la publicidad política

Compromiso 4

Adoptar durante el primer año de vigencia del Código una definición común de publicidad política y temática, de acuerdo con la definición establecida en el Reglamento 2024/900, del Parlamento Europeo y del Consejo, de 13 de marzo de 2024, sobre transparencia y segmentación en la publicidad política. Si no fuera posible, establecer definiciones operativas de publicidad política y publicidad temática en el marco del Grupo de Trabajo.

Compromiso 5

Indicar claramente en sus políticas publicitarias hasta qué punto dicha publicidad política está permitida o prohibida en sus servicios, publicando sus normas o directrices

[30] GARM, IAB Europe u otros.

de un modo accesible fácilmente comprensible, y aplicando los principios de etiquetado eficiente, de transparencia y de verificación en todos los anuncios relevantes.

Compromiso 6

Identificar los anuncios políticos o temáticos como contenido pago, de forma que no induzcan a error a los usuarios. Para ello, desarrollarán un conjunto de mejores prácticas comunes y ejemplos para marcas y etiquetas en dichos anuncios, asegurándose de que la información relevante, como la identidad del patrocinador, se incluya en la etiqueta o sea fácilmente accesible para el usuario desde la etiqueta.

Publicar sus diseños de etiquetado, con ejemplos de cómo la identidad del patrocinador u otra información relevante se adjunta a los anuncios o se ofrece fácilmente accesible para los usuarios desde la etiqueta. Participar y contribuir económicamente a la realización de investigaciones para para comprender cómo los usuarios identifican y comprenden las etiquetas en los anuncios políticos o temáticos, de modo que los resultados permitan mejorar esa identificación y comprensión.

Publicar métricas significativas, a nivel de Estado miembro, sobre el volumen de anuncios aceptados y etiquetados, cantidades gastadas por los anunciantes etiquetados u otras que puedan determinarse.

Garantizar que, una vez que un anuncio político o temático se etiquete como tal en su plataforma, la etiqueta permanezca vigente cuando los usuarios compartan el mismo anuncio en la misma plataforma, de modo que siga identificándose claramente como contenido político o temático de pago.

Realizar esfuerzos razonables para mejorar la visibilidad de las etiquetas aplicadas a la publicidad política compartida a través de servicios de mensajería, siempre que sea posible y en cumplimiento con la ley local, desarrollando soluciones que faciliten a los usuarios el reconocimiento sin debilitar el cifrado y con el debido respeto a la protección de la privacidad.

Compromiso 7

Implementar sistemas de verificación de identidad proporcionados y adecuados para los patrocinadores y proveedores de servicios publicitarios que actúen en nombre de los patrocinadores que publiquen anuncios políticos o temáticos. Garantizar que se cumplen los requisitos de etiquetado y transparencia para el usuario antes de permitir la publicación de dichos anuncios.

Informar sobre las herramientas y procesos establecidos para que los patrocinadores, o los proveedores de servicios publicitarios que actúen en nombre de los patrocinadores, declaren si el servicio publicitario que solicitan constituye publicidad política o temática.

Publicar datos sobre la efectividad y proporcionalidad de las medidas adoptadas para verificar la identidad de los patrocinadores de anuncios políticos o temáticos. Publicar métricas significativas sobre el volumen de anuncios rechazados por incumplimiento de los procesos de verificación pertinentes, por servicio y a nivel de Estado miembro.

Informar sobre las medidas adoptadas contra los actores que evadan de forma demostrable las herramientas y procesos implementados, como suspensiones u otras sanciones que afecten a sus cuentas.

Proporcionar a los usuarios herramientas y funcionalidades que les permitan marcar los anuncios políticos que no estén etiquetados como políticos.

Compromiso 8

Proporcionar información transparente[31] a los usuarios sobre los anuncios políticos o temáticos que ven en su servicio, como la identificación del patrocinador, el período de exhibición, el gasto publicitario y la información agregada sobre los destinatarios del anuncio, proporcionando un enlace directo desde el anuncio al repositorio de los mismos. Publicar las obligaciones mínimas comunes de transparencia.

Compromiso 9

Proporcionar a los usuarios información clara, sencilla, comprensible y completa sobre el motivo por el que ven un anuncio político o temático, así como sobre qué datos utilizan los patrocinadores y los proveedores (demográficos, geográficos, contextuales, de intereses o basados en el comportamiento) para determinar que un anuncio político o temático se muestra específicamente al usuario.

Describir las herramientas y funciones implementadas para proporcionar a los usuarios la información descrita, incluyendo ejemplos relevantes para cada método de segmentación ofrecido por el servicio.

---

[31] Véase el Reglamento (UE) 2024/900 del Parlamento Europeo y del Consejo, de 13 de marzo de 2024, sobre transparencia y segmentación en la publicidad política.

Compromiso 10

Mantener repositorios con funcionalidades mínimas y aplicaciones para acceder a datos de anuncios políticos o temáticos y a asegurar su actualidad, integridad, usabilidad y calidad, de tal manera que contengan toda la publicidad política y temática servida, junto con la información necesaria para cumplir con sus obligaciones legales y con los compromisos de transparencia.

Los repositorios tendrán capacidad de búsqueda (lo más cerca posible del tiempo real, en particular durante los períodos electorales) de todos los anuncios políticos y temáticos publicados, con información relevante para cada anuncio, como la identificación del patrocinador; las fechas de emisión del anuncio; el importe total invertido en el anuncio; la cantidad de impresiones generadas; los criterios de audiencia utilizados para determinar los destinatarios; los datos demográficos y la cantidad de destinatarios que vieron el anuncio, y las zonas geográficas en las que se vio. Esta información estará disponible públicamente durante al menos 5 años.

Detallar la disponibilidad, las características y la cadencia de actualización de los repositorios, proporcionando información cuantitativa sobre el uso periódico (por ejemplo, mensual) de los mismos.

