

Cerrando la brecha de la credibilidad: cómo han evolucionado la desinformación y la confianza digital

‘Cerrando la brecha de la credibilidad: cómo han evolucionado la desinformación y la confianza digital’ ha sido elaborado en el contexto de FERMI (Fake nEws Risk MItigator) [Proyecto 101073980], un proyecto de Horizonte Europa que estudia e intenta contrarrestar las causas de fondo, la propagación y las implicaciones de la desinformación y las noticias falsas.

Este material formativo está inspirado en las ideas compartidas durante el webinar final de FERMI, “Cómo evoluciona la desinformación: narrativas, influencia digital y confianza”, organizado por CONVERGENCE el 10/09/2025. El webinar contó con Katerina Beli (investigadora asociada y gestora de comunicación en The Lisbon Council) y la Dra. Sofía Tipaldou (profesora asistente de Relaciones Internacionales en la Universidad Panteion de Ciencias Sociales y Políticas) como ponentes; sus contribuciones se reconocen y se citan a lo largo del documento.

El objetivo de este material es analizar la dinámica evolutiva de la desinformación y explorar cómo se construyen, cuestionen y reconfiguran la credibilidad, la confianza y la legitimidad en la era digital. Analizando el ciclo de vida de la desinformación y el marco tecnológico de FERMI, se destaca cómo un enfoque interdisciplinario, que combina inteligencia artificial, ciencias del comportamiento y análisis socioeconómico, puede tender puentes en la brecha de credibilidad y fomentar la resiliencia en las sociedades europeas.

Este material complementa los materiales de FERMI “Navigating Disinformation: A Comprehensive Guide” y “Digital Trust: A Practical Path to Combating Disinformation and Fostering Resilience”, ampliando el debate hacia la interacción entre la manipulación narrativa y la confianza pública.

“La desinformación no solo distorsiona hechos; moviliza a la gente.”
(Webinar FERMI, 2025)

PARTNERS



SECCIÓN 1: COMPRENDIENDO LA DESINFORMACIÓN – CONCEPTOS Y CONTEXTO

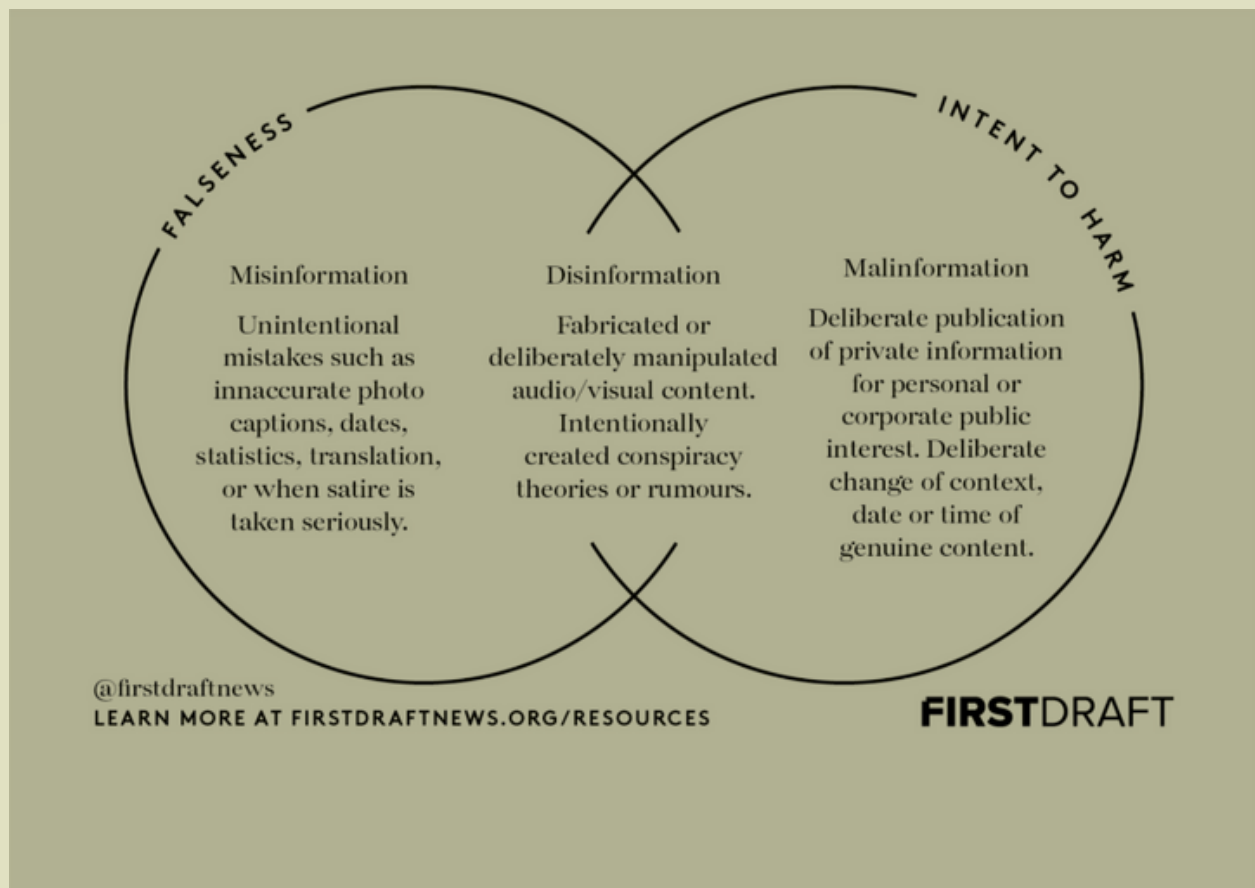
Como subrayó la Dra. Sofía Tipaldou en el webinar, la desinformación no es simplemente una publicación falsa aislada, sino un proceso estratégico: se crea deliberadamente para engañar, manipular o polarizar, a menudo explotando vulnerabilidades emocionales e ideológicas. Para comprender su impacto, debemos considerar el concepto más amplio de desorden informativo, es decir, cómo la información se crea, se comparte y se recibe en la era digital, donde información errónea, desinformación e información maliciosa interactúan y evolucionan.

