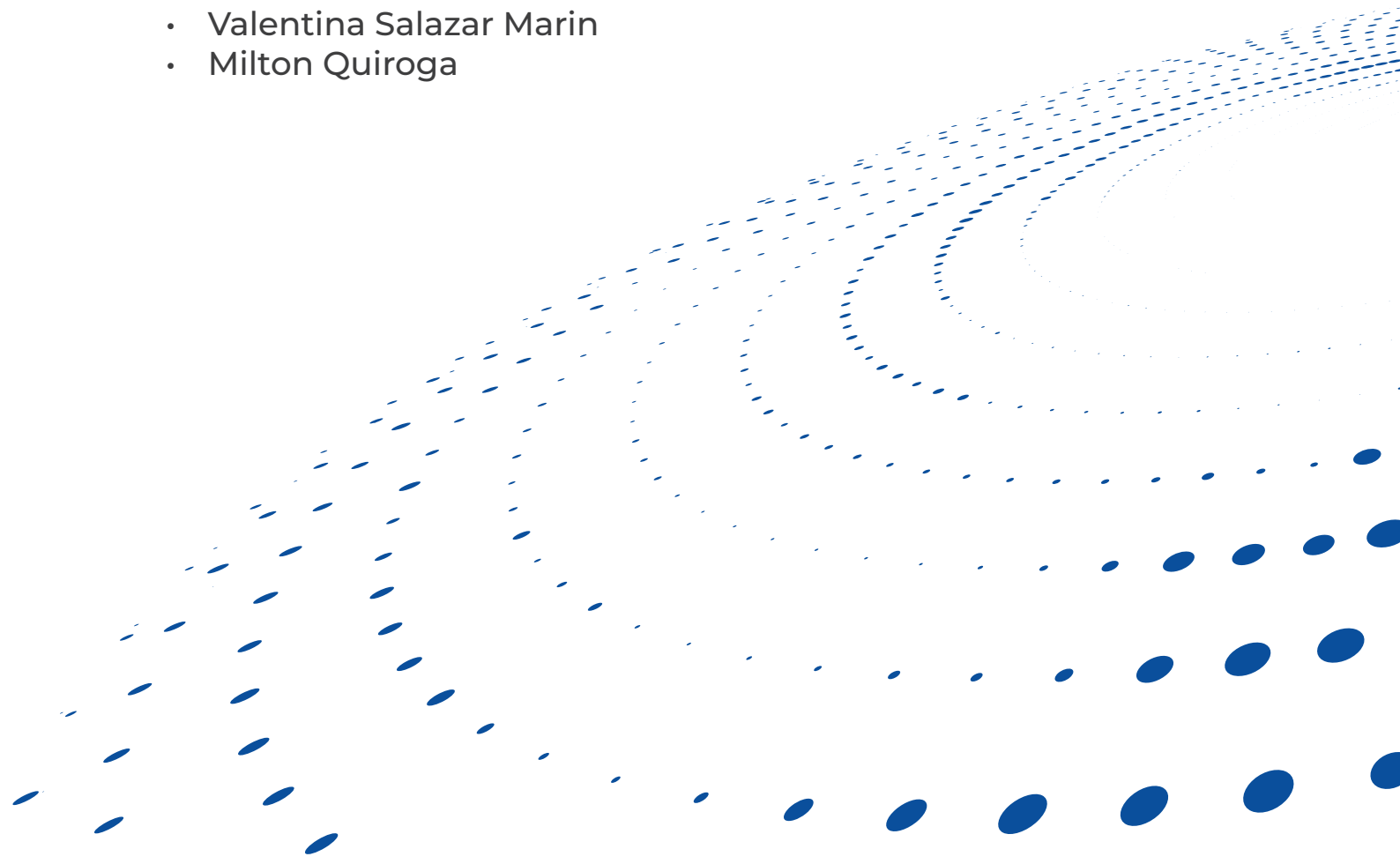
FLAR
FONDO LATINOAMERICANO DE RESERVAS

Factor humano y resiliencia cibernética en bancos centrales

Gobernanza, formación continua y respuesta operativa frente a ingeniería social

Autores:

- Fondo Latinoamericano de Reservas - FLAR
- Bancos centrales miembros
- Valentina Salazar Marin
- Milton Quiroga



Índice general

Índice de tablas

Introducción

La convergencia de sistemas financieros críticos, amenazas basadas en ingeniería social y marcos regulatorios cada vez más exigentes demanda un compendio técnico capaz de articular la dimensión humana de la ciberseguridad con prácticas de mejora continua. El presente documento busca incrementar la resiliencia operativa de los bancos centrales mediante la integración de modelos de madurez, capacitación adaptativa, procedimientos de respuesta, simulacros, alineación de competencias profesionales y consolidación de una cultura organizacional robusta frente a ataques focalizados en las personas.

El alcance comprende seis dominios complementarios que articulan un recorrido progresivo: (i) un marco estratégico sustentado en una gobernanza alineada con CERT-RMM y NIST CSF; (ii) programas formativos modulares que contemplan motivadores conductuales, procesos de aprendizaje y recursos de gamificación; (iii) esquemas de ejercicios realistas sin previo aviso que miden tiempos de reacción y calidad de contención; (iv) directrices operativas para la detección temprana de *phishing* mediante herramientas con analítica avanzada, clasificación estandarizada conforme a MITRE ATT&CK y preservación forense rigurosa; (v) guías para coordinar equipos multidisciplinarios y favorecer la comunicación expedita durante incidentes; y (vi) estrategias de sensibilización sostenida apoyadas en métricas de clima organizacional.

Con este entramado conceptual y práctico, la guía se erige como una brújula técnica que orienta a las instituciones participantes hacia una defensa proactiva, coherente con las exigencias regulatorias y respaldada por la excelencia profesional de su capital humano.

1 Gobernanza estratégica de la seguridad humana

1.1 Autoevaluación de la madurez institucional

Una autoevaluación estructurada es indispensable para calibrar la postura de ciberseguridad de una institución en ámbitos como gobernanza, gestión de riesgos, salvaguarda de activos y respuesta a incidentes [1] [2]. Los cuestionarios alineados con las cinco funciones del marco de seguridad cibernética (CSF) del NIST —identificar, proteger, detectar, responder y recuperar— permiten descubrir brechas y jerarquizar las iniciativas de mejora [3]. Las preguntas deben explorar, entre otros puntos, la existencia y actualización de políticas, la periodicidad de sus revisiones, la designación formal de responsables, la definición de rutas de escalamiento, la integración de la ciberseguridad en la estrategia corporativa y el patrocinio activo de la alta dirección [1].

Con este enfoque, el formulario trasciende la mera enumeración de carencias. Por el contrario, al involucrar a múltiples áreas, se transmite la idea de que la ciberseguridad es un compromiso transversal y no una prerrogativa exclusiva de TI. Tal reflexión colectiva estimula el diálogo interfuncional, desplazando la percepción de un problema técnico hacia la comprensión de un riesgo empresarial compartido. Así, al invitar a empleados y líderes a examinar sus responsabilidades, se consolida la premisa de que la seguridad es tarea de todos, lo cual constituye el fundamento de una cultura de protección resiliente [3].

De acuerdo con lo anterior, se propone el siguiente esquema.

Introducción

Cada ítem debe calificarse en una escala del 1 al 5, siendo:

- **1 (Muy deficiente)**: Ausencia total o grave deficiencia.
- **2 (Deficiente)**: Algunas iniciativas aisladas, pero sin consistencia.
- **3 (Aceptable)**: Prácticas definidas, aunque con áreas importantes por mejorar.

- **4 (Bueno)**: Buenas prácticas implementadas consistentemente con resultados visibles.
- **5 (Excelente)**: Mejores prácticas consolidadas, monitoreo continuo y mejora constante.

Cada sección incluye, además, un espacio para comentarios cualitativos que expliquen la puntuación otorgada y contribuyan al análisis posterior.