Compromiso 11

Proporcionar un amplio acceso y disponibilidad a los usuarios e investigadores para realizar búsquedas personalizadas en los repositorios, a partir de un conjunto de funcionalidades y de criterios de búsqueda mínimos en formatos estándar. Incluyendo, por ejemplo, búsquedas por anunciante o candidato, por área geográfica o país, por idioma, palabra clave, elección u otros criterios de segmentación, para permitir la investigación y el seguimiento.

Informarán sobre la interacción con los investigadores, incluyendo su experiencia con las funcionalidades de las API, y sobre las mejoras resultantes de dicha interacción.

Compromiso 12

Investigar, monitorear e informar sobre el uso de la publicidad política o temática en línea en los Estados miembros mediante herramientas, paneles y otros datos para asegurar un análisis adecuado de la misma, particularmente durante los períodos electorales.

Alertar a otros firmantes sobre problemas en la implementación o cumplimiento de las políticas de publicidad política o temática o de este Código, así como en el seno del Grupo de Trabajo del Código.

Compromiso 13

Identificar riesgos nuevos y cambiantes de desinformación en los usos de la publicidad política o temática de forma particular, en el seno del Grupo de Trabajo y con otras partes interesadas para evaluar la oportunidad y el impacto de los períodos de silencio o veda electoral para la publicidad política o temática en sus servicios en todos los Estados miembros.

Evaluar, en esos marcos, si existen suficientes análisis independientes de la publicidad política o temática en los Estados miembros. Para ello se convocará al menos una vez al año, a través del Grupo de Trabajo, un debate sobre riesgos nuevos en la publicidad política para desarrollar políticas coordinadas; períodos de silencio; análisis independiente de la publicidad política o temática.

Reducción de las prácticas de manipulación

Compromiso 14

Elaborar un listado terminológico (glosario) común de los comportamientos, actores y prácticas de manipulación según servicios, en el seno del Grupo de Trabajo en los primeros seis meses tras la firma del Código, que debería revisarse al menos anualmente a la luz de la evidencia más reciente sobre las TTP empleadas por actores maliciosos[32], como la creación y el uso de cuentas falsas; apropiaciones de cuentas y amplificación impulsada por bots; operaciones de piratería y filtración; suplantación de identidad; ultrafalsificaciones (deep fakes); compra de compromisos falsos; mensajes pagados o no transparentes o promoción por parte de personas influyentes; creación y uso de cuentas falsas coordinadas; conductas de usuario dirigidas a amplificar artificialmente el alcance o el apoyo público percibido para la desinformación.

[32] Marco de Tácticas, Técnicas y Procedimientos de Desinformación de AMITT,

Adoptar, reforzar e implementar políticas claras respecto a estos comportamientos, enumerando en detalle las acciones proactivas desarrolladas, disponibles públicamente, aclarando qué comportamientos y prácticas están prohibidos en sus servicios.

Aportar indicadores y métricas adicionales a nivel de Estado miembro sobre: el número de casos de TTP identificados y el tipo de contenido; el impacto/eficacia de las medidas adoptadas; la penetración estimada y el impacto que las cuentas falsas/no auténticas entre los usuarios (tendencias en las audiencias objetivo; narrativas utilizadas, etc.); visualizaciones/impresiones e interacciones ("me gusta", "compartir", comentarios), en relación con cada TTP identificado, antes y después de que se adoptara la correspondiente medida.

Compromiso 15

Quienes desarrollen u operen sistemas de IA, o que difundan IA, establecerán políticas para contrarrestar las prácticas generativas de manipulación de contenidos, advirtiendo a los usuarios sobre cómo detectar proactivamente dicho contenido[33].

Garantizar que los algoritmos utilizados para la detección, moderación y sanción de conductas y contenidos inadmisibles sean confiables, respeten los derechos de los usuarios finales y no constituyan prácticas de manipulación prohibidas que distorsionen su comportamiento.

Compromiso 16

Establecer canales de intercambio para compartir proactivamente información sobre operaciones de influencia entre plataformas, interferencias extranjeras en el ámbito de la información e incidentes relevantes que surjan en sus respectivos servicios, con el objetivo de prevenir su difusión o el resurgimiento en otros firmantes. Para ello compartirán información relevante a través de otros foros, divulgando su existencia y los aprendizajes derivados de dicho intercambio.

Informar sobre el número de medidas adoptadas como resultado de la colaboración y el intercambio de información, especificando, si es posible, qué Estados miembros se han visto afectados. Compartir información sobre la migración de actores conocidos de la

---

[33] Véase el Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Ley de Inteligencia Artificial).

desinformación a través de diferentes plataformas como una forma de eludir las políticas de moderación, involucrando a diferentes audiencias o coordinando acciones en plataformas con menos supervisión. Como resultado de este monitoreo, se compartirán ejemplos y estudios de caso, según lo observado por los equipos de moderación y/o por colaboradores externos (expertos académicos, verificadores).

Empoderamiento de los usuarios

Compromiso 17

Reforzar sus acciones en el área de la alfabetización mediática y el pensamiento crítico, también con el objetivo de incluir a los grupos vulnerables[34], diseñando e implementando herramientas a tal efecto, por ejemplo, proporcionando a los usuarios contexto sobre los contenidos y la posibilidad de evaluar dichos contenidos desde el punto de vista de su veracidad. Describir las herramientas utilizadas e informar sobre su implementación en cada Estado miembro, con métricas pertinentes para evaluar sus efectos: recuento total de impresiones de las herramientas; información sobre las interacciones con ellas.

Desarrollar, promover y/o apoyar o mantener campañas para crear conciencia sobre la desinformación y las TTP que están siendo utilizadas por actores maliciosos, entre el público general y entre las comunidades vulnerables. Describir las actividades lanzadas o apoyadas y en qué Estados miembros; el número de actividades de alfabetización mediática y sensibilización organizadas y/o en las que hayan participado; lugares donde se llevaron a cabo campañas, alcance de las mismas y efectos obtenidos: participación, interacciones online.