¿Por qué el término “desorden informativo” en lugar de “noticias falsas”?

El término “fake news” (“noticias falsas”) resulta engañoso y se ha politizado cada vez más. Gran parte del contenido problemático no son noticias en absoluto (por ejemplo, rumores, memes, imágenes manipuladas, anuncios dirigidos); a veces un contenido mezcla elementos genuinos con descontextualización; además, la etiqueta “fake news” suele usarse para desacreditar al periodismo legítimo. Desorden informativo supone un enfoque más preciso y exhaustivo para describir el complejo entorno informativo actual.

Terminología clave del webinar:

- **Información errónea (misinformation):** información falsa o inexacta compartida sin intención de engañar.
- **Desinformación (disinformation):** información falsa o engañosa compartida con la intención de engañar o manipular.
- **Información maliciosa (malinformation):** información verdadera utilizada de manera que cause daño (por ejemplo, difundiendo datos reales en un contexto malicioso destinado a perjudicar).



Tipos de desinformación y desinformación

Según la taxonomía de Claire Wardle (First Draft, 2017–2019), la información falsa o engañosa abarca un espectro basado en la intención de engañar y el potencial de causar daño. Este marco distingue entre diferentes formas, incluyendo

- sátira o parodia,
- conexión falsa,
- contenido engañoso,
- contexto falso,
- contenido impostor,
- contenido manipulado y contenido fabricado.

Reconocer estas categorías ayuda tanto a analistas como a ciudadanos a identificar técnicas de manipulación, evaluar la intención detrás del contenido y responder de manera adecuada dentro de un ecosistema informativo en constante evolución.

¿Por qué es importante todo esto?

La desinformación opera como un sistema, combinando encuadres narrativos, manipulación de medios y amplificación selectiva para lograr objetivos políticos, económicos o ideológicos. Por esta razón, debemos estudiarla como un ecosistema, no como publicaciones aisladas.

Esta perspectiva conduce directamente a la siguiente sección del webinar, donde Tipaldou presentó el “ciclo de vida de la desinformación”, un modelo procesual que muestra cómo evolucionan y persisten tales narrativas.

HIGH



FABRICATED CONTENT

New content that is 100% false, designed to deceive and do harm.