Tabla 1.1. Ámbitos de gobernanza estratégica en ciberseguridad.

Pregunta	Calificación
¿Cuenta la organización con políticas formales, aprobadas por alta dirección, que definen claramente las responsabilidades en ciberseguridad?	
¿Existen revisiones periódicas (mínimo anual) documentadas de las políticas de ciberseguridad para reflejar cambios regulatorios, tecnológicos y organizacionales?	
¿Está formalmente designado y reconocido un responsable ejecutivo de ciberseguridad (CISO o equivalente)?	
¿La alta dirección recibe y revisa regularmente informes de desempeño en ciberseguridad (KPI)?	
¿La ciberseguridad se integra explícitamente en los objetivos estratégicos y operativos del banco central?	

Comentarios y observaciones adicionales sobre gobernanza estratégica en ciberseguridad: (Espacio abierto para comentarios)

Tabla 1.2. Gestión del riesgo en ciberseguridad.

Pregunta	Calificación
¿La organización realiza evaluaciones formales de riesgo cibernético al menos anualmente?	
¿Se emplean metodologías reconocidas (p. ej. NIST, ISO 27005) para evaluar y priorizar riesgos cibernéticos?	
¿Están definidas claramente las rutas de escalamiento para riesgos identificados que exceden la tolerancia preestablecida?	

Continúa en la página siguiente

Tabla 1.2. Gestión del riesgo en ciberseguridad. (Continuación)

¿Existen procedimientos documentados de gestión de riesgos que involucran explícitamente a diferentes áreas de negocio (no solo TI)?
¿Las decisiones críticas del banco central consideran explícitamente los riesgos cibernéticos en su análisis previo?

Comentarios y observaciones adicionales sobre gestión del riesgo en ciberseguridad:
(Espacio abierto para comentarios)

Tabla 1.3. Cultura y conducta organizacional.

Pregunta	Calificación
¿Los empleados reciben capacitaciones regulares (al menos anualmente) sobre ingeniería social y prácticas seguras relacionadas con ciberseguridad?	
¿La organización promueve activamente la cultura de reporte inmediato de incidentes sospechosos, sin temor a represalias?	
¿Existen campañas internas de comunicación frecuentes que refuercen la idea de que la ciberseguridad es responsabilidad de todos?	
¿El personal demuestra una conducta proactiva frente a la seguridad (por ejemplo, mediante participación voluntaria en iniciativas o reporte temprano de incidentes)?	
¿La alta dirección y líderes intermedios exhiben comportamientos ejemplares en relación con las políticas y prácticas de seguridad definidas?	

Comentarios y observaciones adicionales sobre cultura y conducta organizacional:
(Espacio abierto para comentarios)

Tabla 1.4. Capacidades operativas para respuesta a incidentes.

Pregunta	Calificación
¿Existen procedimientos formales y documentados para la detección temprana, notificación y escalamiento interno de incidentes de ciberseguridad?	

Continúa en la página siguiente

Tabla 1.4. Capacidades operativas para respuesta a incidentes. (Continuación)

¿La organización ejecuta simulacros prácticos de incidentes de ciberseguridad (por ejemplo, *phishing*) al menos una vez al año?	
¿Se mide regularmente la eficacia de la respuesta operativa (por ejemplo, tiempo promedio de detección y respuesta)?	
¿Existe una clara definición de roles y responsabilidades durante un incidente cibernético, utilizando modelos reconocidos (ej. RACI)?	
¿La organización documenta sistemáticamente las lecciones aprendidas tras cada incidente real o simulado, y realiza ajustes en procedimientos o capacitaciones?	

Comentarios y observaciones adicionales sobre capacidades operativas de respuesta a incidentes: (Espacio abierto para comentarios)

Tabla 1.5. Monitoreo y mejora continua.

Pregunta	Calificación
¿Cuenta la organización con métricas claras y cuantificables (KPI) para evaluar la eficacia global de la ciberseguridad?	
¿Se monitorea periódicamente (mínimo trimestralmente) el estado general de la seguridad mediante herramientas automatizadas (por ejemplo, dashboards)?	
¿Existe un proceso definido y activo de mejora continua (por ejemplo, *Plan-Do-Check-Act*) aplicado específicamente a la seguridad de la información?	
¿Los resultados del monitoreo son reportados regularmente a los líderes organizacionales para toma de decisiones y ajustes estratégicos?	
¿La autoevaluación institucional de ciberseguridad se realiza regularmente (mínimo anual), con un análisis formal y reporte a la alta dirección?	

Comentarios y observaciones adicionales sobre monitoreo y mejora continua: (Espacio abierto para comentarios)

Puntuación total y diagnóstico general:

Suma total de puntuaciones: ____/100 puntos (cada una de las 5 secciones tiene un máximo de 25 puntos posibles)

Interpretación:

- **0-40**: Muy baja madurez institucional. Se requieren acciones urgentes y priorizadas para implementar políticas y prácticas básicas.
- **41-65**: Madurez institucional intermedia. Existen prácticas aceptables, pero múltiples oportunidades de mejora estratégica.
- **66-85**: Buena madurez institucional, aunque restan aspectos por optimizar, especialmente en integración estratégica y cultura.
- **86-100**: Excelente madurez institucional. La organización demuestra prácticas robustas de gobernanza, cultura y respuesta, con capacidad de adaptación continua.

Espacio para conclusiones y recomendaciones finales: (Espacio para elresumen del diagnóstico global y recomendaciones concretas de mejora a implementar.)

Llevar a cabo la actividad requiere (1) que cada representante clave (ciberseguridad, TI, riesgos, RR. HH., comunicaciones y negocios críticos) complete inicialmente el cuestionario de forma individual para captar percepciones diversas; (2) celebrar luego una sesión conjunta facilitada por el equipo de Seguridad de la Información, donde se comparan las respuestas, se debate y se consensúa la puntuación definitiva, documentando argumentos y brechas; y (3) elaborar un informe con los resultados, conclusiones e intervenciones priorizadas, que se eleva a la alta dirección como línea base para el plan de mejora y se repite semestral o anualmente para medir el progreso cultural y técnico.

Como alternativa, el cuestionario puede distribuirse por bloques temáticos; cada área contesta únicamente las secciones que le competen de acuerdo con sus funciones y su exposición al riesgo. Dicha modalidad mantiene la rigurosidad metodológica y, al mismo tiempo, reduce la carga de trabajo, debido a que los expertos profundizan en los apartados donde aportan mayor valor. Una vez recopiladas las respuestas parciales, el equipo de Seguridad de la Información integra los hallazgos y convoca la reunión plenaria para conciliar los puntajes, garantizando que la visión agregada refleje con fidelidad la madurez colectiva y las singularidades operativas de cada dominio.

1.2 Adaptación del CERT-RMM al ámbito regulatorio local

El modelo de resiliencia del CERT (CERT-RMM) constituye un marco conceptual robusto para afianzar la resiliencia operativa mediante la mejora sistemática de procesos inter-

nos, instaurando prácticas organizativas esenciales para afrontar situaciones adversas [4]. Dicho enfoque estratégico posibilita a las instituciones reconocer fortalezas y áreas críticas que requieran atención urgente, trazar objetivos precisos y delinear iniciativas encaminadas a reducir vulnerabilidades, con el fin último de consolidar una respuesta madura, coherente y predecible ante eventos disruptivos.

Tras la evaluación inicial de la postura de protección, el siguiente paso consiste en contextualizar las dimensiones del CERT-RMM en la realidad operativa de un banco central, colocando a la seguridad humana en el centro de todas las iniciativas. El CERT-RMM abarca dominios en torno al entrenamiento y la concientización organizacional (*Organizational Training and Awareness*, OTA), los recursos humanos (*Human Resource Management*, HRM) y las comunicaciones (*Communications*, COMM) [4], ajustables a procesos críticos financieros. Para cada dominio conviene describir escenarios de aplicación puntuales, fijar el nivel de madurez deseado y definir una métrica verificable que mida el progreso. Así, en el área OTA se podría aspirar a que la totalidad del personal complete formación avanzada contra phishing, registrando tanto la tasa de finalización como los resultados en evaluaciones posteriores.