Colaborar con expertos en alfabetización mediática en el diseño, implementación y medición del impacto de las herramientas, como es el caso del Grupo de Expertos en Alfabetización Mediática de la Comisión Europea, el Grupo de Acción en Alfabetización Mediática de ERGA, EDMO y sus sucursales específicas de cada país, o las universidades u organizaciones de los Estados miembros con experiencia acreditada, describiendo las características y objetivos de esa asociación o colaboración.

---

[34] A la luz de las iniciativas de la Comisión Europea en el área de la alfabetización mediática, incluido el nuevo Plan de Acción de Educación Digital.

Compromiso 18

Mitigar los riesgos de propagación viral de desinformación mediante la adopción de prácticas de "diseño seguro" de la arquitectura de los servicios, políticas transparentes y rendición de cuentas de los sistemas de recomendación, diseñando dichos sistemas para mejorar la preminencia de la información autorizada y aplicando otros enfoques sistémicos en el diseño de sus productos, políticas o procesos, como las pruebas previas. Informar sobre su implementación en cada Estado miembro, detallando los principales parámetros de sus sistemas de recomendación y proporcionando datos y métricas significativas, así como una estimación de la efectividad de las medidas y sus efectos.

Desarrollar políticas proporcionadas, de acceso público, para limitar la propagación de información falsa o engañosa, como prohibir, bajar la clasificación o no recomendar este tipo de información según su gravedad, salvaguardando la libertad de expresión e información, y tomando medidas contra páginas web o actores que violen persistentemente estas políticas. Informar sobre las acciones realizadas en respuesta a las violaciones de las medidas adoptadas a nivel de Estado miembro, incluyendo en las métricas el número total de infracciones, y de medidas para medir su impacto (visibilidad e interacción con el contenido).

Participar en investigaciones sobre la difusión de desinformación, poniendo sus resultados a disposición del público y del Grupo de Trabajo del Código, con el fin de mejorar las prácticas y características de diseño seguro existentes o desarrollar otras nuevas. Describir estas investigaciones, las organizaciones externas con las que se colabora, las aportaciones económicas realizadas y los resultados conseguidos.

Compromiso 19

Transparencia de los sistemas de recomendación en cuanto a los principales criterios y parámetros utilizados para priorizar o despriorizar la información, poniendo a disposición de los usuarios, de forma clara, accesible y fácilmente comprensible, el conocimiento de esos sistemas (en los términos y condiciones).

Ofrecer la posibilidad de que los destinatarios del servicio seleccionen y modifiquen en cualquier momento sus preferencias de recomendación. Proporcionar información agregada sobre las configuraciones efectivas de los usuarios; número de veces que han interactuado activamente con estas configuraciones; cambios realizados en los patrones de configuración.

Compromiso 20

Capacitar a los usuarios con herramientas para evaluar la procedencia y el historial de edición o la autenticidad o precisión del contenido digital. Desarrollar soluciones tecnológicas que ayuden a los usuarios a verificar la autenticidad o identificar la procedencia o la fuente del contenido digital.

Adoptar medidas para unirse y apoyar iniciativas globales y organismos de normalización[35] enfocados en el desarrollo de herramientas de identificación de procedencia. Proporcionar detalles de esas iniciativas o del apoyo brindado a organizaciones, aportando los enlaces a sus sitios web.

Mejorar el equipamiento de los usuarios para identificar la desinformación; en particular, facilitar, en todos los idiomas de los Estados miembros en los que prestan sus servicios los firmantes, el acceso de los usuarios a herramientas para evaluar la exactitud fáctica de las fuentes mediante la información proporcionada por los verificadores, así como etiquetas de advertencia de otras fuentes autorizadas.

Compromiso 21

Desarrollar y aplicar políticas, funciones o programas en todos los Estados miembros y los idiomas de la UE para ayudar a los usuarios a beneficiarse del contexto y la información proporcionada por verificadores independientes o fuentes autorizadas. Por ejemplo, mediante etiquetas como las que indican las calificaciones de los verificadores; avisos a los usuarios que intentan compartir o han compartido previamente el contenido calificado; paneles informativos.

Actuar sobre el contenido notificado por los verificadores de datos que incumple sus políticas sobre desinformación. Informar sobre las acciones implementadas para cumplir con esta medida, y sobre su disponibilidad en los Estados miembros.

Informar sobre los verificadores de datos independientes con los que trabajan para etiquetar el contenido en sus servicios (a menos que una organización de verificación se oponga basándose en un temor razonable a represalias o violencia), los idiomas en los que operan, las políticas bajo las que trabajan y el etiquetado aplicado.

---

[35] (por ejemplo, C2PA).

Aportar métricas sobre las acciones desarrolladas a nivel de Estado Miembro, incluyendo en su caso información sobre el alcance de las etiquetas o verificaciones de datos y de otras fuentes autorizadas (impresiones totales de verificaciones de datos; proporción con respecto a las impresiones originales del contenido verificado, u otras métricas explicando su pertinencia) y su impacto (número de artículos publicados por verificadores de datos independientes; número de etiquetas aplicadas al contenido, por ejemplo, sobre la base de dichos artículos; métricas significativas sobre el impacto de las acciones realizadas y su impacto en las interacciones de los usuarios, o las veces que éstos comparten, contenido verificado como falso o engañoso).

Emprender o apoyar, a la luz de la evidencia científica y las especificidades de sus servicios y de las preferencias de privacidad de los usuarios, investigaciones y pruebas sobre advertencias o actualizaciones dirigidas a usuarios que han interactuado con contenido que posteriormente ha sido objeto de acciones por violación de las políticas de desinformación. Informar sobre esas iniciativas o apoyos y sobre sus resultados poniéndolos a disposición del público general.

Informar sobre los procedimientos para desarrollar e implementar sistemas de etiquetado o advertencia y sobre cómo tienen en cuenta la evidencia científica y las necesidades de los usuarios para maximizar su utilidad.