MANIPULATED CONTENT

When genuine information or imagery is manipulated to deceive.



IMPOSTER CONTENT

When genuine sources are impersonated.



FALSE CONTEXT

When genuine content is shared with false contextual information.



MISLEADING CONTENT

Misleading use of information to frame an issue or individual.



FALSE CONNECTION

When headlines, visuals or captions don't support the content.



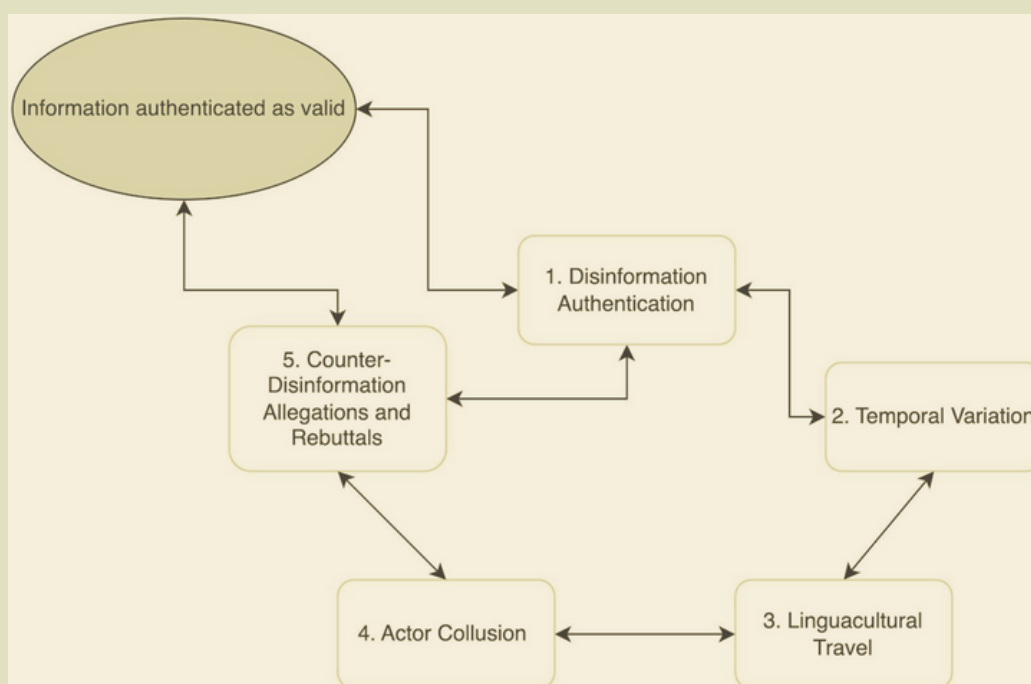
SATIRE OR PARODY

No intention to cause harm but has potential to fool.

LOW

SECCIÓN 2: EL CICLO DE VIDA DE LA DESINFORMACIÓN – CINCO ETAPAS INTERRELACIONADAS

Tal y como lo presentó la Dra. Sofía Tipaldou, el “ciclo de vida de la desinformación” (Tolz, Hutchings, Kazakov, Tipaldou, 2025) ofrece un modelo orientado a procesos que captura cómo la desinformación evoluciona a través de un ciclo constante de validación, transformación y disputa dentro del ecosistema informativo. Este ciclo consta de cinco etapas interrelacionadas que describen cómo las narrativas falsas o engañosas se desarrollan y persisten:



- **Autenticación de la desinformación** – Procesos y prácticas mediante los cuales los actores presentan, etiquetan o “acreditan” ciertos contenidos como creíbles, preparándolos para una mayor difusión.
- **Variación temporal** – La misma narrativa cambia de forma con el tiempo (se reinterpreta, se ajusta en el tiempo o se actualiza selectivamente) para adaptarse a nuevos acontecimientos y mantener su vigencia.
- **Viaje lingüístico-cultural** – Las narrativas se trasladan a través de idiomas y culturas, incorporando referencias, metáforas y marcos que resuenan en contextos locales.
- **Coordinación entre actores** – Coordinación (formal o tácita) entre medios, plataformas, influencers o comunidades para aumentar la visibilidad y reforzar la legitimidad percibida de la narrativa.
- **Acusaciones y refutaciones** – Ciclos de acusación y respuesta que, paradójicamente, pueden seguir difundiendo y afianzando la narrativa original bajo nuevas formas (por ejemplo, acusando a otros de “desinformación” e insistiendo en el mensaje inicial).