Por otro lado, la adopción local requiere armonizar el CERT-RMM con lineamientos sectoriales, en particular con el NIST CSF. Este último introduce la función «gobernar», que sitúa la ciberseguridad dentro de la dirección estratégica y exige a la alta gerencia participar en la creación de métricas y en la revisión periódica de las cinco funciones operativas (identificar, proteger, detectar, responder y recuperar). Integrar ambas perspectivas refuerza la coordinación de todas las áreas, puesto que alinea objetivos corporativos con prácticas cotidianas enfocadas en la persona: la **oficina de riesgos** establece tolerancias; **recursos humanos** diseña programas de sensibilización; **tecnología** instrumenta controles técnicos; y **comunicación** difunde mensajes que convierten a cada empleado en un «sensor» activo frente a amenazas sociales

El CERT-RMM, más que un catálogo de buenas prácticas, proporciona un andamiaje para la mejora continua. Su filosofía de «proceso sobre proyecto» señala que ningún nivel de madurez constituye un punto final; por el contrario, la entidad debe revisar de forma iterativa sus rutinas a medida que emergen nuevas tácticas hostiles. Tal bucle de retroalimentación transforma al banco central en una organización que aprende a detectar puntos débiles, ajustar procedimientos y perfeccionar la cultura de protección humana de manera sostenida. De este modo, la resiliencia deja de ser un objetivo estático y se convierte en una competencia institucional dinámica, nutrida por la sinergia entre liderazgo, tecnología y cultura organizacional.

La Tabla 1.6 ilustra cómo las dimensiones del RMM pueden aplicarse en el contexto de un banco central, sugiriendo niveles de madurez deseados:

Tabla 1.6: Dimensiones del CERT-RMM aplicadas a bancos centrales.

Dimensión	Objetivos / prácticas clave	Escenario de aplicación en un banco	Nivel de madurez deseado (1-5)
Organizational Training and Awareness (OTA)	- Identificar las necesidades formativas por rol. - Diseñar contenidos alineados con amenazas actuales. - Custodiar evidencias de cursos completados. - Impartir los programas definidos. - Valorar el aprendizaje y la conducta posterior.	El área de talento organiza un itinerario modular en el que todo el personal realiza, al ingresar y después de cada semestre o periodo estipulado, simulaciones de *spear phishing* basadas en incidentes reales del ecosistema financiero latinoamericano.	**4 -** Prácticas estandarizadas y revisadas periódicamente, con indicadores cuantitativos y retroalimentación continua.
Human Resource Management (HRM)	- Definir perfiles y competencias de seguridad. -Alinear talentos con funciones críticas. - Asignar metas de comportamiento seguro. - Manejar cambios en la situación laboral de un empleado.	Recursos humanos, en coordinación con seguridad de la información, define requisitos de confiabilidad para puestos de alto impacto —por ejemplo, operadores SWIFT y administradores de infraestructura crítica— que contemplan revisiones de antecedentes, pruebas de honestidad y conocimiento de protocolos de respuesta. Las evaluaciones de desempeño incorporan métricas de conducta cibernética —como puntualidad en el reporte de mensajes sospechosos—. Para suplir vacantes o ausencias, existe un cuadro de reemplazos entrenados con la misma rigurosidad, con lo que se garantiza la continuidad operativa sin exponer al banco a riesgos por personal improvisado.	**4** – Procesos formales de alineación talento y misión, con vigilancia de competencias y planes de mejora que se mantengan documentados.

Communications (COMM)	-Catalogar interlocutores internos y externos. - Estructurar un plan que detalle canales, mensajes y frecuencia. - Ejecutar la difusión durante operaciones normales y crisis. - Revisar la efectividad de la comunicación y corregir sus deficiencias.	La oficina de comunicación mantiene un plan multicanal avalado por la junta directiva que contempla boletines trimestrales sobre tendencias de ingeniería social, avisos mediante mensajería segura ante incidentes y un protocolo de medios para coordinar la difusión externa cuando un ataque pudiera impactar la confianza pública. Durante simulacros, el «equipo blanco» evalúa si cada audiencia clave —directores, analistas de mercado, proveedores y reguladores— recibe la información adecuada en tiempo y forma, sin términos técnicos ambiguos. Los resultados se archivan en el repositorio de continuidad operativa para auditorías futuras.	**5** – Prácticas optimizadas, validadas por ejercicios realistas y con refinamiento sistemático en cada ciclo de retroalimentación.

1.3 Roles y responsabilidades con matriz RACI

La asignación inequívoca de funciones es crítica en áreas de riesgo elevado, como la protección de la información en entornos financieros. El esquema RACI («Responsable», «Aprobador», «Consultado» e «Informado») constituye un instrumento robusto para distinguir deberes y proporcionar transparencia. El sector bancario opera bajo marcos normativos rigurosos —GLBA, FFIEC, ISO 27001 y NIST CSF— que subrayan la necesidad de roles perfectamente demarcados; por ende, las matrices RACI resultan beneficiosas para la preparación de auditorías, la reducción de retrasos y el fortalecimiento de la cooperación interdepartamental [5]. Una matriz RACI puede aplicarse tanto a tareas sencillas (como actualizaciones de políticas de seguridad de la información) como a procesos complejos (por ejemplo, evaluación de riesgos), precisando quién ejecuta la labor (R), quién asume la salida (A), quién aporta insumos (C) y quién debe quedar informado (I) [5]. Conviene revisar el cuadro periódicamente, sobre todo tras modificaciones en el organigrama, en las herramientas tecnológicas o en los requisitos regulatorios [5].

Más allá de la distribución de tareas, un esquema RACI bien implantado impulsa una cultura de rendición de cuentas y propiedad compartida, neutralizando la percepción de que la ciberseguridad es un asunto exclusivo de TI. En un banco central, donde las interdependencias son recurrentes y un fallo puede adquirir carácter sistémico, esta responsabilidad distribuida se vuelve indispensable para salvaguardar el factor humano. El valor del método radica en su facultad de designar un único custodio del resultado (A), mientras promueve la participación extensa de quienes ejecutan, asesoran o requieren información (R, I y C). Tal estructura impide que las tareas queden relegadas por ambigüedad. Al mismo tiempo, en un entorno con alta aversión al riesgo y necesidad de continuidad operativa, la claridad obtenida se traduce en respuestas más ágiles a incidentes y menor exposición. Además, el personal puede reconocer su aporte a la postura colectiva, trascendiendo perspectivas aisladas y consolidando una cultura protectora cohesionada [6].

La Tabla 1.7 presenta una plantilla RACI, ajustada al contexto de la seguridad del factor humano en un banco central:

Tabla 1.7. Matriz RACI para seguridad del factor humano.

Tarea clave de seguridad humana	Responsable (R)	Aprobador (A)	Consultado (C)	Informado (I)
Desarrollo de política de seguridad para el factor humano.	Analista de seguridad de la información.	CISO	Legal, RR.HH. y auditoría interna.	Alta dirección y jefes de departamento.
Diseño de módulos de capacitación en ingeniería social.	Especialista en capacitación de seguridad.	CISO	Expertos en amenazas, RR.HH. y jefes de departamento.	Empleados y comité de ciberseguridad.
Ejecución de simulacros de *phishing* no anunciados.	Equipo de Operaciones de Seguridad (SOC).	CISO	Comunicaciones, legal y auditoría interna.	Jefes de departamento y alta dirección.
Análisis posterior al simulacro y compilación de lecciones aprendidas.	Analista de incidentes o equipo SOC.	CISO	Jefes de departamento, afectados y RR.HH.	Alta dirección y comité de ciberseguridad.
Gestión de incidentes de *phishing* (triaje inicial).	SOC de nivel 1.	Gerente de SOC o CISO.	Equipo de forense digital y TI.	Usuario afectado y jefes de departamento.
Actualización de procedimientos de respuesta a incidentes.	Equipo de respuesta a incidentes.	CISO	Legal, auditoría interna y TI.	Todos los equipos de seguridad y alta dirección.

