Proyecto 2 Parte 1

Seguridad Informatica II

Mauricio Josafat Salinas Carrillo
Profesor: Servando Lopez Contreras
6 de marzo de 2026

1. Marco Teórico y Fundamentación Técnica

La ciberseguridad corporativa ha experimentado una transformación paradigmática fundamental. Históricamente, las organizaciones dependían de la Formación en Concientización sobre Seguridad (SAT), un modelo centrado en el cumplimiento normativo que medía el éxito mediante métricas superficiales (tasas de finalización de cursos). Sin embargo, la evidencia demuestra la existencia de una "brecha entre concientización y acción": el conocimiento teórico no evita que los empleados interactúen con enlaces maliciosos.

Para mitigar esta vulnerabilidad, la industria evolucionó hacia la Gestión del Riesgo Humano (HRM). Este marco es continuo e impulsado por datos empíricos. Se basa en modelos de cambio de comportamiento, como el modelo de BJ Fogg, el cual postula que para modificar una acción de ciberseguridad deben converger simultáneamente la motivación, la capacidad y un detonante adecuado.

Enfoques de Entrenamiento Modernos y Diseño Ético

El diseño ético de una campaña de ingeniería social asume que los empleados son una capa de defensa activa (sensores), no una vulnerabilidad que deba ser engañada o castigada. Las metodologías contemporáneas de HRM se rigen por los siguientes principios:

- **Seguridad psicológica y amnistía:** Las plataformas modernas reemplazan el castigo (asignar horas de cursos remediales) con retroalimentación positiva. Ante un fallo, se despliegan módulos de microaprendizaje instantáneo (*just-in-time*) de un minuto. Se fomenta un modelo de "amnistía" donde se prioriza el reporte rápido sin temor a represalias laborales.
- **Límites éticos en señuelos:** Se prohíbe el uso de correos que simulen despidos, resultados médicos o crisis familiares. Las campañas deben proporcionar un "realismo que enseñe" sin destruir la confianza interdepartamental.
- **Ventanas de confianza y medición científica:** Se protegen períodos de alta sensibilidad corporativa (auditorías, reestructuraciones) suavizando las simulaciones. Para estandarizar la dificultad, se utiliza la *NIST Phish Scale*, garantizando que las pruebas sean pedagógicamente justas y estadísticamente medibles evaluando tanto señales de engaño como el contexto del mensaje.

Análisis Exhaustivo de Métricas Telemétricas

La madurez de un programa se define por su transición desde el seguimiento de la susceptibilidad hacia el análisis holístico de la resiliencia:

1. **Tasa de Clics (Susceptibilidad):** Mide interacciones negativas con la amenaza. En programas maduros, esta métrica genera una "meseta de concientización"; una tasa baja puede reflejar inacción (ignorar el correo) más que una defensa activa.
2. **Tasa de Reporte y Resiliencia:** El estándar de oro empírico. Cuantifica a quienes identifican proactivamente la amenaza y alertan al Centro de Operaciones de Seguridad (SOC). El *Resilience Ratio* (Tasa de Resiliencia) se calcula dividiendo los reportes generados entre los clics registrados.
3. **Puntuación de Riesgo (Risk Score):** La culminación analítica basada en la ecuación de riesgo clásica: $Risk Score = Probabilidad \times Impacto$. La ecuación evalúa la probabilidad analizando el historial de fallos y la frecuencia de reportes. El factor de impacto se ajusta asimétricamente según el nivel de privilegios del usuario; un fallo de un administrador de bases de datos genera una penalización algorítmica mayor que el de un empleado de nivel de entrada.

2. Análisis Individual de las 8 Plataformas (Fichas Descriptivas)

2.1. Hoxhunt

- **Arquitectura y Capacidades:** Plataforma nativa en la nube que abandona las campañas masivas en favor de la entrega individualizada. Utiliza aprendizaje automático para adaptar dinámicamente la dificultad, tematización y vector de ataque. Su principal diferenciador técnico son sus conectores SOC nativos bidireccionales, capaces de reducir el ruido de alertas falsas positivas hasta en un 97%.
- **Enfoque Ético:** Su arquitectura se centra en la ciencia conductual y la gamificación estructurada. Los usuarios obtienen puntos e insignias por reportar correos, fomentando el refuerzo positivo. Para perfiles avanzados,

incluye un modo dinámico (*Spicy mode*) que eleva la sofisticación sin requerir reconfiguración manual.

2.2. Proofpoint (Security Awareness Training)

- **Arquitectura y Capacidades:** Integrada directamente a su pasarela de correo seguro (Secure Email Gateway). Inyecta Inteligencia de Amenazas Legible por Máquinas para identificar algorítmicamente a las Personas Muy Atacadas (VAPs). Sus simulaciones poseen un realismo forense extremo basado en campañas activas de *malware* a nivel global.
- **Enfoque Ético:** Su ética se fundamenta en la "relevancia clínica". Al basar los señuelos en amenazas reales que acechan la red, asegura que el usuario se entrene contra vectores estadísticamente probables, mitigando la frustración de enfrentar simulaciones teóricas o irrelevantes.