El ciclo es iterativo: las acusaciones, refutaciones y nuevas “autenticaciones” retroalimentan el sistema, a veces logrando que un contenido se acepte ampliamente como válido a pesar de sus orígenes manipuladores.

SECCIÓN 3: LECCIONES DE UN ESTUDIO DE CASO – LA ESTRATEGIA DEL KREMLIN Y LAS ELECCIONES EUROPEAS DE 2024

La investigación empírica de la Dra. Sofía Tipaldou, presentada durante el webinar final de FERMI, examinó las estrategias digitales coordinadas empleadas por medios vinculados a Rusia en el período previo a las elecciones europeas de 2024. Este estudio de caso demuestra cómo opera en la práctica el ciclo de vida de la desinformación, revelando un sistema de influencia en red que se apoya en la repetición, la referencia cruzada y la adaptación lingüística para sostener la credibilidad de narrativas engañosas a lo largo del tiempo.

Medios como Voice of Europe (VoE), RRN, Pravda y France et EU actuaron como nodos centrales dentro de una constelación más amplia de actores mediáticos. Estas plataformas combinaron contenido auténtico, semi-auténtico y fabricado, borrando la línea entre periodismo, opinión y propaganda para influir en cómo se perciben la política, la identidad y la moral en Europa.

Entre las características clave identificadas se incluyen:

- Referencias cruzadas a medios de comunicación occidentales para generar una legitimidad percibida, a menudo sin citar las fuentes de forma transparente.
- Citas falsas de medios tradicionales, reemplazando referencias auténticas con fuentes vinculadas al Estado ruso, tales como TASS, RT o RIA Novosti.
- Hipervínculos extensos entre sitios web de extrema derecha, extrema izquierda y teorías conspirativas, tejiendo una red engañosa de validación mutua.
- Enmarque temático centrado en la soberanía, el nacionalismo y la “decadencia moral”, diseñado para apelar emocionalmente a distintos públicos ideológicos.

El análisis de enlaces entrantes (backlinks) revela que estas narrativas fueron amplificadas tanto por ecosistemas mediáticos extremistas como por medios “alternativos”, creando cámaras de eco donde comunidades polarizadas reforzaban inadvertidamente la misma desinformación. Estos hallazgos ilustran cómo la desinformación explota la lógica de las redes: depende más de la visibilidad, la repetición y la colaboración entre extremos ideológicos que de una coherencia ideológica tradicional.

SECCIÓN 4: EL ENFOQUE DE FERMI – ANÁLISIS HOLÍSTICO Y HERRAMIENTAS BASADAS EN IA

Presentado por Katerina Beli, el proyecto FERMI responde a esta complejidad con un enfoque holístico y multidisciplinario que integra inteligencia artificial, análisis de comportamiento y datos socioeconómicos.

Objetivos:

- Apoyar a las autoridades policiales europeas (Law Enforcement Authorities, LEAs) y a otros actores relevantes en la detección, análisis y mitigación de amenazas de desinformación.
- Predecir y mapear las potenciales consecuencias offline de la desinformación, incluidas manifestaciones extremistas, delitos de odio y erosión de la confianza pública.
- Desarrollar y difundir materiales formativos para fortalecer la confianza digital y la resiliencia en toda Europa.

Módulos y entregables clave:

- Analizador de fuentes y difusión de la desinformación – identifica el origen, las vías de propagación y el escalado de las campañas de desinformación.
- Analizador de sentimientos – rastrea los tonos emocionales y patrones de interacción del público en torno a ciertos contenidos.
- Modelador de flujos dinámicos – predice cómo las narrativas online pueden traducirse en riesgos en el mundo real, anticipando posibles escaladas.
- Perfilador de comportamiento y analizador socioeconómico – evalúa las vulnerabilidades dentro de comunidades específicas que podrían hacerlas más susceptibles a la desinformación.
- Modelo de resiliencia comunitaria – mide la capacidad de recuperación de comunidades y su potencial para reconstruir la confianza tras incidentes de desinformación.
- Módulo de aprendizaje de enjambre (Swarm Learning Module) – un sistema de IA descentralizado que permite la colaboración entre las autoridades policiales europeas, preservando la privacidad de los datos.