En cuanto a las atribuciones mínimas por rol en torno a la misma temática, la Tabla 1.8 describe algunas de ellas.

Tabla 1.8. Responsabilidades mínimas por rol.

Rol	Responsabilidades
CISO	- Definir la estrategia general de seguridad del factor humano. - Aprobar políticas, programas de capacitación y metodologías de simulacros. - Asignar recursos y asegurar el cumplimiento normativo. - Reportar el desempeño de la seguridad del factor humano a la alta dirección.
Analista de seguridad de la información / equipo SOC	- Monitorear alertas de seguridad relacionadas con ingeniería social. - Diseñar y ejecutar simulacros de phishing. - Realizar análisis inicial de incidentes de ingeniería social. - Proponer mejoras en defensas y formación.
Especialista en capacitación de seguridad / RR.HH.	- Desarrollar y actualizar el contenido de los módulos de capacitación. - Implementar programas de concientización y formación continua. - Evaluar la efectividad de la capacitación en cuanto a conocimiento, percepción y comportamiento. - Fomentar una cultura de seguridad positiva.
Jefes de departamento	- Asegurar la participación de su equipo en las capacitaciones y simulacros. - Reforzar las buenas prácticas de seguridad dentro de su área. - Reportar cualquier incidente o sospecha de ingeniería social. - Proporcionar retroalimentación sobre la relevancia de la capacitación.
Equipo legal	- Revisar políticas y procedimientos para asegurar el cumplimiento legal y normativo. - Asesorar durante la gestión de incidentes sobre implicaciones legales y de privacidad. - Participar en la definición de protocolos de comunicación externa.

Continúa en la página siguiente

Tabla 1.8. Responsabilidades mínimas por rol. (Continuación)

Auditoría interna	- Evaluar la eficacia de los controles de seguridad del factor humano. - Corroborar el cumplimiento de políticas y procedimientos. - Revisar los resultados de los simulacros y de las acciones correctivas.

Ambos recursos constituyen un soporte operativo para los bancos centrales, dado que permiten afianzar la responsabilidad en todas las fases del factor humano. Las referencias trasladan el marco RACI desde el plano teórico hacia la práctica cotidiana: asignan sin dilación un responsable final, acotan el ámbito de actuación de colaboradores o consultores y precisan qué niveles jerárquicos deben mantenerse al tanto en cada hito.

1.4 Indicadores ejecutivos y monitoreo de avances

Los indicadores clave de rendimiento (*Key Performance Indicators*, KPI) resultan esenciales para cuantificar la eficacia de la ciberseguridad, seguir la eficiencia de la respuesta ante amenazas y alinear iniciativas con los objetivos del negocio [7]. En consecuencia, los KPI actúan como puente informacional entre equipos técnicos y alta dirección, al aportar datos medibles que impulsan decisiones proactivas y respaldan la priorización de inversiones [7].

La Tabla 1.9 propone un conjunto de KPI y umbrales de referencia [7] [8] [9].

Tabla 1.9. KPI de gobernanza para la dimensión humana (visión ejecutiva).

KPI	Definición	Método de cálculo	Umbral recomendado u objetivo
Tasa de adopción de MFA	Porcentaje de cuentas con MFA activa.	(Usuarios con MFA / Total de usuarios). $\times$ 100.	$\geq$ 95 % global; 100 % en perfiles privilegiados.

Continúa en la página siguiente

Tabla 1.9. KPI de gobernanza para la dimensión humana (visión ejecutiva). (Continuación)

Violaciones de acceso de origen humano	Intentos que contravienen políticas de identidad y acceso (p. ej., aprobación indebida de consentimiento OAuth, compartición de credenciales o almacenamiento no autorizado).	Conteo mensual normalizado por 100 usuarios; clasificación por severidad.	Tendencia descendente; cero incidentes críticos en cuentas privilegiadas.
Cobertura de microcápsulas formativas	Consumo efectivo de cápsulas breves en la frecuencia definida.	(Cápsulas completadas / Cápsulas asignadas) $\times$ 100, por mes y por área.	$\geq$ 85 % mensual; alerta a rezagos persistentes.
Retención a 60–90 días	Conservación del conocimiento transcurrido un intervalo posformación.	(Puntuación posprueba tardía / Puntuación posprueba inmediata) $\times$ 100.	$\geq$ 80 % de retención; refuerzo dirigido bajo ese umbral.
Puntuación promedio en evaluaciones	Media de resultados en pruebas de conocimiento posteriores a cada módulo.	$\sum$ de puntuaciones / Número de participantes	$\geq$ 85 %; plan de refuerzo para valores inferiores.
Índice de clima cultural	Medida agregada de actitudes, hábitos y percepciones en seguridad.	Puntuación global y proporción de dimensiones bajo umbral; comparación interanual.	Mejora $\geq$ 5 puntos porcentuales. semestral; ninguna dimensión $<$ 60/100.
Incidentes reportados por usuarios	Avisos provenientes de la primera línea (reales o sospechosos).	Conteo normalizado por 100 empleados; tendencia por trimestre.	Tendencia ascendente sostenida; objetivo interno por área.

Continúa en la página siguiente

Tabla 1.9. KPI de gobernanza para la dimensión humana (visión ejecutiva). (Continuación)

Time-to-Lesson (TTL)	Días desde un incidente/simulacro hasta la publicación de lección aprendida y asignación de responsable.	Mediana de días por caso cerrado en el periodo.	$\leq$ 30 días para publicación; $\leq$ 60 días para cierre de acción clave.
Verificación fuera de banda (OOB) en procesos críticos	Proporción de instrucciones sensibles validadas por un canal independiente antes de ejecutar.	(Transacciones críticas verificadas OOB / Transacciones críticas revisadas) $\times$ 100.	$\geq$ 95 % en cambios de beneficiario, pagos excepcionales y altas privilegios.

Dicho cuadro otorga a los bancos centrales un repertorio ordenado de métricas operativas, diseñado para observar y robustecer, de manera sistemática, la dimensión humana de la seguridad institucional. El esquema fortalece los procesos decisorios con datos empíricos, ayuda a reconocer con precisión los ámbitos que requieren intervención prioritaria y brinda una vía concreta para demostrar progresos sostenidos, tanto ante órganos internos de gobernanza como frente a entidades supervisoras o autoridades regulatorias.

2 Capacitación modular continua y dinámica

2.1 Diseño de estrategias formativas frente a ingeniería social

El programa formativo debe estructurarse en módulos flexibles, calibrados según el nivel jerárquico y las funciones de cada unidad organizacional; así, el personal recibe contenidos pertinentes y acordes a sus responsabilidades [10]. A su vez, la revisión periódica del temario resulta imprescindible para que la enseñanza refleje la realidad cambiante del panorama ofensivo y la vigencia de las prácticas reconocidas por la industria [10]. Por consiguiente, se sugiere que dichos módulos sean impartidos en sesiones breves y recurrentes que se renueven con frecuencia, alejándose de los tradicionales cursos anuales que no responden oportunamente a los escenarios y a las dinámicas del mundo real.

Entre los vectores contemporáneos, la ingeniería social asistida por inteligencia artificial (IA) ocupa un lugar destacado. Los adversarios ahora cuentan con algoritmos que segmentan objetivos, redactan mensajes persuasivos y elaboran insumos multimedia —correos, SMS, llamadas o publicaciones en redes— con apariencia impecable [11]. Asimismo, los denominados *deepfakes* pueden imitar con precisión voces e imágenes de funcionarios clave, abriendo paso a órdenes fraudulentas y a la divulgación maliciosa de información corporativa [12].