Compromiso 22

Proporcionar a los usuarios herramientas de ayuda para tomar decisiones informadas, cuando encuentren información en línea que pueda ser falsa o engañosa, que permitan evaluar la fiabilidad de las fuentes de información: indicadores de fiabilidad (marcas de confianza centradas en la integridad de la fuente y la metodología detrás de dichos indicadores) en colaboración con terceros independientes como los medios de comunicación, las asociaciones de periodistas, las organizaciones de libertad de prensa, verificadores y otras entidades que pueden ayudar a los usuarios a tomar esas decisiones informadas.

Informar sobre cómo permiten que los usuarios de sus servicios se beneficien de los indicadores o marcas de confianza, así como sobre el porcentaje de éstos que han habilitado el indicador de fiabilidad en cada Estado miembro.

Ofrecer a los usuarios la opción de disponer de señalizaciones relacionadas con la fiabilidad de las fuentes de los medios de comunicación en los sistemas de recomendación. Informar sobre si introducen dichas señalizaciones y, en ese caso, cómo lo hacen, explicando la justificación de su enfoque, proporcionando a los usuarios los detalles de las políticas adoptadas y los principales parámetros que emplean sus sistemas de recomendación.

Garantizar que las fuentes de información se revisen de manera transparente, apolítica, imparcial e independiente, con criterios aplicados por igual a todas las fuentes y permitiendo auditorías independientes por parte de autoridades reguladoras u otros organismos competentes. Proporcionar ejemplos de esa aplicación a una gama representativa de diferentes editores. Informar sobre quién contribuyó a la evaluación de la fuente o qué organismo de certificación la evaluó.

Proporcionar mecanismos de cumplimiento y corrección que respeten el derecho de los editores a ser escuchados y a participar en el proceso de evaluación antes de que se apliquen los indicadores y tener sus respuestas a disposición de los consumidores después de que se publiquen las evaluaciones.

Informar sobre las cifras que miden el volumen de tráfico de fuentes confiables, generado gracias a los indicadores de confiabilidad y sobre las correcciones periódicas en sus calificaciones o indicadores si se producen actualizaciones o errores. Informar sobre el número de editores que han mejorado sus prácticas periodísticas y sus puntuaciones de confiabilidad después de ser evaluados. Proporcionar ejemplos de intercambios con los editores.

Informar sobre los indicadores de confiabilidad mediante estándares europeos voluntarios, autorreguladores y certificables[36], desarrollados y revisados basándose en las mejores prácticas y normas éticas aceptadas internacionalmente, poniéndolos a disposición del público accesibles de una manera no propietaria y neutral[37]. Publicar los resultados de la autoevaluación y certificación, así como las estadísticas y análisis, incluyendo la gestión de quejas.

Diseñar e implementar productos y funciones (p. ej., paneles informativos, banners, ventanas emergentes, mapas y avisos, indicadores de fiabilidad) que dirijan a los usuarios a fuentes fidedignas sobre temas de especial interés público y social o en situaciones de crisis, especificando si están disponibles en todos los Estados miembros.

Informar sobre el alcance o las interacciones de los usuarios con los productos o funciones, a nivel de Estado miembro, a través de las métricas de impresiones e interacciones (clics, tasas de clics, acciones, según corresponda a las herramientas y servicios).

---

[36] Estándares técnicos como el CWA17493:2019.

[37] De acuerdo con la Acreditación Europea y el Reglamento (CE) n.º 765/2008, del Parlamento Europeo y del Consejo, de 9 de julio de 2008, por el que se establecen los requisitos de acreditación y vigilancia del mercado relativos a la comercialización de los productos.

Compromiso 23

Proporcionar a los usuarios una funcionalidad fácil de usar para marcar información falsa o engañosa en todos sus servicios y en todos los idiomas de los Estados miembros que infrinja sus políticas o sus términos de servicio. Esta funcionalidad debería permitir acciones de seguimiento apropiadas, proporcionadas y coherentes, con pleno respeto a la libertad de expresión.

Informar sobre la disponibilidad de sistemas de alerta para sus políticas relacionadas con información falsa o engañosa perjudicial en los Estados miembros de la UE, especificando los diferentes pasos necesarios para activar los sistemas. Adoptar las medidas necesarias para garantizar que esta funcionalidad esté debidamente protegida contra el abuso humano o informático (por ejemplo, la táctica de alerta masiva para silenciar otras voces).

Informar sobre las medidas generales adoptadas para garantizar la integridad de los sistemas de denuncia y recurso de apelación, evitando al mismo tiempo divulgar datos que puedan ayudar a explotar vulnerabilidades en sus defensas en el desarrollo de prácticas abusivas e ilícitas.

Compromiso 24

Informar a los usuarios, cuyas cuentas o contenidos difundidos hayan sido objeto de medidas por infracción de las políticas de desinformación, sobre el motivo por el cual han sido etiquetados, degradados o se les ha aplicado alguna otra medida. Contar con un mecanismo de apelación transparente contra dichas medidas que permita gestionar las quejas de manera oportuna, diligente, transparente y objetiva, y revertir la acción sin demora indebida cuando la queja se considere fundada.

Informar sobre la disponibilidad de esos sistemas de notificación y apelación en los diferentes Estados miembros e idiomas, proporcionando detalles sobre los pasos del procedimiento de apelación: número y naturaleza de las medidas de cumplimiento de las políticas descritas; número de medidas que fueron posteriormente apeladas: resultados de estas apelaciones; métricas, en la medida de lo posible, que proporcionen información sobre la duración o la eficacia de la tramitación del proceso de apelación.

Compromiso 25

En el caso de las aplicaciones de mensajería, desarrollar e implementar herramientas y funciones que ayuden a los usuarios a identificar la desinformación, sin debilitar el cifrado

y con el debido respeto a la protección de la privacidad, con el fin de limitar el reenvío y viralización de dicha desinformación a través de las conversaciones, haciendo visible ese etiquetado.