Juntos, estos módulos permiten una respuesta multidimensional que aborda los aspectos digitales y sociales de la desinformación.

Los disturbios de Dublín (23 de noviembre de 2023)

Durante el webinar también se enfatizó que las consecuencias de la desinformación trascienden el ámbito digital. Los disturbios ocurridos en Dublín, Irlanda, fueron un ejemplo contundente de cómo las narrativas falsas online pueden desencadenar en estallidos de violencia en el mundo real, ilustrando los riesgos tangibles para la estabilidad democrática y la seguridad pública. El sistema de FERMI responde a este “nexo con la seguridad pública” ayudando a las autoridades a identificar señales tempranas de escalada, evaluar la vulnerabilidad de la comunidad y fortalecer la resiliencia social mediante prevención basada en datos y respuestas coordinadas.

EL INCIDENTE:

El 23 de noviembre de 2023, un hombre atacó a tres niños y un adulto frente a una escuela en Dublín, Irlanda. El hecho fue trágico, pero acabó siendo catastrófico cuando comenzó a difundirse información falsa sobre los antecedentes del atacante.

LA NARRATIVA FALSA:

En cuestión de horas, circularon en redes sociales y foros en línea afirmaciones falsas que aseguraban que el atacante era un migrante o refugiado. Estas aseveraciones no se basaban en ninguna información oficial, pero se presentaron como hechos, a menudo con un marco provocador del tipo: “otro ejemplo de la fallida política de inmigración en Irlanda” o “¿qué esperabas cuando dejas entrar a esta gente?”.

LA DIFUSIÓN:

La narrativa falsa se propagó rápidamente a través de:

- Cuentas en redes sociales de orientación derechista y canales de Telegram asociados.
- Foros y tableros de discusión en línea.
- Campañas de difusión coordinadas que aparentaron ser consensuadas.
- Aprovechamiento de las tensiones existentes en torno a la política migratoria irlandesa.

Los disturbios de Dublín (23 de noviembre de 2023)

EL IMPACTO EN EL MUNDO REAL:

Esa misma tarde, estallaron disturbios masivos en Dublín, concentrados en el centro de la ciudad. La violencia incluyó:

- Destrucción de propiedad pública.
- Quema de vehículos, incluidos autobuses.
- Ataques a comercios.
- Enfrentamientos con la policía.
- Heridos tanto entre civiles como entre agentes del orden.

El Ayuntamiento de Dublín estimó posteriormente que los daños materiales ascendieron a aproximadamente 20 millones de euros.

¿Qué nos enseña este caso?

- **Velocidad de la mentira vs. velocidad de la verdad:** La información falsa se difundió en cuestión de horas, mientras que las correcciones oficiales tardaron mucho más y alcanzaron a un público menor.
- **Aprovechamiento de tensiones existentes:** La desinformación no creó el sentimiento anti-inmigración; más bien utilizó como arma las ansiedades y divisiones culturales preexistentes para incitar el conflicto.
- **Consecuencias reales:** Esto demuestra que la desinformación no es meramente un fenómeno online. Puede conducir directamente a la violencia, daños económicos y graves perjuicios sociales en el mundo real.
- **Dificultad de corrección:** Incluso después de ser desmentida la narrativa falsa, muchos de los participantes o simpatizantes de los disturbios continuaron creyendo en ella o justificaron sus acciones basándose en la mentira original.

SECCIÓN 5: CREDIBILIDAD, LEGITIMIDAD Y BRECHA DE CONFIANZA

La brecha de credibilidad (según Aikaterini Beli) se define como la creciente divergencia entre lo que la ciudadanía percibe como digno de confianza y lo que las instituciones (gobierno, medios, ciencia, etc.) reconocen como legítimo.

¿Por qué es relevante?

La desinformación se aprovecha de esta brecha al:

- dar prioridad a la emoción sobre la evidencia, ya las señales de identidad por encima del conocimiento experto;
- preferir la simplificación frente a la complejidad, ofreciendo culpables claros y respuestas simplistas;
- anteponer la rapidez y el impacto visual por encima de la verificación;
- utilizar contenido generado o manipulado por IA para imitar la autenticidad a gran escala.