Ante esta coyuntura, la capacitación debe trasladar al alumno desde el clásico «detecta la trampa» hacia un pensamiento crítico avanzado. Como potencial víctima, limitarse a buscar errores ortográficos o remitentes extraños es infructuoso cuando los modelos lingüísticos producen material sin fallos visibles. El énfasis, por tanto, recae en la identificación de anomalías de comportamiento, la corroboración fuera de banda (contacto directo por canales independientes) y la instauración de una cultura de escepticismo. Los empleados necesitan ser motivados para cuestionar el contexto y la intención de las comunicaciones, especialmente las solicitudes urgentes o inusuales, y para veri-

ficar a través de vías alternativas y confiables cuando la sensibilidad de la operación lo amerite [13]. Adicionalmente, se requiere instruir al personal sobre el reporte inmediato de ataques confeccionados con IA, puesto que tales artefactos pueden esquivar filtros automáticos tradicionales. Con estas competencias cognitivas y procedimentales, la fuerza laboral contribuye a la protección integral contra campañas cada vez más sofisticadas.

La Tabla 2.1 ofrece un punto de partida estructurado para diseñar un currículo formativo alineado con las necesidades de los bancos centrales. Su propósito es articular un recorrido pedagógico que cubra las distintas manifestaciones de la ingeniería social [11] [14] [15], tanto las tradicionales como aquellas amplificadas por tecnologías emergentes, mapeando señales observables y respuestas concretas esperadas [16] [17]. Este recurso apunta a la elaboración de un cuerpo temático fielmente contextualizado, dirigido a enfrentar las amenazas más frecuentes y disruptivas que afectan actualmente al sector financiero, con un enfoque claro en la mitigación del riesgo humano.

Tabla 2.1: Módulos temáticos de capacitación frente a ingeniería social.

Tipo de ataque	Descripción breve	Indicadores o señales de alerta	Acciones defensivas para empleados
Phishing (general)	Mensajes engañosos (correo, SMS, web) dirigidos a robar credenciales o desplegar *malware*.	Solicitudes inesperadas o urgentes; sutiles alteraciones en URL/remitente; tono inusual; y enlaces que no coinciden con el texto.	Evitar clics; comprobar remitente y dominio; acceder escribiendo la dirección oficial en el navegador; reportar de inmediato.
Spear-phishing / Whaling	Campañas dirigidas a perfiles concretos o ejecutivos, con datos personalizados.	Mensajes que conocen rol, proyectos o jerarquía; e instrucciones de alto impacto y prisa.	Validar por canal alterno (llamada a número verificado); cuestionar la urgencia y escalar dudas.
Vishing (voice phishing)	Engaño telefónico o por buzón de voz que apela a autoridad o premura.	Llamadas imprevistas de banco/soporte pidiendo datos sensibles; y voz clonada mediante IA.	Colgar y devolver la llamada a un número oficial; desconfiar del identificador; consultar con un colega antes de actuar.
Smishing (SMS phishing)	Texto con enlaces maliciosos o solicitud de información.	Mensajes sobre verificación de cuenta, premios o «bloqueos»; y acortadores de URL.	No pulsar enlaces; eliminar; bloquear remitente; confirmar con la entidad por canal institucional.
***Deepfakes* (video/audio)**	Contenido sintético generado por IA que suplantan voces o imágenes.	Desajustes en sincronía labial; artefactos visuales; tono robótico; e instrucciones de pago «confidenciales».	Corroborar instrucciones críticas mediante verificación independiente y regla de dos personas; jamás confiar en el material audiovisual como prueba de identidad.

Ataques con IA (general)	Uso de IA para personalizar, automatizar y escalar campañas.	Gramática impecable, pero en peticiones inusuales; e interacción prolongada para ganar confianza.	Mantener escepticismo; comprobar identidad por múltiples puntos; reportar cualquier anomalía.
Business Email Compromise (BEC)	Fraude para ordenar transferencias o modificar beneficiarios, con suplantación de dominio o buzón comprometido.	Cambio «por única vez» de la cuenta; veto a llamadas; presión temporal; y reuniones virtuales con audio o vídeo alterado.	Confirmación fuera de banda; doble control para variaciones de pago; aviso inmediato a seguridad y finanzas.
Telephone-oriented attack delivery [18]	Correo que induce a llamar a un «soporte» para completar la estafa, frecuentemente seguido de acceso remoto.	Facturas falsas con número telefónico; suscripciones que «expiran hoy»; e instrucciones de instalación por voz.	Colgar y contactar al proveedor mediante canal oficial; prohibir instalación guiada por teléfono; escalar y registrar el intento.
Fatiga MFA [19]	Ráfagas de notificaciones de segundo factor para forzar aprobaciones por cansancio.	Solicitudes de aprobación en serie sin inicio de sesión; y peticiones en horario atípico.	Rechazar cada aviso; activar *number matching* y geolocalización de intentos; cambiar credenciales y reportar.
***Quishing* (QR phishing)**	Códigos QR adulterados que redirigen a portales fraudulentos o descargas.	Adhesivos superpuestos en POS/estacionamientos; URL no oficial tras el escaneo; y urgencias de pago o multas.	Teclear manualmente la dirección; revisar la vista previa del enlace; evitar los códigos en superficies dudosas; y notificar a seguridad ante sospecha.

***Consent phishing* / Illicit OAuth grant** [20]	Aplicaciones que solicitan consentimientos OAuth excesivos para acceder a correo, archivos y calendario sin robar contraseñas.	Pantallas con permisos desproporcionados u origen desconocido; e insistencia en «autorizar para continuar».	Rechazar permisos no solicitados; confirmar si la app está aprobada; reportar y revocar accesos en el portal; y llevar a cabo concientización sobre señales de consentimiento engañoso.
***USB baiting* (cebo USB)** [21]	Dispositivos dejados en áreas comunes que buscan ser conectados para ejecutar cargas maliciosas.	Memorias «olvidadas» con etiquetas llamativas.	No conectar; entregar a seguridad; campañas internas con simulaciones controladas y recordatorio de prácticas seguras.

2.2 Implementación a través de ciclos recurrentes y micro-aprendizaje

La capacitación en ciberseguridad debe contrarrestar tanto la pérdida progresiva de conocimientos como la apatía cognitiva que puede surgir entre los empleados frente a temarios repetitivos o desactualizados. La transición desde esquemas formativos anuales, instaurados meramente como requisito de cumplimiento, hacia modelos de microaprendizaje dinámico, gamificado y adaptativo [10] [22] no responde a una lógica de entretenimiento, sino a una necesidad estratégica de fomentar un comportamiento seguro e instintivo. En ecosistemas exigentes, como en los que se desenvuelven los bancos centrales, donde las decisiones deben tomarse bajo presión y sin margen de error, la capacidad de respuesta ante intentos de manipulación social se construye a través de un refuerzo continuo y consciente, no mediante la memorización. El propósito último no reside únicamente en transmitir teoría, sino en transformar patrones conductuales, fortaleciendo lo que puede denominarse una «memoria procedimental» en materia de seguridad.

La Tabla 2.2 describe diversas metodologías didácticas junto con su frecuencia sugerida. Este recurso actúa como una guía estructurada que facilita a los bancos centrales transitar desde modalidades convencionales y poco funcionales hacia enfoques envolventes, diseñados para propiciar transformaciones reales en el comportamiento humano frente a amenazas sociotécnicas.

Tabla 2.2: Metodologías y frecuencia sugerida para formación continua.