Informar sobre las herramientas, políticas, alianzas, programas y campañas implementadas para cumplir con esta medida y sobre su disponibilidad en los Estados miembros, incluyendo, cuando sea posible, detalles relevantes sobre los resultados obtenidos: número de reclamaciones presentadas por los usuarios a los verificadores de datos, y alcance de los controles realizados por la plataforma en función de las reclamaciones presentadas.

Contar con colaboradores externos como verificadores, expertos y organizaciones de la sociedad civil para desarrollar campañas que ayuden a los usuarios a reflexionar sobre la necesidad de evitar promover la desinformación reenviando los mensajes que reciben. Participar con estos colaboradores en investigaciones basadas en la evidencia sobre el uso y el impacto de las herramientas, funciones y campañas implementadas para cumplir con las medidas.

Empoderamiento de los investigadores

Compromiso 26

Proporcionar un acceso estable, continuo, en tiempo real o casi real y con capacidad de búsqueda, a datos anonimizados, agregados o manifiestamente públicos pertinentes para fines de investigación sobre desinformación. Contar para ello con medios automatizados como API u otras soluciones técnicas abiertas y accesibles que permitan el análisis de dichos datos.

Describir las herramientas y procesos para proporcionar acceso público a estos datos, así como las salvaguardias establecidas para abordar los riesgos de su uso inadecuado. Proporcionar datos sobre cómo se ha concretado la participación de investigadores y organizaciones sociales sin ánimo de lucro que puedan realizar esa labor.

Detallar los protocolos técnicos que se utilizarán para acceder a los puntos de datos disponibles, en el centro de ayuda correspondiente, proporcionando información técnica y metodológica sobre cómo se crearon esos puntos y la representatividad de sus datos. Proporcionar un acceso legible a dichos datos, describiendo el proceso de solicitud establecido para acceder a los mismos.

Proporcionar métricas significativas sobre la adopción, la rapidez y la aceptación de las herramientas y los procesos de estas medidas, como el número de usuarios mensuales

(durante un período representativo definido como muestra); número de solicitudes recibidas, rechazadas y aceptadas; tiempo promedio de respuesta.

Implementar procedimientos para que los investigadores informen sobre el posible mal funcionamiento de los sistemas de acceso, para restablecer el mismo y, en su caso, reparar las funcionalidades defectuosas en un plazo razonable. Informar sobre las incidencias ocurridas, sobre su procedimiento de respuesta y el tiempo necesario para solucionarlas.

Compromiso 27

Proporcionar a los investigadores acceso a los datos, en cooperación con un organismo tercero independiente que pueda examinar a los investigadores y las propuestas de investigación. Colaborar con otras entidades (Comisión Europea, organizaciones sociales, DPA) para desarrollar ese organismo independiente, teniendo en cuenta la propuesta de EDMO de contar con un Código de Conducta sobre el Acceso a los Datos de la Plataforma.

Cofinanciar el desarrollo del organismo independiente, revelando su financiación y comprometiéndose a cooperar con él para permitir el intercambio de datos personales necesarios para realizar investigaciones sobre desinformación con investigadores aprobados, de conformidad con los protocolos que definirá el organismo. Describir dicha cooperación.

Detallar los programas en los que han participado; cuántos proyectos de investigación han sido aprobados por el organismo independiente sobre los presentados, su naturaleza y los resultados obtenidos.

Participar en programas piloto para compartir datos con investigadores sin esperar a que el organismo independiente externo esté completamente establecido, que operarán de acuerdo con todas las leyes aplicables en materia de intercambio/uso de datos.

Compromiso 28

Apoyar la investigación ("de buena fe") sobre la desinformación que involucre sus servicios. Contar con los recursos humanos adecuados para facilitar la investigación, y establecer y mantener un diálogo abierto con los investigadores para realizar un seguimiento de los tipos de datos que serán solicitados por éstos y ayudarles a encontrar puntos de contacto relevantes en sus organizaciones.

Describir los recursos y procesos implementados para facilitar la investigación e interactuar con la comunidad investigadora, incluyendo API, equipos dedicados, herramientas, centros

de ayuda, programas o eventos. Aclarar qué tipo de datos se ponen a disposición de los investigadores.

No prohibir ni desalentar la investigación sobre desinformación "de buena fe, genuina y de interés público demostrable" en sus plataformas. No emprender acciones adversas contra usuarios investigadores o cuentas que realicen o participen en esas investigaciones. Colaborar con EDMO para realizar una consulta anual a investigadores europeos con el fin de evaluar si han experimentado esas acciones adversas o impedimentos.

Poner a disposición fondos para que los investigadores gestionen y definan de forma independiente las prioridades y los procedimientos de asignación transparentes basados en el mérito científico, una metodología transparente y estándares éticos. Divulgar los recursos puestos a disposición y los procedimientos establecidos para garantizar que los recursos se gestionen de forma independiente.

Compromiso 29

Proporcionar informes sobre las investigaciones y su desarrollo metodológico, y compartir los resultados con las audiencias relevantes (miembros del Grupo de Trabajo, EDMO, ERGA, otros firmantes, y en última instancia público en general).

Colaborar con actividades de investigación que tengan como objetivo determinar la eficacia relativa de diversas medidas de fomento de la resiliencia implementadas en el Código y en otros ámbitos (por ejemplo, etiquetas, advertencias, notificaciones ex post), con vistas a fundamentar futuras intervenciones regulatorias y operativas.

Verificación de datos

Compromiso 30

Establecer un marco para la cooperación transparente, estructurada, abierta, financieramente sostenible y no discriminatoria con la comunidad de verificadores de datos de la UE. Establecer acuerdos para lograr la cobertura de la verificación en todos los Estados miembros, cumpliendo con altos estándares éticos y profesionales; basados en condiciones transparentes, abiertas, consistentes y no discriminatorias, y garantizando la independencia de los verificadores.

Informar y explicar la naturaleza de los acuerdos, los resultados esperados, así como información cuantitativa relevante (por ejemplo: contenidos verificados, cobertura) y estándares y condiciones comunes.

