Habilidades Prácticas en el Ciberespacio - CEM 2025

Estudiante: MYCIM Yerson Alejandro Torres Bueno

1. ¿Cuál es la diferencia fundamental, según el texto, entre "misinformation" y "disinformation"?

La diferencia principal entre *misinformation* y *disinformation* radica en la intencionalidad. *Misinformation* se refiere a información incorrecta que se difunde sin intención de causar daño; el emisor cree que es veraz. En cambio, *disinformation* implica una acción deliberada de manipulación, donde el emisor sabe que la información es falsa y la utiliza con fines políticos, sociales o económicos (Gómez Galán et al., 2023, p. 15). Por ejemplo, compartir un remedio falso contra la COVID-19 creyendo en su efectividad sería *misinformation*, mientras que hacerlo para manipular audiencias o provocar daño social constituye *disinformation*. Esta distinción es fundamental en ciberdefensa, ya que implica diferentes estrategias de mitigación frente a amenazas informativas.

Referencia: Gómez Galán, J., Alzate, C., & Martínez, L. (2023). *Luchando contra la desinformación mediante la inteligencia artificial*. Universidad Nebrija.

2. Según el Reuters Institute Digital News Report 2023, ¿qué tendencia preocupante se observa en España con respecto al interés por las noticias?

El *Reuters Institute Digital News Report 2023* revela una tendencia decreciente en el interés por las noticias en España: del 85% en 2015 al 51% en 2023, una reducción de 34 puntos porcentuales. Este descenso va acompañado de una desconfianza creciente hacia los medios de comunicación, especialmente entre jóvenes menores de 45 años, donde la desconfianza alcanza el 40% (Gómez Galán et al., 2023, pp. 14–15). Esta pérdida de confianza socava la capacidad de las sociedades para discernir información confiable, un desafío crítico en ciberdefensa ante campañas de desinformación masiva.

Referencia: Gómez Galán et al., 2023.

3. ¿Cómo se comparan, según los experimentos de Vosoughi, Roy y Aral (2018), la velocidad y facilidad de difusión de noticias falsas frente a las verdaderas?

Según el estudio de Vosoughi, Roy y Aral (2018), las noticias falsas se propagan de forma significativamente más rápida y amplia que las verdaderas. El 1% de las fake news más virales alcanzó entre 1.000 y 100.000 personas, mientras que el 1% de las verdaderas rara

vez superó las 1.000. La explicación no radica en bots, sino en el comportamiento humano: las noticias falsas suelen ser más novedosas y emocionales, lo que incrementa su viralidad (Gómez Galán et al., 2023, p. 21). Esto plantea un riesgo real para la seguridad informacional de los Estados.

Referencia complementaria: Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

4. ¿Qué ventaja clave ofrecen las redes latentes de difusión sobre los modelos epidemiológicos para el estudio de la desinformación?

Las redes latentes de difusión superan a los modelos epidemiológicos tradicionales al permitir la identificación precisa de los nodos clave en la propagación de desinformación. Mientras los modelos epidemiológicos representan la diseminación como un fenómeno de contagio poblacional, las redes latentes mapean las interacciones específicas entre usuarios, detectando "superpropagadores" e influencias ocultas (Gómez Galán et al., 2023, p. 24). Este enfoque es vital para aplicar contrainteligencia digital y ejecutar operaciones defensivas eficaces en entornos cibernéticos.

Referencia: Gómez Galán et al., 2023.

5. ¿Qué son los "grandes modelos de lenguaje" y cuál es su principal riesgo en el contexto de la desinformación?

Los grandes modelos de lenguaje (LLMs, por sus siglas en inglés), como GPT-3 o GPT-4, son arquitecturas de inteligencia artificial entrenadas con volúmenes masivos de datos para generar texto coherente, responder preguntas y producir contenido contextualizado. Si bien son útiles para tareas legítimas, su principal riesgo es que pueden automatizar la creación de narrativas falsas con una apariencia de veracidad casi indistinguible para el lector promedio, facilitando campañas de desinformación masiva (Gómez Galán et al., 2023, pp. 28–29). Esto se agrava con la proliferación de herramientas de acceso libre y la falta de regulación global.

Referencia: Gómez Galán et al., 2023.
OpenAI. (2023). *GPT-4 Technical Report*. https://openai.com/research/gpt-4

6. ¿Cómo facilita la accesibilidad de los modelos de IA la generación de desinformación?

La disponibilidad pública de modelos de IA generativa y herramientas de código abierto (como ChatGPT, Stable Diffusion o Midjourney) permite que individuos sin conocimientos

técnicos puedan producir contenido falso creíble. Esta facilidad reduce la barrera de entrada para ciberdelincuentes, propagandistas o actores estatales hostiles que buscan manipular la opinión pública, especialmente durante contextos electorales o conflictos híbridos (Gómez Galán et al., 2023, p. 27; Brundage et al., 2018).

Referencia complementaria: Brundage, M., et al. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. https://arxiv.org/abs/1802.07228

7. ¿Qué son las "cajas negras" en el contexto de la IA explicativa y cuál es el desafío asociado?

Una "caja negra" en IA se refiere a modelos cuyos procesos internos son opacos o incomprensibles para humanos, incluso para sus propios desarrolladores. Aunque son altamente precisos, su falta de explicabilidad genera desafíos éticos, legales y operacionales, sobre todo cuando se aplican en contextos como la detección de desinformación, donde es necesario justificar decisiones automatizadas (Gómez Galán et al., 2023, p. 48; Ribeiro et al., 2016).

Referencia complementaria: Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD*, 1135–1144.

8. ¿Qué implicaciones tiene el concepto de "Inteligencia Artificial General (AGI)" para la lucha contra la desinformación?

La AGI representa un paradigma de IA con capacidades cognitivas equiparables a las humanas. Si bien podría revolucionar los sistemas de verificación y ciberdefensa, también podría ser empleada por actores maliciosos para diseñar desinformación más sofisticada, contextualizada y adaptativa (Gómez Galán et al., 2023, p. 49). En escenarios de guerra híbrida, el uso ofensivo de AGI supondría un reto sin precedentes para la soberanía informativa de los Estados.

Referencia: Gómez Galán et al., 2023.

9. ¿Qué normativas europeas importantes se mencionan en relación con la IA y la privacidad?

El texto destaca dos instrumentos clave: el Reglamento General de Protección de Datos (RGPD), que garantiza la protección de datos personales, y el *Libro Blanco sobre la Inteligencia Artificial* de la Comisión Europea, que promueve una IA ética, segura y transparente. Estas normativas proporcionan un marco para evitar abusos en el tratamiento

automatizado de datos e impulsan la explicabilidad y supervisión humana de los sistemas (Gómez Galán et al., 2023, pp. 49–50).

Referencia complementaria: Comisión Europea. (2020). *Libro Blanco sobre la inteligencia artificial*. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52020DC0065

10. ¿Cómo garantiza FacTeR-Check el cumplimiento de la normativa de protección de datos al analizar redes sociales?

FacTeR-Check se adhiere a los principios del RGPD al utilizar exclusivamente datos de usuarios con perfiles públicos o consentimiento informado. Además, se basa en fuentes verificadas y aplica técnicas de explicabilidad en sus modelos, lo que permite auditar sus decisiones. Esta metodología respeta la privacidad, la transparencia y el derecho a la protección de datos personales en el contexto europeo (Gómez Galán et al., 2023, p. 50).