Metodología	Descripción y beneficios	Ejemplos de implementación	Frecuencia sugerida
Gamificación	Transforma el aprendizaje en una experiencia interactiva y divertida, aumentando tanto la participación como la retención.	Concursos de ciberseguridad con tablas de clasificación, *escape rooms* virtuales, desafíos de detección de amenazas con puntos y recompensas [23].	Mensual (minijuegos o cuestionarios) y trimestral (competiciones mayores).
Microaprendizaje	Enseña por medio de pequeñas dosis fáciles de digerir, por lo que se adapta a las agendas ocupadas y mejora la retención a largo plazo.	Vídeos cortos (2-5 min) sobre una amenaza puntual, infografías interactivas y consejos de seguridad diarios por correo electrónico [10].	Diaria/semanal (cápsulas informativas) y mensual (módulos cortos).
Capacitación adaptativa	Personaliza la ruta de aprendizaje según el rendimiento del usuario en simulacros y evaluaciones, enfocándose así en áreas de mejora.	Simulaciones de *phishing* que se ajustan en dificultad y tipo de engaño según la respuesta del empleado [24].	Continua (basada en el rendimiento individual).
Talleres prácticos	Permite a los empleados aplicar conocimientos en escenarios realistas, desarrollando habilidades de respuesta y toma de decisiones.	Simulaciones de respuesta a incidentes de ingeniería social, ejercicios de mesa para discutir protocolos de comunicación en crisis [23].	Trimestral (para equipos de respuesta) y semestral (para toda la organización).
Videos y casos reales	Ilustra las amenazas y consecuencias de forma vívida, tal que las reflexiones son más memorables.	Videos cortos con ejemplos de ataques reales y análisis de casos de estudio de ingeniería social en el sector financiero.	Mensual (como parte de microaprendizaje) y trimestral (encuentros para discutir sobre los recursos multimedia provistos).

2.3 Evaluación de la efectividad (conocimiento, retención y transferencia)

La valoración del impacto formativo exige un marco que combine métodos heterogéneos: encuestas, ejercicios de simulación, registros de incidentes, supervisión continua y evaluaciones estructuradas [10] [25]. Focalizar la medición en la transformación conductual, más allá de la simple retención de información, constituye un estadio avanzado en la protección del componente humano. Para un banco central, esta perspectiva trasciende la lógica de cumplimiento («¿se completó el curso?») y acoge una mentalidad de mitigación del peligro («¿se observa un actuar prudente en el día a día?»), lo que influye directamente en la resiliencia de la organización frente a los ataques dirigidos a las personas. Conocer un protocolo dista de aplicarlo cuando impera la presión o el engaño. En instituciones cuya exposición a la ingeniería social acarrea pérdidas cuantiosas, deterioro reputacional e incluso riesgo sistémico, la métrica definitiva reside en la manifestación de comportamientos seguros, verificables en pruebas controladas o incidentes reales. Asociar dichos resultados con la capacitación en conciencia cibernética puede apuntar a una reducción palpable del riesgo [24], evidencia que refuerza la noción de un «cortafuegos humano» que emerge como pilar fundamental de la continuidad operativa.

La Tabla 2.3 brinda a los bancos centrales un mecanismo para valorar el impacto real de sus iniciativas. Al cuantificar el conocimiento, la retención y la transferencia, este instrumento posibilita una visión matizada de las fortalezas y los ámbitos susceptibles de mejora [25]. Dicha metodología propicia un refinamiento permanente y garantiza que la inversión en formación se refleje en incrementos tangibles de resiliencia humana.

Tabla 2.3: Técnicas y KPI para evaluar conocimiento, retención y transferencia.

Categoría de evaluación	Técnica recomendada	KPI asociado	Interpretación y vinculación a efectividad
Conocimiento	**Cuestionarios pre y post por módulo**	Puntuación promedio; ganancia de aprendizaje (%): (post − pre) / pre × 100	Mide comprensión conceptual y progresión inmediata. Ganancia positiva sostenida por tema indica asimilación; valores bajos sugieren refuerzo focalizado [26].
	Evaluaciones recurrentes	Retención a 60–90 días (%): (posprueba tardía / posprueba inmediata) × 100.	Estima conservación del conocimiento y necesidad de recordatorios. Caídas pronunciadas orientan microcápsulas dirigidas y ajustes de frecuencia.
Percepción	**Encuestas breves de utilidad y relevancia**	Utilidad percibida; pertinencia del contenido (escalas Likert).	Alta valoración correlaciona con mayor adopción de prácticas; respuestas bajas guían revisión de ejemplos, lenguaje y contexto sectorial [25].
	Autoeficacia para responder a incidentes	Índice de autoeficacia por escenario (capacidad percibida para actuar acertadamente)	Aumentos consistentes tras cada ciclo reflejan confianza operativa; estancamiento exige práctica guiada y ejercicios de roles.
	Encuestas de clima de ciberseguridad	Índice DVMS-CAT / HAIS-Q / SeBIS	Explora actitudes y hábitos; identifica fricciones culturales y comunica prioridades de intervención [27].
Transferencia	**Evaluaciones situacionales**	Tasa de decisiones acertadas por caso; tiempo medio de respuesta; índice de pensamiento crítico.	Evidencia aplicación en contexto: reconocimiento de señales, elección de OOB y escalamiento adecuado. Tendencia ascendente indica internalización.

	Hábitos en herramientas de trabajo	Uso del botón de reporte en el cliente de correo electrónico; tasa de OOB en flujos críticos.	Observa conductas clave en el puesto. Incrementos estables reflejan transferencia; valores bajos disparan educación puntual y recordatorios.
	Refuerzo con microaprendizaje	Cobertura de microcápsulas; mejora en % de aciertos poscápsula.	Vincula consumo con desempeño inmediato. Mejora sin cobertura sugiere aprendizaje informal; cobertura sin mejora exige rediseño instruccional.

3 Simulacros y evaluación operativa en tiempo real

3.1 Metodología para simulacros no anunciados

En este ámbito, los simulacros comprenden dos categorías complementarias, ajustables al grado de madurez institucional. La primera, formativa y de bajo impacto, reproduce campañas de ingeniería social mediante señuelos controlados (correo, SMS, mensajería o códigos QR), con énfasis en el actuar del personal —detección, reporte oportuno y adhesión a los canales oficiales—, sin intervención sobre activos productivos ni exposición de información [28].

La segunda, de mayor sofisticación, corresponde a ejercicios de *red/blue team* —y, cuando procede, *purple*— los cuales emulan cadenas de ataque extremo a extremo y evalúan la eficacia de las salvaguardas de detección y defensa bajo reglas de compromiso estrictas: telemetría en tiempo real, coordinación y cooperación, tiempos de reacción y resiliencia operativa.

Ambas modalidades replican tácticas plausibles sin afectar la integridad de los sistemas ni la confidencialidad de los datos y aportan insumos cuantitativos para la mejora continua. Puntualmente, funcionan como verdaderas pruebas de estrés para la defensa humana: exponen no solo lagunas cognitivas, sino también reacciones ante condiciones verosímiles de presión, emergencia o manipulación emocional. Este tipo de actividades trasciende la conciencia meramente conceptual y se orienta hacia la resiliencia operativa, un aspecto crucial para un banco central. Su carácter sorpresivo y realista, inclusive con amenazas de última generación, somete al personal a decisiones espontáneas, revelando la respuesta instintiva más allá del conocimiento memorizado. De este modo, emergen las susceptibilidades que los atacantes explotan justamente cuando impera la urgencia [13].

Para una entidad emisora, tales ejercicios configuran un laboratorio controlado y de bajo impacto donde es plausible detectar grietas críticas del cortafuegos humano antes de enfrentar un incidente genuino con consecuencias sistémicas. Adicionalmente, los

criterios cuantitativos recabados (tiempos de reporte, rutas de escalamiento seguidas, patrones de clics o divulgación de credenciales) nutren un programa de capacitación hiperfocalizado, centrado en reforzar conductas y áreas funcionales con mayor exposición. Así, se optimiza el retorno de la inversión en concientización, puesto que se canalizan recursos hacia segmentos donde la intervención pedagógica produce el mayor beneficio en términos de reducción del riesgo vinculado con el recurso humano.