Implicaciones

A medida que esta brecha se ensancha, la confianza se fragmenta, los mensajes institucionales enfrentan un mayor escepticismo, y las narrativas engañosas encuentran menos barreras para su aceptación y recirculación. Esta dinámica se sustenta en la interacción narrativa: cada afirmación falsa provoca reacciones y contra-afirmaciones que, a su vez, amplifican la visibilidad de la narrativa original. En este sentido, la desinformación no es un mensaje unidireccional, sino un diálogo continuo, un bucle de retroalimentación que difumina los límites de la verdad y desestabiliza la confianza pública.

La metodología de FERMI proporciona herramientas para trazar estas cadenas de credibilidad y legitimidad, rastreando cómo el aprovechamiento de las emociones, la política identitaria y los mecanismos de amplificación digital configuran la percepción de qué es confiable.

“La credibilidad no se decreta; se gana con coherencia y transparencia.”

(Webinar FERMI, 2025)

SECCIÓN 6: TU KIT DE HERRAMIENTAS PERSONAL DE RESPUESTA

Como destacó Katerina Beli durante el webinar final de FERMI, combatir la desinformación requiere no solo salvaguardas institucionales, sino también conciencia personal.

La confianza digital, explicó, comienza con las decisiones cotidianas de los usuarios individuales: cómo consumimos, interpretamos y compartimos información. Inspirada en sus comentarios sobre la importancia del pensamiento crítico, el uso responsable de la tecnología y el papel activo de la ciudadanía en reforzar la resiliencia, esta sección convierte esas ideas en una guía práctica.

Si bien las medidas estructurales y las herramientas de detección impulsadas por IA son una parte de la solución, las acciones individuales en conjunto dan forma al ecosistema informativo. Al tomarnos un momento antes de reaccionar, cuestionar el contenido emocional y verificar la credibilidad, cada uno de nosotros puede contribuir a un espacio digital más confiable.

A continuación, ofrecemos sugerencias simples y accionables para poner en práctica la alfabetización digital en el día a día.



El método STOP: Antes de compartir, detente

Cuando te encuentres con información online –especialmente si es sorprendente, indignante o apela fuertemente a tus emociones– utiliza el método STOP antes de reaccionar o compartirla:



S - STOP - (Detente):

Pausa antes de reaccionar emocionalmente o compartir. Nuestro instinto inmediato ante un contenido impactante suele ser difundirlo de inmediato. Resiste ese impulso inicial.

T - THINK - (Piensa):

Hazte preguntas críticas:

- ¿Quién creó este contenido?
- ¿Por qué fue creado?
- ¿Cuál es la fuente original?
- ¿Están intentando manipularme emocionalmente?
- ¿Confirma esto lo que ya creo (podría ser sesgo de confirmación)?

O - OBSERVE - (Observa):

Busca señales de alerta de posible contenido engañoso:

- Lenguaje emocional o sensacionalista extremo.
- Ausencia de fuentes confiables o referencias verificables.
- URLs sospechosas o sitios web poco conocidos.
- Faltas ortográficas o gramaticales inusuales (aunque a veces son deliberadas para evadir detección).
- Peticiones de acción o de compartir de forma inmediata.

P - PROBE - (Investiga):

Verifica antes de dar algo por cierto o compartirlo:

- Consulta sitios web de verificación de datos confiables.
- Realiza búsquedas inversas de imágenes para ver su origen.
- Comprueba si medios de reconocida reputación están reportando la misma información.
- Busca la fuente original de la afirmación o noticia.

Esta breve comprobación ayuda a frenar el “pensamiento rápido”, esas reacciones automáticas que la desinformación busca explotar.

Ejercicio práctico: Pon a prueba tus habilidades y haz de la pausa un hábito. Realiza la breve actividad interactiva de la Oficina del Comisionado de Derechos Humanos de Columbia Británica (BC Human Rights Commission) para practicar el método STOP con ejemplos reales, obtener comentarios al instante y reforzar estos hábitos de verificación en situaciones reales:

<https://bchumanrights.ca/resources/awareness-campaigns/stop/>.