Enumerar las entidades de verificación con las que tienen acuerdos (a menos que una organización de verificación de hechos se oponga a dicha divulgación sobre la base de un temor razonable a represalias o violencia).

Informar sobre los recursos asignados para los acuerdos con las organizaciones de verificación de datos; número total de acuerdos segmentados por Estado miembro, por idioma y, cuando corresponda, por servicio.

Proporcionar contribuciones financieras justas, que podrían adoptar la forma de acuerdos particulares, con múltiples verificadores o con un organismo electo representativo de las organizaciones europeas independientes de verificación de datos que tenga el mandato de celebrar dichos acuerdos. Informar sobre las acciones desarrolladas y sobre los criterios generales utilizados para asegurar esas contribuciones financieras justas por el trabajo realizado; sobre los criterios utilizados en dichos acuerdos para garantizar altos estándares éticos y profesionales, así como la independencia de los verificadores y las condiciones de transparencia, apertura, consistencia y no discriminación de los acuerdos.

Participar en revisiones regulares con sus asociados sobre la naturaleza y efectividad del programa de verificación.

Contribuir a la cooperación transfronteriza entre verificadores de datos, informando sobre las acciones tomadas para facilitar dicha cooperación incluyendo ejemplos de verificaciones de datos, idiomas o Estados miembros en los que se facilitó dicha cooperación.

Consultar, para el desarrollo de estas medidas, con el EDMO y con un organismo electo representativo de las organizaciones europeas independientes de verificación de datos. Informar ex ante sobre los planes para involucrar a dichas entidades y ex post sobre las acciones realizadas, incluyendo el propio desarrollo del marco de cooperación.

Las organizaciones europeas de verificación de datos informarán, directamente, como firmantes del Código) o indirectamente (p. ej. a través de encuestas por EDMO o un organismo electo representante de las organizaciones europeas independientes de verificación de datos) sobre la equidad de las compensaciones individuales que se les proporcionan a través de estos acuerdos.

Compromiso 31

Utilizar ("o usar consistentemente de otra manera") el trabajo de los verificadores en los servicios, procesos y contenidos de sus plataformas, con cobertura completa de todos los Estados miembros e idiomas, incluyendo los generados por los usuarios (UGC). Colaborar con los verificadores para tal fin, comenzando por realizar y documentar investigaciones y pruebas.

Emplear mecanismos rápidos y eficientes para integrar esas verificaciones de datos en sus productos o procesos, como el etiquetado, los paneles informativos o la aplicación de políticas para ayudar a aumentar el impacto de las verificaciones de datos en las audiencias.

Informar sobre sus actividades e iniciativas específicas, incluyendo la metodología aplicada los resultados obtenidos, mediante métricas estimatorias significativas a nivel de Estado miembro y servicio: número de artículos de verificación de datos publicados y su alcance; número de piezas de contenido revisadas; número de piezas de contenido etiquetadas sobre la base de artículos de verificación de datos; interacciones de los usuarios con la información verificada como falsa o engañosa. Ello permitirá a los investigadores, a los verificadores de datos, a la Comisión, al regulador (ERGA /EBMS) y al público comprender y evaluar el impacto de las medidas adoptadas.

Incluir información cuantitativa de referencia que ayude a contextualizar los indicadores, que se presentará y debatirá en el seno del grupo de trabajo permanente.

Crear un repositorio de contenidos verificados en un plazo máximo de 12 meses, en colaboración con el EDMO y con los verificadores (así como su órgano de representación), y gestionado por éstos. Contribuir a su financiación, que se reevaluará anualmente en el Grupo de Trabajo. Informar sobre dicha financiación, así como de otras contribuciones de en apoyo técnico y aportación de recursos, con métricas relevantes.

Explorar dentro del Grupo de Trabajo soluciones tecnológicas para facilitar el uso eficiente de este repositorio común por parte de las diferentes plataformas firmantes, en los diferentes idiomas. Realizar un seguimiento e informar sobre estas soluciones.

Compromiso 32

Proporcionar a los verificadores un acceso rápido y, siempre que sea posible, automatizado, a la información pertinente para maximizar la calidad y el impacto de su actividad, según lo definido en un marco que se diseñará en coordinación con EDMO y el organismo electo representativo de las organizaciones europeas independientes de verificación de datos.

Proporcionar información que les ayude a cuantificar el impacto del contenido verificado a lo largo del tiempo, como las acciones tomadas en función de ese contenido dependiendo del servicio, las impresiones, los clics o las interacciones.

En el caso de difusión de contenidos generados por los usuarios, proporcionar interfaces apropiadas para que los verificadores puedan acceder a la información sobre el impacto de los contenidos en sus plataformas y garantizar la coherencia en la forma en la que loa prestadores firmantes utilizan, acreditan y proporcionan retroalimentación sobre su trabajo. Proporcionar detalles sobre estas interfaces u otras herramientas implementadas para proporcionar a los verificadores la información necesaria, incluyendo información cuantitativa sobre su utilización; por ejemplo, usuarios mensuales.

Intercambiar información regularmente con los verificadores para fortalecer la cooperación, informando sobre los canales de comunicación y los intercambios realizados para ese fin y sobre os resultados (éxito y satisfacción) obtenidos.

Hay que referirse también al

Compromiso 33

que no va dirigido a las plataformas firmantes sino a los verificadores, y que contempla el compromiso de éstos de operar sobre la base de estrictas normas éticas y de transparencia y de proteger su independencia. Para ello, cumplirán con requisitos como ser firmantes del Código de Principios de la International Fact-Checking Network (IFCN)[38] o el Código de Integridad Profesional para Organizaciones Independientes Europeas de Verificación de Datos[39]. Informarán sobre el estado de su membresía y de las acciones tomadas como resultado de ello para asegurar reglas estrictas de ética y transparencia y para proteger su independencia. Informarán sobre el número de verificadores de datos europeos que están certificados por la IFCN o son miembros del futuro Código de Integridad Profesional.