3.1.1 Simulación de phishing

Respecto a la primera categoría, el diseño de un simulacro debe contemplar las siguientes etapas:

1. **Formulación de objetivos estratégicos**: Delimitar las metas a las que apuntan las iniciativas, las cuales pueden abarcar: cuantificar la tasa actual de clics en enlaces maliciosos, examinar la sensibilidad del recurso humano frente a técnicas emergentes de phishing, registrar la proporción de usuarios que reportan de modo proactivo la amenaza, descubrir falencias de comportamiento en áreas expuestas a mayor riesgo o medir los efectos concretos de intervenciones formativas recientes [28].
2. **Construcción del escenario**:
    - Elaborar plantillas realistas que reflejen amenazas reales sin causar estrés indebido [28].
    - Incorporar elementos contextuales relevantes (noticias del sector financiero, procesos internos o eventos corporativos vigentes), así como recursos visuales auténticos (gestión de marca de servicios legítimos), lenguaje creíble y disparadores (*triggers*) psicológicos comunes (urgencia, autoridad, curiosidad o miedo) [28].
    - Desplegar una amplia gama de modalidades: spear-phishing (personalizado), vishing (vía telefónica), smishing (mensajes de texto), pretexting (suplantación con historia falsa), baiting (ofertas llamativas), whaling (dirigido a ejecutivos) y variantes más sofisticadas potenciadas por inteligencia artificial, como los deepfakes.
3. **Distribución controlada**: El envío debe llevarse a cabo de manera escalonada o segmentada para crear situaciones plausibles y evitar alertas colectivas que distorsionen los resultados [26]. La aleatorización por unidades organizativas, niveles jerárquicos o zonas horarias contribuye a obtener una lectura más fidedigna del comportamiento espontáneo del personal.

4. **Monitoreo de respuestas**: Durante el ejercicio, es fundamental registrar indicadores clave: aperturas del mensaje, clics en enlaces, descargas de archivos adjuntos, ingreso de credenciales en formularios fraudulentos, reenvíos internos y reportes voluntarios al canal oficial de seguridad [26]. Estas métricas no solo reflejan el grado de exposición, sino también la capacidad de reacción bajo presión.
5. **Retroalimentación personalizada y refuerzo educativo**: A los receptores se les debe redirigir, de forma inmediata o diferida, a una página explicativa que revele la naturaleza del ejercicio, ofrezca pistas sobre las señales de fraude omitidas y proporcione recomendaciones para impedir que se materialicen incidentes reales. La intervención puede complementarse con sesiones formativas adicionales, enfocadas en las debilidades exhibidas durante el simulacro [26]. Tal enfoque favorece la asimilación de conductas seguras y fortalece la cultura organizacional frente a amenazas de manipulación psicológica.

Entre las consideraciones metodológicas más relevantes se encuentran: no recurrir a patrones repetitivos que puedan ser reconocidos por el personal y, por ende, neutralizar la espontaneidad de la respuesta; dosificar las campañas para prevenir el fenómeno conocido como fatiga de simulación; y alinear la programación con los ciclos del banco central, de modo que las pruebas no interfieran con las funciones críticas [28]. Asimismo, se sugiere incrementar progresivamente la frecuencia de las simulaciones, con el objetivo de consolidar una cultura de vigilancia sostenida sin inducir al agotamiento cognitivo [28].

Respecto a la fijación de blancos estratégicos, las iniciativas dirigidas a perfiles con exposición ampliada —como administradores de sistemas, personal de pagos interbancarios o departamentos estratégicos— son esenciales debido a que concentran mayores niveles de acceso [28] y, por consiguiente, representan vectores privilegiados desde la perspectiva del atacante.

3.1.2 Ejercicios de red team y purple team

Red team. Su propósito consiste en emular con plena fidelidad la cadena de ataque de un adversario real [29]. Así, además de procedimientos puramente técnicos (*exploits*, movimiento lateral o exfiltración), el equipo atacante suele incluir fases de reconocimiento social, pretexting, vishing o phishing dirigidos [29]. En un banco central, esta aproximación se conduce bajo estricta confidencialidad: un conjunto reducido del personal (designado como el equipo blanco) conoce la actividad a realizar, tal que las reacciones sean genuinas. Entre los marcos y metodologías que regulan este tipo de ejercicios, destacan:

- CBEST Intelligence Led Testing del Banco de Inglaterra.

- Threat Intelligence-Based Ethical Red Teaming (TIBER-EU), a cargo del Banco Central de Europa (BCE).
- Red Team: Adversarial Attack Simulation Exercises – ABS (Association of Banks of Singapore).
- intelligence-led Cyber Attack Simulation Testing (iCAST) – HKMA (Hong Kong Monetary Authority).
- G-7 Fundamental Elements for Threat-Led Penetration Testing (G7FE-TLPT)
- A Framework for the Regulatory Use of Penetration Testing and Red Teaming in the Financial Services Industry – GFMA (Global Financial Markets Association).

TIBER-EU [30] o CBEST [31] exigen, de hecho, que el vector humano forme parte del escenario, puesto que la mayoría de las intrusiones graves apuntan a este.

Purple team Cuando la organización busca un aprendizaje más inmediato, se convoca a los equipos ofensivos (red team) y defensivos (blue team) en sesiones conjuntas. La parte ofensiva emprende la campaña de ingeniería social, mientras que la parte defensiva vigila alertas y despliega contramedidas en tiempo real. Al culminar cada fase, ambas comparten indicadores y refinan tácticas o mecanismos de detección [30]. Este formato es idóneo para reforzar procedimientos de reporte, validar el tiempo de reacción frente a correos maliciosos o llamadas fraudulentas, así como afinar reglas de correlación en el SIEM.

3.1.3 Plataformas comerciales y soluciones de código abierto

El mercado ofrece soluciones comerciales consolidadas como KnowBe4, Proofpoint Security Awareness Training, Cofense PhishMe, Infosec IQ y Hoxhunt [32]. Las funcionalidades que otorgan incluyen desde bibliotecas de plantillas extensas, microaprendizaje con analíticas granulares e informes avanzados sobre tasas de reporte, hasta campañas gamificadas [22] y módulos didácticos de alta calidad, lo que resulta idóneo para organizaciones sin personal propio de formación o que demandan paneles ejecutivos exhaustivos [28].

En contraste, entornos con restricciones presupuestarias pueden recurrir a plataformas abiertas como Gophish, Social-Engineer Toolkit o King Phisher [28]. Estas herramientas permiten crear campañas, diseñar plantillas y extraer estadísticas básicas, pero exigen una mayor pericia técnica para realizar la configuración manual de servidores SMTP o certificados, por ejemplo. Además, prescinden del tipo de experiencia de usuario, soporte dedicado y amplitud de catálogos que ofrecen las alternativas comerciales.

Para un banco central, la decisión suele sustentarse en la cobertura de contenido, la

inteligencia de amenazas, la recopilación de métricas y la elaboración automática de reportes, las integraciones con otras plataformas y factores diferenciadores tales como: entrenamiento en torno a requisitos legislativos y de privacidad de los datos [33], alto grado de personalización, retroalimentación instantánea, gamificación adaptativa [22], entre otros.

3.2 KPI para supervisar el desempeño conductual y la capacidad operativa

Integrar indicadores de rendimiento centrados en el factor humano (como el tiempo de reporte o TTR) con métricas funcionales de mayor espectro (por ejemplo, tiempo medio de detección o MTTD y tiempo medio de contención o MTTD) en los ejercicios de simulación brinda una panorámica amplia acerca de la resiliencia institucional. Dicha convergencia revela con exactitud los puntos en los que la reacción del personal incide en la velocidad y la calidad de la respuesta técnica, guiando las acciones posteriores para abordar las deficiencias.

La Tabla 3.1 constituye una herramienta sistemática y cuantificable para que los bancos centrales examinen en profundidad el desempeño de sus ejercicios de simulación vinculados al componente humano de la seguridad [24] [28]. El recurso facilita el seguimiento longitudinal de los resultados, permitiendo trazar patrones de mejora o estancamiento a lo largo del tiempo. Asimismo, posibilita la identificación de unidades organizativas o perfiles funcionales con mayor propensión o exposición a vulnerabilidades de esta naturaleza, lo cual habilita intervenciones focalizadas.

