Subí mi propio LinkedIn a NotebookLM.

Hace unas semanas hice un experimento con mi propio perfil profesional. Cogí mi LinkedIn, dos artículos que había publicado y una entrevista pública en un podcast. Lo metí todo en NotebookLM, la herramienta de Google que medio mundo utiliza para estudiar o resumir informes. Lo que me devolvió no fue un resumen.

Fue un manual de instrucciones para manipularme.

No hubo truco, ni filtración, ni dato privado involucrado. Solo información que yo mismo había decidido publicar a lo largo de los años, dispersa en distintos sitios. La IA hizo algo tan sencillo como demoledor: conectarla. Y al conectarla, me entregó un retrato psicológico que cualquier ingeniero social profesional firmaría con los ojos cerrados.

Este artículo cuenta qué hice, qué encontré y por qué creo que esto cambia las reglas de la ciberseguridad doméstica y empresarial para siempre.

Qué es NotebookLM (y por qué todo el mundo lo usa)

NotebookLM es una herramienta gratuita de Google diseñada para trabajar con fuentes propias. Le subes documentos —PDFs, textos, transcripciones, enlaces— y la IA los lee, los conecta entre sí y te permite preguntarle cualquier cosa sobre ellos. Puede resumirlos, generar esquemas mentales, preparar materiales de estudio o incluso producir conversaciones tipo podcast donde dos voces sintéticas discuten el contenido con naturalidad asombrosa.

Es, probablemente, una de las mejores herramientas de productividad que han aparecido en los últimos años. Estudiantes, consultores, abogados, investigadores, médicos, periodistas. Todos la usan. Y con razón.

El problema no es NotebookLM. El problema es que estamos asumiendo que una herramienta diseñada para organizar conocimiento solo sirve para eso. No nos hemos parado a pensar qué pasa cuando el "conocimiento" que alguien le entrega no es un libro técnico, sino las huellas digitales de una persona real.

El experimento: solo con lo que ya es público

Quiero dejar algo claro desde el principio, porque es la clave de todo lo que viene: **en este experimento no se vulneró nada, no se rompió ninguna política de uso y no se accedió a ningún dato privado**. Ese es precisamente el punto incómodo.

Lo que subí a NotebookLM fue esto:

- Mi perfil público de LinkedIn, descargado directamente desde la propia plataforma.
- Dos artículos largos que había publicado en medios abiertos, firmados con mi nombre.
- La transcripción de una entrevista pública en un podcast donde hablaba de mi trayectoria.
- Una selección de comentarios y respuestas mías en publicaciones virales de otros profesionales.

Todo material que cualquier persona con diez minutos y conexión a internet puede recopilar sobre mí. No hay hackeo. No hay filtración. No hay nada que un buscador no me ofrezca ya.

Entonces le hice a la IA tres peticiones concretas. La primera, que redactara un resumen profesional de "quién es esta persona". La segunda, que identificara patrones de decisión, valores personales y posibles temas sensibles a partir del material. La tercera —la más inquietante— que generara una conversación tipo podcast donde dos voces discutieran el perfil como si estuvieran analizando a un desconocido.

Tres minutos. Eso es lo que tardó en devolverme los tres resultados.

Los resultados: no reveló secretos, hizo algo peor

Si esperas que te cuente que la IA desveló información oculta sobre mí, te voy a decepcionar. No hay revelaciones jugosas. No hay secretos. Lo que ocurrió fue mucho más interesante y, sinceramente, mucho más preocupante.

La IA **conectó puntos**.

Supo que valoro la autenticidad en los negocios porque usé esa palabra en cuatro contextos distintos durante dos años. Identificó que tuve una etapa profesional complicada hace unos años porque lo mencioné de pasada en un post motivacional.

Detectó que admiro a una figura pública concreta porque había comentado varias publicaciones suyas con un tono cercano. Observó que mi estilo de comunicación se vuelve más emocional los domingos por la noche —algo que yo mismo no había notado— y lo señaló como un posible patrón vinculado a mi forma de procesar la semana.

Juntó mi ciudad, mi sector, el colegio donde estudié (lo mencioné una vez en un post sobre profesores que me marcaron), los nombres propios de dos mentores, el proyecto que acabo de cerrar, el tipo de clientes con los que trabajo, las posturas que defiendo en debates públicos y las que evito. Construyó una hipótesis sobre mis horarios, mi estilo de toma de decisiones y mis probables puntos ciegos.

Todo eso estaba ahí. Público. Disperso. Aparentemente inofensivo. La IA lo unificó en un perfil psicológico coherente y accionable **en el tiempo que tardas en hacerte un café**.

Ese es el punto que necesita entenderse: la amenaza no es la revelación, es la agregación. Cada publicación tuya, por sí sola, es un grano de arena. La IA construye la playa.

De un resumen simpático a un ataque quirúrgico

Aquí es donde el experimento deja de ser una anécdota curiosa y se convierte en un problema estratégico serio. Porque si yo, como analista, puedo hacer esto con mis propios datos en tres minutos, la pregunta obvia es: ¿qué puede hacer alguien con intenciones distintas con los datos de un directivo financiero, de un responsable de compras, de un padre preocupado por sus hijos o de un adolescente con un perfil sobreexpuesto?

La ingeniería social siempre ha sido el arte de manipular a las personas usando lo que se sabe de ellas. Lo que cambia en 2026 no es la técnica, es la escala y la precisión.

El fraude ya no se parece a aquel correo mal traducido que pedía una herencia nigeriana. Hoy se parece a una llamada en la que alguien con la voz de tu jefe te saluda por el apodo que solo usas en Instagram, te pregunta por el proyecto que cerraste la semana pasada —del que hablaste en LinkedIn el martes— y te pide que hagas una transferencia urgente antes de una reunión que sabe que tienes bloqueada en tu calendario público. Una llamada diseñada para que no tengas tiempo de dudar, con una familiaridad que solo se construye cuando alguien te ha estudiado.