[38] https://ifcncodeofprinciples.poynter.org/

[39] El Código se puso en marcha en 2021, por iniciativa de Maldita.es, con la participación de AFP, Demagog, Correctiv, Pagella Política/Facta y EU Lab, en el marco del consorcio European Fact-Checking Standars Projec.

LOS INFORMES DE TRANSPARENCIA

Desde el año 2023, el Centro de Transparencia ha recogido los informes elaborados por los firmantes del Código de buenas prácticas (actualmente, Código de Conducta en Materia de Desinformación) sobre su grado de cumplimiento de los compromisos contemplados en dicho Código.

En su página web[40] se señala que los firmantes del Código, de acuerdo con el Compromiso 41 del mismo, se comprometieron a desarrollar Indicadores Estructurales, diseñados para evaluar la eficacia del Código en la reducción de la propagación de desinformación en línea para cada signatario pertinente y para todo el ecosistema digital en la UE y a nivel de los Estados miembros. Para ello crearon un Grupo de Trabajo en junio de 2022.

En cuanto a los informes presentados por los principales signatarios del Código, en 2025 han sido los siguientes:

[40] https://disinfocode.eu/

INFORMES DE TRANSPARENCIA VLOP VLOSE 2025	
GOOGLE Google Search https://disinfocode.eu/reports/google-search/5/text	marzo
INFORMES GOOGLE (II) Google Ads https://disinfocode.eu/reports/google-ads/5/text Youtube https://disinfocode.eu/reports/youtube/5/text	marzo/septiembre marzo/septiembre
INFORMES META https://disinfocode.eu/reports/meta/5/text Facebook https://disinfocode.eu/reports/facebook/5/text Messenger https://disinfocode.eu/reports/messenger/5/text Instagram https://disinfocode.eu/reports/instagram/5/text Whatsapp https://disinfocode.eu/reports/whatsapp/5/text	marzo/septiembre marzo marzo marzo marzo
INFORMES MICROSOFT Bing https://disinfocode.eu/reports/microsoft-bing/5/text Linkedin https://disinfocode.eu/reports/microsoft-linkedin/5/text	marzo/septiembre marzo/septiembre
INFORME TIKTOK https://disinfocode.eu/reports/tiktok/5/text	marzo/septiembre
INFORME TWITCH https://disinfocode.eu/reports/twitch/5/text	marzo

Visión general de los informes

Los informes de transparencia presentados por los grandes prestadores han ido ofreciendo más información y más detallada desde la aprobación del Reglamento de Servicios Digitales y la integración del Código de Conducta en Materia de Desinformación en el marco de ese Reglamento, observándose un esfuerzo de concreción a la hora de dar cuenta del cumplimiento de los compromisos asumidos, las medidas adoptadas y los mecanismos empleados para evaluar su eficacia frente a la desinformación.

En términos generales, ello se observa en relación con los indicadores cuantitativos y cualitativos contemplados (Indicadores de Nivel de Servicio, ILS y Elementos Cualitativos de Informe, ECI), así como sobre las tácticas, técnicas y procedimientos empleados (TTP).

Los grandes prestadores mencionan avances relevantes en la lucha contra la desinformación o los intentos de amplificar contenidos de forma engañosa, como el fortalecimiento de los términos y condiciones y el desarrollo de políticas más robustas contra las prácticas manipulativas, es decir, contra intentos de engañar r a los usuarios mediante informaciones erróneas, vídeos alterados, campañas coordinadas y operaciones de influencia.

Dan cuenta de las medidas adoptadas con respecto a aspectos como las normas de la comunidad para identificar ilícitos en los contenidos generados por los usuarios (UGC), incluyendo la difusión de *spam*; las cuentas falsas (*botnets*), el robo de cuentas y suplantación de identidad; la manipulación de contenidos, pero también de los algoritmos y las funciones de búsqueda; la no identificación de fuentes, o los perfilados. Ofreciendo a los usuarios orientación sobre cómo denunciar ese tipo de contenido.

En ocasiones, los informes aportan enlaces a información complementaria sobre las medidas proactivas que adoptan los operadores, por ejemplo, ante el fraude y las estafas, o ante la desinformación en materia de salud, así como la implementación de avisos a los usuarios ante estos contenidos ilícitos.

Mencionan iniciativas e inversiones en tecnologías relacionadas con la IA: por ejemplo, para la detección de perfiles falsos mediante algoritmos avanzados que detectan similitudes de contenido y comportamiento y de *deepfakes*; detección de conductas de riesgo y anómalas; modelos de aprendizaje para detectar secuencias de actividad asociadas con la automatización abusiva.

En materia de alfabetización mediática e informacional, los operadores declaran participar en diferentes proyectos y campañas basados en la investigación y las mejores prácticas de la industria (*The Trust Project*, *Verified*, *News Literacy Project* y otros.). Forma parte de

la alfabetización En el aspecto de la alfabetización mediática e informacional se recurre a paneles informativos junto a los vídeos o en los resultados de búsqueda para ofrecer contexto adicional y se desarrollan campañas de sensibilización que animan a los usuarios a reflexionar antes de compartir contenidos dudosos.

De forma no sistemática se publican datos estadísticos desagregados por tipo de servicio y Estado miembro (incluyendo España), así como para el conjunto de la UE, sobre las mencionadas cuentas falsas, pero también sobre la falsedad de reacciones (*likes*, etc.), seguidores / suscriptores grupos de chat, foros o dominios igualmente falsos. También en relación a las visitas, visitantes, usuarios, páginas visitadas y descargas.

En los informes se hace referencia a la Coalición para la Procedencia y Autenticidad del Contenido (C2PA), una iniciativa internacional orientada a garantizar la trazabilidad y autenticidad de los contenidos digitales. La integración de metadatos permite verificar el origen y las modificaciones de los contenidos, funcionando como una especie de sello de procedencia o etiquetado que facilita distinguir los materiales auténticos de los alterados.