2.3. KnowBe4

- **Arquitectura y Capacidades:** La plataforma de mayor adopción, albergando la inmensa biblioteca *ModStore*. Su evolución técnica reciente incluye el *SmartRisk Agent* y los *Artificial Intelligence Defense Agents* para un cálculo de riesgo multidimensional. Su módulo PhishER es crítico para la automatización de triage y respuesta a incidentes.
- **Enfoque Ético:** La abrumadora versatilidad de la herramienta delega el filtro ético al administrador. Al disponer de decenas de miles de plantillas estáticas, existe un riesgo documentado de que administradores implementen tácticas de "trampa" que erosionen la moral si no se rigen por políticas de recursos humanos estrictas.

2.4. Cofense (PhishMe)

- **Arquitectura y Capacidades:** Prioriza radicalmente la madurez de las operaciones de seguridad. Su núcleo es el suministro de Inteligencia de Amenazas Legible por Máquinas, validada al 100% por analistas humanos,

garantizando una fidelidad del 99.998% en los indicadores de compromiso. Se integra bidireccionalmente a gran escala con SIEM y SOAR.

- **Enfoque Ético:** Empodera tácticamente al usuario, elevándolo de ser el "eslabón más débil" a un "sensor orgánico activo". Al carecer deliberadamente de gamificación superficial, requiere fuertes campañas de comunicación interna para mantener el entusiasmo a largo plazo.

2.5. Phished (Phished.io)

- **Arquitectura y Capacidades:** Orientada a la eliminación de la carga administrativa mediante IA generativa agresiva. Su innovación técnica principal es el *Zero Incident Mail*, una arquitectura que crea un entorno aislado (confianza cero) donde los usuarios interactúan con enlaces maliciosos de prueba sin riesgo de comprometer la red real.
- **Enfoque Ético:** Se fundamenta en "aprender fallando de manera segura". Al aislar criptográficamente el error en el entorno ZIM, se erradica el miedo sistémico a represalias laborales, cumpliendo rigurosamente con marcos de privacidad de datos (GDPR, ISO 27001).

2.6. NINJIO

- **Arquitectura y Capacidades:** Difiere del mercado al focalizarse en el microaprendizaje episódico de alta producción audiovisual. Despliega animaciones de estilo Hollywoodense narradas por actores, basadas en violaciones de datos documentadas. Sus simulaciones técnicas son básicas, careciendo de canalizaciones profundas para ingesta SIEM o análisis forense.
- **Enfoque Ético:** Es la plataforma con mayor enfoque en el compromiso emocional y la empatía corporativa. Aborda la seguridad desde un ángulo puramente humano, logrando reducir drásticamente la "fatiga de aprendizaje" y la hostilidad hacia el departamento de TI.

2.7. Mimecast (Awareness Training)

- **Arquitectura y Capacidades:** Su *Human Risk Command Center* (HRCC) calcula un *SAFE Score* procesando hasta 2.5 años de latencia histórica (comportamientos, fallos, sentimiento). Su capacidad técnica más disruptiva es el *defanging*: intercepta ataques de *phishing* reales en la pasarela, los "desarma" criptográficamente y los recicla como simulaciones hiperrealistas inmediatas.
- **Enfoque Ético:** Se rige por la intervención proporcional y *just-in-time*. Respeta el ancho de banda cognitivo corporativo entrenando quirúrgicamente solo al personal que exhibe comportamientos de alto riesgo, salvaguardando la productividad de la organización.

2.8. Infosec IQ

- **Arquitectura y Capacidades:** Estructurada bajo el paradigma de un Sistema de Gestión del Aprendizaje clásico. Posee un soporte inamovible para protocolos estándar como SCORM. Muestra rezagos técnicos en la generación dinámica por IA, enfrentando limitaciones documentadas en la latencia de sus microservicios y APIs durante sincronizaciones masivas.
- **Enfoque Ético:** Profundamente arraigada en la certificación institucional y el cumplimiento normativo. Sus defectos operativos pueden derivar en la clasificación accidental de correos válidos como fallos de seguridad, comprometiendo la equidad en la evaluación y generando desmoralización injustificada.