Y para estudiarte ya no hace falta contratar a un investigador. Hace falta una herramienta gratuita y diez minutos.

La realidad del mercado: los datos que ya están sobre la mesa

No hace falta especular. Los casos reales empiezan a acumularse a velocidad incómoda.

En febrero de 2024, la multinacional británica de ingeniería Arup perdió aproximadamente 25 millones de dólares en una estafa en Hong Kong en la que un empleado participó en una videollamada con varios supuestos directivos de la empresa —todos ellos generados por IA a partir de material público— que le instruyeron para realizar transferencias. El empleado ejecutó la orden porque reconoció las caras y las voces. Eran clones.

En España, las denuncias por clonación de voz en estafas familiares —el clásico "mamá, he tenido un problema"— han crecido de forma sostenida en los últimos dos años, con variantes cada vez más sofisticadas que incluyen conocimiento específico del contexto familiar. Información que, en muchos casos, procede de perfiles públicos de redes sociales.

No son anécdotas aisladas. Son la primera ola de un fenómeno que la Agencia de la Unión Europea para la Ciberseguridad (ENISA) lleva alertando desde hace más de un año: el uso combinado de IA generativa y OSINT —inteligencia a partir de fuentes abiertas— para diseñar ataques personalizados a bajo coste.

El dato, para quienes trabajamos en esto, no sorprende. Lo que sorprende es la complacencia con la que seguimos alimentando el modelo.

Qué hacer: cinco acciones ordenadas de menor a mayor compromiso

No te voy a pedir que abandones las redes sociales. Tampoco lo voy a hacer yo. Lo que sí te voy a pedir es que cambies la relación que tienes con lo que publicas. Estas son las cinco acciones que yo recomiendo, en el orden en que las implementaría.

1. Audita lo que es público. Busca tu nombre en Google. Revisa tu LinkedIn como si fueras un desconocido que acaba de recibir el encargo de "investigarte". Qué cuenta de ti sin que tú estés presente para matizarlo. Te sorprenderá lo que encuentras y, sobre todo, lo que se puede deducir de lo que encuentras. Dedica una hora al año a este ejercicio. Solo eso.

2. Reduce la granularidad emocional. Un post que comparte tu vulnerabilidad es valioso humanamente; para un ingeniero social, es un regalo. No se trata de volverse frío ni impostado, sino de encontrar el punto medio: compartir la lección sin regalar el diagnóstico psicológico. Cuéntame que superaste una etapa difícil; no me cuentes exactamente cuándo fue, qué la provocó y cómo te afectó emocionalmente durante meses.

3. Establece protocolos verbales de verificación. Con tu familia y con tu equipo. Una palabra de seguridad acordada para peticiones inusuales, especialmente las que llegan por voz o vídeo y transmiten urgencia. Es la medida más barata y más efectiva contra los deepfakes: una IA no puede saber lo que vosotros decidisteis en privado.

4. Desvincula capas de información. Tu fecha de nacimiento, el nombre de tu colegio, tu primera mascota y el apellido de tu madre son, combinados, las respuestas a la mitad de las preguntas de seguridad de tu banco. No los publiques juntos. No los publiques seguidos. Idealmente, no los uses como respuestas reales.

5. Asume el nuevo paradigma. Si usas IA, perfecto. Yo la uso. Lo que no puedes permitirte es usarla sin tener presente que otras personas, con otras intenciones, pueden hacer lo mismo con tus datos. La ciberseguridad de 2026 ya no es un problema técnico reservado a los departamentos de sistemas. Es una higiene cotidiana, tan básica como cerrar con llave al salir de casa.

La reflexión que deja el experimento

Cada herramienta de productividad que aparece también es, por definición, una herramienta de reconocimiento. Lo que ayuda a un estudiante a preparar un examen ayuda a un atacante a preparar un guion. Lo que resume un informe para un consultor resume una vida para un estafador. No es que la herramienta sea mala; es que el uso depende de quién la tiene en las manos, y en 2026 la tiene cualquiera.

Lo que esto exige no es miedo, es madurez. Entender que tu exposición digital ya no la miden tus amigos, tus seguidores o tus colegas. La mide, cada vez más, una máquina capaz de procesar en tres minutos lo que a un humano le llevaría semanas. Y las máquinas, a diferencia de las personas, no olvidan lo que ya vieron.

No escribo esto para que dejes de publicar. Escribo esto para que publiques con conciencia. Para que cuando decidas compartir algo, lo hagas sabiendo que ese fragmento puede vivir para siempre y conectarse con otros que ni recuerdas haber

publicado. La autenticidad en redes no tiene por qué desaparecer; lo que tiene que llegar es el criterio.

Cierre y próximo paso

Si has llegado hasta aquí, haz una cosa antes de cerrar el artículo: abre tu perfil de LinkedIn, tu Instagram o tu X, y míralos con la pregunta "¿qué podría hacer alguien con todo esto en tres minutos?". Solo eso. No hace falta borrarlo todo de golpe; hace falta empezar a mirar distinto.

¿Te atreves a hacer el experimento con tus propios datos públicos? Cuéntame por privado qué te encontró la IA sobre ti. Lo leo todo. Y si este artículo te ha hecho pensar, compártelo con alguien a quien quieras proteger: una persona informada es una persona mucho más difícil de manipular.

Para más análisis sobre ingeniería social, ciberseguridad aplicada y cultura digital, visita el blog. Hay recursos gratuitos esperándote.

Isaac Ruiz Romero