SECCIÓN 7: PROPUESTAS DE ACTUACIÓN Y REPERCURSIONES SOCIALES

Tanto Katerina Beli como Sofía Tipaldou subrayaron la importancia de equilibrar las contramedidas eficaces contra la desinformación con la protección de los derechos fundamentales. Las respuestas en materia de políticas deben ser proporcionadas, transparentes y basadas en principios democráticos, evitando caer en excesos o en censura arbitraria. A continuación, se presentan algunas ideas clave en materia de políticas públicas expuestas en el webinar:

- **No exagerar la amenaza:** Evitar sobredimensionar el alcance percibido de las campañas de desinformación, ya que hacerlo podría, involuntariamente, amplificarlas aún más.
- **Precisión terminológica:** Emplear una terminología consistente, diferenciando con claridad entre desinformación, información errónea y comunicación política legítima.
- **IA + análisis humano:** Complementar las herramientas de detección automatizada con análisis cualitativo y contextual para prevenir interpretaciones erróneas o falsos positivos/negativos.
- **Fomentar la alfabetización mediática:** Promover la alfabetización mediática y digital, así como el pensamiento crítico, como habilidades básicas para todos los ciudadanos desde la educación básica.
- **Colaboración intersectorial:** Impulsar la colaboración intersectorial entre gobiernos, empresas tecnológicas, investigadores y sociedad civil para reforzar la confianza digital de manera conjunta.

SECCIÓN 8: CONCLUSIONES

El proyecto FERMI replantea la desinformación como un ecosistema en evolución, no simplemente como una amenaza estática, y ofrece un camino práctico para la construcción de resiliencia. Mediante la combinación de innovación tecnológica, gobernanza ética y empoderamiento ciudadano, FERMI fortalece la capacidad europea para navegar la complejidad de los flujos informativos digitales.

FERMI aporta:

- **Una comprensión dinámica de la desinformación como un proceso iterativo y arraigado en lo social.**
- **Herramientas analíticas que conectan la evolución de las narrativas con riesgos sociales medibles.**
- **Un enfoque centrado en la confianza, que acerca a las instituciones y la ciudadanía para cerrar las brechas de la credibilidad.**
- **Un modelo de colaboración que alinea la IA, la ética y la responsabilidad democrática en la lucha contra la desinformación.**

Puntos clave:

- La desinformación prospera al aprovechar de los puntos débiles de la sociedad, los detonantes emocionales, la ambigüedad informativa, las y brechas en la credibilidad institucional. Luego se reproduce a través de un ciclo de cinco etapas (autenticación, variación temporal, desplazamiento lingüístico-cultural, coordinación entre actores y contranarrativas/refutaciones) que mantiene las narrativas dañinas en circulación.
- Cerrar la brecha de credibilidad —es decir, la distancia entre lo que la gente percibe como confiable y lo que las instituciones consideran legítimo— es central para la resiliencia. El sistema de FERMI aborda este desafío mediante el análisis con IA de fuentes, difusión, sentimiento y riesgo, junto con modelos de resiliencia comunitaria, permitiendo anticipar cambios narrativos en lugar de solo reaccionar a ellos.
- Las contramedidas deben ser efectivas pero proporcionadas, respetuosas con los derechos fundamentales, garantizando la libertad de expresión al tiempo que promueven la transparencia y la rendición de cuentas.
- Finalmente, la confianza digital es una responsabilidad compartida que exige un esfuerzo coordinado entre las políticas públicas, la tecnología, los medios de comunicación, la sociedad civil, y la ciudadanía en general.

Lecturas adicionales:

- [Navigating Disinformation: A Comprehensive Guide](#)
- [Digital Trust: A Practical Path to Combating Disinformation and Fostering Resilience](#)
- [Disinformation as process: modeling the lifecycle of deceit](#)
- [Information disorder: Toward an interdisciplinary framework for research and policy making.\(2017\)](#)
- [Misinformation Has Created a New World Disorder](#)
- [Misinformation, Disinformation, and Bias: Identifying misinformation, its various types, and both personal and news media biases](#)
- [Misinformation: Can you STOP it?](#)

Y para explorar más materiales y recursos relacionados, visita:

- <https://fighting-fake-news.eu/materials/training-materials>