3. Tabla Técnica Comparativa

Criterio / Plataform	Analítica y Métricas Predictivas	Capacidades de Integración y Operación	Diseño Ético y Enfoque Cultural	Limitaciones Estructurales
Hoxhunt	Telemetría en *Time-to-Report* e índice de resiliencia.	Conectores SOC nativos bidireccionales; reducción de alertas	Refuerzo positivo y amnistía total ante el error humano.	Exige esfuerzo organizacional para mantener activa la gamificación a largo plazo.
Proofpoint	Algoritmos VAP (Personas Muy Atacadas)	Integración nativa exclusiva con pasarelas Proofpoint globales.	Relevancia clínica; simula amenazas que el usuario enfrentará	*Lock-in* tecnológico; menor adaptabilidad algorítmica para personalizar la
KnowBe4	*SmartRisk Agent* e IA de defensa (AIDA).	Ecosistema masivo: APIs dinámicas, Azure, PhishER para orquestación.	Transferido al cliente; versatilidad absoluta que permite enfoques	Riesgo de fricción interna si los administradores utilizan plantillas estáticas
Cofense	Analítica empresarial orientada a triage	Integración SOAR/SIEM crítica (Splunk, Cortex XSOAR).	Convierte al usuario en un sensor orgánico de la red	Ausencia deliberada de gamificación; alta curva técnica para
Phished	*Behavioral Risk Score* y adaptación algorítmica sin	Entorno de despliegue *Zero Incident Mail* (ZIM) de confianza cero.	Falla-segura: erradica el miedo a represalias aislando tecnológicamente	Alta dependencia de la IA generativa, requiere escrutinio en sectores altamente
NINJIO	Métricas orientadas a Recursos Humanos (adopción y	Exportaciones estadísticas simples; ideal como servicio gestionado	Empatía extrema basada en narrativa audiovisual emocional.	Arquitectura técnica de simulación básica; inviable para operaciones defensivas

Mimecast	*SAFE Score* interactivo procesando 2.5 años de	Desarme (*defanging*) de cargas reales integradas a la	Intervención quirúrgica *just-in-time*; entrena solo a perfiles	La agresiva extracción histórica puede generar percepciones
Infosec IQ	Analítica LMS orientada al cumplimiento normativo y	Soporte legal estructural, pero con latencia documentada en sincronización	Centrado rígidamente en la instrucción formativa legal clásica.	Carencia de plantillas dinámicas generadas por IA y fallos esporádicos en evaluación.

4. Análisis Crítico y Conclusiones

La progresión de la industria desde herramientas genéricas de concientización hacia ecosistemas complejos de Gestión de Riesgo Humano refleja la necesidad de integrar la ingeniería conductual con la respuesta a incidentes en tiempo real. Este escrutinio profundo demuestra que la elección de una arquitectura debe estar supeditada al nivel de madurez técnica de la organización, sus obligaciones regulatorias y su clima interno.

Segmentación Estratégica y Casos de Uso:

1. **Entornos de Alta Madurez Operativa:** Para corporaciones con equipos activos de caza de amenazas, la ventaja técnica se inclina hacia Cofense o Hoxhunt. La capacidad paramétrica de estas infraestructuras para enlazarse con plataformas SOAR permite correlacionar reportes de simulaciones con movimientos laterales reales, acortando drásticamente el tiempo de permanencia de los atacantes en la red.
2. **Multinacionales y Escrutinio Regulador:** En sectores como la banca o infraestructura crítica, herramientas colosales como KnowBe4 y Proofpoint dominan gracias a su granularidad en el control de acceso y exhaustiva reportería, proporcionando un blindaje legal ante auditorías gubernamentales.
3. **Organizaciones con Alta Fricción Interna:** Si el clima laboral es hostil hacia el departamento de TI debido a metodologías punitivas previas, NINJIO es la opción preeminente para reconstruir la confianza mediante narrativas empáticas. Alternativamente, Phished proporciona una automatización segura a través de su entorno ZIM, ideal para equipos que buscan delegar la operación a la IA sin comprometer la moral del usuario.

Conclusión:

El diseño ético de las campañas de simulación no es un imperativo puramente moral, sino una necesidad de estabilidad operativa. Desplegar un programa agresivo carente de seguridad psicológica y contrapesos erosiona irreversiblemente la confianza del usuario. Este deterioro metodológico crea un riesgo de sabotaje pasivo o inacción (no reportar anomalías por miedo al despido) mucho más destructivo para la continuidad del negocio que las amenazas de *phishing* que el programa intentaba mitigar en primer lugar. La convergencia hacia simulaciones adaptativas e intervenciones proporcionales impulsadas por IA traza el estándar ineludible de la ciberseguridad corporativa moderna.

5. Referencias

- Baker, E. (2023, 7 de febrero). *How security behavior change lets you measure and manage True Risk*. Hoxhunt. https://hoxhunt.com/blog/security-behavior-change-to-measure-true-risk-and-manage-human-risk
- Cybersecurity and Infrastructure Security Agency. (2023). *Phishing Guidance: Stopping the Attack Cycle at Phase One*. Departamento de Seguridad Nacional de los Estados Unidos. https://www.cisa.gov/sites/default/files/2023-10/Phishing%20Guidance%20-%20Stopping%20the%20Attack%20Cycle%20at%20Phase%20One_508c.pdf
- Dawkins, S., & Jacobs, J. (2023). *NIST Phish Scale User Guide (NIST TN 2276)*. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.TN.2276
- Gartner. (2024). *Security Awareness Computer-Based Training Software Reviews*. https://www.gartner.com/reviews/market/security-awareness-computer-based-training
- Infosec Institute. (2026). *Security Awareness Training vs Human Risk Management*. https://www.infosecinstitute.com/resources/security-awareness/security-awareness-training-vs-human-risk-management/
- National Institute of Standards and Technology. (2022). *NIST IR 8420A*. https://nvlpubs.nist.gov/nistpubs/ir/2022/NIST.IR.8420A.pdf