Hay que tener en cuenta, sin embargo, que la información que se recibe de los operadores depende de su decisión de ofrecerla. Al tratarse de un instrumento voluntario, son ellos quienes deciden, como ya se ha comentado, a qué compromisos y medidas se adhieren y sobre qué indicadores dan cuenta y cómo. Debido a esa voluntariedad, propia de los códigos de conducta, el contenido de los informes de transparencia es muy desigual, incluso entre los diferentes servicios o filiales de una misma plataforma, y a pesar de la existencia de la plantilla de referencia.

Es frecuente que los operadores se limiten en sus informes a reproducir en un apartado lo que recoge el Código con respecto al mismo sobre lo que debería hacerse, sin aportar nada, o recogen las tablas de indicadores cuantitativos sobre diferentes indicadores vacías o con valores cero. En ocasiones los informes se limitan a señalar que sigue mejorando sus sistemas internos para detectar comportamientos de desinformación, pero sin detallar cambios importantes o medidas futuras.

Como hemos señalado, se indican datos por países, por ejemplo, sobre la retirada de contenidos, pero en muchas ocasiones no se ofrece información sobre el impacto real de las acciones. Por ejemplo, visualizaciones o comentarios tenían esos vídeos antes y después de ser retirados, lo que ayudaría a entender si las medidas son efectivas.

Los propios informes señalan a veces que los datos aportados deben tomarse como aproximados, ya que pueden incluir errores o duplicaciones en los recuentos.

Cuando se mencionan, las medidas adoptadas frente a la propaganda política son muy destacadas por los operadores, especialmente en relación con los procesos electorales. También las referidas al ámbito de las comunicaciones comerciales. En relación con esta últimas, algunos informes destacan que el control sobre la ubicación de los anuncios permite a los anunciantes excluir determinadas categorías de contenido, como política, noticias, deportes, belleza o moda, para evitar que su publicidad aparezca junto a materiales potencialmente polémicos o poco fiables.

No obstante, si bien las marcas pueden recurrir a decisiones de bloqueo o suspensión para evitar financiar indirectamente sitios o contenidos desinformativos, reduciendo los incentivos económicos que sostienen estas prácticas, no siempre se aclara cuáles son los criterios empleados ni el proceso resulta claro y accesible.

Las informaciones que se ofrecen sobre la colaboración con agentes externos que monitorizan la calidad y seguridad de los contenidos son más concretas en el caso de los verificadores y, en menor medida, de los expertos. Pero son muy genéricas a la hora de referirse a la participación de las organizaciones sociales en el seguimiento y supervisión del cumplimiento del código y del Reglamento.

Por último, cabe señalar que, en relación con la cooperación entre plataformas más allá del Grupo de Trabajo, se mencionan pocas iniciativas actuales o futuras.

CONCLUSIONES FINALES

El marco regulatorio de la actividad de los operadores de internet, especialmente de las grandes plataformas y buscadores, tiene en el Reglamento de servicios Digitales y en el Código de Conducta en Materia de Desinformación herramientas de gran utilidad para garantizar las buenas prácticas y el empoderamiento social.

La integración del Código de Conducta en el marco de la DSA lo convierte en un referente relevante para determinar el cumplimiento de este Reglamento en relación con los riesgos de desinformación por parte de las VLOP y VLOSE que se adhieran al mismo.

La evolución de los informes de rendición de cuentas pone de relieve importantes avances en la adopción de compromisos y medidas contra la desinformación, en materias como la actuación contra las cuentas falsas, la publicidad engañosa, la propaganda política ilícita, la notificación y bloqueo de contenidos o la creación de cauces de notificación por los usuarios. Es más, se van conociendo decisiones de algunas grandes plataformas firmantes del Código en el sentido de ir reduciendo el papel de los verificadores de datos externos, e incluso de los moderadores de contenidos propios, en favor de las notas de comunidad, que hacen pivotar la responsabilidad última de la desinformación en los usuarios finales.

Por otra parte, buena parte de los informes de transparencia presentados, registran logros e iniciativas importantes, aunque en algunos casos, se podría echar en falta algo más de concreción. Los datos ofrecidos son, en ocasiones, demasiado generales, por lo que sería recomendable un poco más de claridad sobre cómo se aplican las medidas.

Es importante que la información sobre las medidas específicas adoptadas en aspectos como la protección de los menores, la salvaguarda de los procesos electorales y la propaganda política, la publicidad comercial, y, en esos ámbitos, las acciones *ex ante* y *ex post* contra, la difusión de desinformación, información errónea, noticias falsas o bulos y operaciones de influencia, sin duda de gran utilidad, sea lo más transparente posible.

En todo caso, no se puede olvidar que el Código es un instrumento de carácter voluntario al que se suscriben las plataformas y que puede coadyuvar al cumplimiento de la DSA pero sin sustituirla. De hecho, tal y como se indica en las conclusiones de la Junta Europea de Servicios Digitales sobre el Código de Conducta, la DSA prevalece siempre respecto al Código. En este sentido, la suscripción a un código de conducta por parte de una VLOP o VLOSE, cuando proceda, puede ser considerada como una medida de mitigación de riesgos sistémicos. En cualquier caso, corresponde a la Comisión Europea analizar si las

VLOP y VLOSE están cumpliendo con sus obligaciones para evitar la difusión de contenido ilícito (incluida la desinformación) y, en su caso, imponer la sanción que corresponda.

En relación con España, esperamos que el Parlamento apruebe, a la mayor brevedad posible, el texto normativo que habilite a la CNMC para el desarrollo de sus funciones como coordinador de servicios digitales, lo que a su vez permitirá, entre otras medidas, poder certificar a aquellos alertadores fiables y órganos de resolución extrajudicial de conflictos que cumplan los requisitos, y contribuir a garantizar un entorno en línea seguro, predecible y digno de confianza, objetivo principal de la DSA.