Tabla 3.1: KPI sugeridos para simulacros no anunciados.

Categoría	KPI	Definición / Fórmula	Meta recomendada
Actuación individual	Tasa de clic	(usuarios con acción no deseada / mensajes entregados a destinatarios válidos) $\times$ 100	Tendencia descendente; $\leq$ 3 %
	Tasa de reporte	(primer reporte válido / mensajes entregados a destinatarios válidos) $\times$ 100	Tendencia ascendente; $\geq$ 90 %
	TTR–Usuario (Tiempo hasta el primer reporte)	Mediana de minutos desde la entrega del mensaje hasta el primer reporte humano	Tendencia descendente; $\leq$ 5 minutos
	Verificación fuera de banda (OOB)	(instrucciones sensibles verificadas por canal alterno antes de actuar / instrucciones sensibles expuestas) $\times$ 100	$\geq$ 75 % inicial; mejora por ciclo
	Reincidencia	(personas que repiten la falla en 12 meses / personal total) $\times$ 100	$\leq$ 1 %
Madurez cultural	Índice de pensamiento crítico	Porcentaje que reconoce las anomalías contextuales en un cuestionario posterior al simulacro	$\geq$ 75 %
	Tasa de esfuerzo compensatorio	(consultas a mesa de ayuda o compañero frente a un mensaje sospechoso / mensajes entregados) $\times$ 100	Incremento estable por ciclo
	Índice de escalamiento correcto	(eventos que siguieron el flujo de escalamiento previsto / eventos que requerían escalamiento) $\times$ 100	$\geq$ 90 %
Capacidad SOC (*blue team*)	MTTD – Mean Time to Detect	Minutos entre entrega y primera alerta del SOC (humana o SIEM)	Tendencia descendente; correlacionar con TTR–Usuario

	MTTC – Mean Time to Contain	Minutos desde la alerta hasta el bloqueo global (URL, remitente o firma)	Tendencia descendente; $\leq$ 15 minutos
	Contención en ventana	(mensajes o dominios bloqueados en $\leq$ 15 minutos / total a bloquear) $\times$ 100	$\geq$ 90 %
	Tasa de erradicación	(indicadores retirados en $\leq$ 24 horas / indicadores totales del ejercicio) $\times$ 100	$\geq$ 95 %
Calidad de ejercicios de *purple team*	Índice de compromiso inicial	objetivos humanos comprometidos / objetivos totales	30–50 % para ejercicios de caja negra
	Profundidad de movimiento lateral	saltos sucesivos logrados por el *red team* tras acceso inicial	Contención antes de 2 saltos
	Tasa de cierre de brechas	hallazgos corregidos en $\leq$ 30 días / hallazgos totales	$\geq$ 80 %
Rendimiento agregado	Reducción porcentual de riesgo humano	(tasa base − tasa actual) / tasa base (definir línea base del año)	$\geq$ 10 % de mejora semestral
	Índice compuesto de conciencia de seguridad	ponderación de KPI de conducta (clic, reporte, TTR–Usuario y OOB) y conocimiento (pospruebas)	Tendencia ascendente; objetivo $\geq$ 85 %

La relación causal entre la preparación del personal y la reducción de los tiempos de respuesta ante incidentes es contundente. Cuando se refuerza la atención y la capacidad de discernimiento de los empleados mediante programas formativos rigurosos y simulaciones diseñadas para replicar condiciones reales de presión, el TTR vinculado a eventos iniciados a través de ingeniería social tiende a disminuir. La reducción no solo refleja un mayor nivel de conciencia individual, sino también una interiorización de los protocolos de actuación ante señales de manipulación.

Adicionalmente, al correlacionar este indicador con parámetros de SOC, es factible descubrir dinámicas reveladoras. Por ejemplo, una caída simultánea de TTR y MTTD podría interpretarse como evidencia de que el personal actúa como un sensor distribuido dentro del tejido organizacional, lo que favorece la visibilidad temprana de amenazas por parte del equipo técnico. Por el contrario, si el MTTC no se reduce de forma paralela al MTTD, podría sugerir que el cuello de botella no reside en la detección automática ni en la vigilancia ejercida por los empleados, sino en los procedimientos de contención y escalamiento posteriores. Procediendo así, la lectura integrada permite refinar no solo las estrategias de formación, sino también los flujos operativos, cerrando brechas donde la reacción técnica no se halla alineada con la alerta humana.

Por último, al recopilar activamente datos de seguimiento, el banco central demuestra el retorno de los esfuerzos invertidos en las actividades de los individuos y el rol que desempeñan en la arquitectura de respuesta, reforzando la justificación financiera para sostener e incluso ampliar la inversión en esta dimensión estratégica [24].

3.3 Lecciones aprendidas y mejora continua

La extracción sistemática de aprendizajes tras cada incidente o ejercicio es esencial para comprender las causas raíz y articular actividades que atiendan factores organizacionales, técnicos y humanos [34].

La documentación de los resultados de los simulacros debe incluir [26] [34]:

- **Descripción exhaustiva del escenario**: vector reproducido, técnica empleada, público objetivo y qué tan fielmente se recrean los atributos de un ataque real.
- **Bitácora del personal**: operaciones culminadas (clics, reportes u omisiones) y su correspondencia con los roles funcionales de cada área.
- **Caracterización de las brechas latentes**: aspectos técnicos (lagunas de telemetría, segmentación deficiente o fallos de contención) y procedimentales (ausencia de rutas de escalamiento precisas, comunicación interna escasa o planes de continuidad incompletos).

En cuanto al informe de lecciones aprendidas, la alta dirección debe respaldar explícitamente su revisión, promoviendo un ambiente donde el error se estudie sin recriminaciones ni represalias. Tal clima motiva declaraciones honestas y acelera la detección de factores subyacentes [35]. En suma a ello, se sugiere que la sesión de retrospectiva se lleve a cabo con la mayor inmediatez posible, cuando la información aún se halla fresca y los recuerdos del equipo permanecen nítidos.

Respecto al contenido, es necesario consignar los planes de respuesta activados, cronologías operativas, bitácoras de sistemas, artefactos forenses y cualquier otro recurso que ilustre la línea de tiempo, la toma de decisiones y las consecuencias hipotéticas del ataque. Asimismo, es indispensable que se reconozcan todos los actores involucrados: personal técnico, equipos interdisciplinarios, áreas de negocio impactadas y, cuando proceda, entidades externas —proveedores, organismos reguladores o consultores especializados— que contribuyeron a la resolución o al diagnóstico posterior [35].

Algunas preguntas guía se listan a continuación [35]:

- ¿Cómo respondieron el personal y los directivos durante la contingencia?
- ¿Los procedimientos existentes mostraron suficiencia o demandaron improvisaciones?
- ¿Qué información adicional habría agilizado la toma de decisiones?
- ¿Existieron determinaciones que ralentizaron la recuperación?
- ¿Qué mecanismos de intercambio de inteligencia conviene reforzar con organismos afines?

Los simulacros no anunciados son infructuosos si las lecciones no se plasman en un proceso de afianzamiento meticuloso. Al completar cada ejercicio, la entidad debe consignar los hallazgos de modo preciso y exhaustivo, para luego incorporarlos en la arquitectura de seguridad preventiva y en los planes operativos de contención [34]. Una vez entendidos los puntos débiles, las correcciones han de situarse en la matriz de prioridades y ejecutarse con prontitud [34]. Dicha dinámica se alinea con la filosofía DevSecOps, donde el flujo constante de retroalimentación y la consolidación de criterios de rendimiento propician el ajuste iterativo de procedimientos [36]. De tal modo, el circuito «simular $\rightarrow$ analizar $\rightarrow$ aprender $\rightarrow$ mejorar $\rightarrow$ simular» convierte cada ejercicio en un eslabón de un ciclo de perfeccionamiento ininterrumpido [36].

Al diseccionar tanto los errores como los aciertos, la entidad trasciende la remediación puntual de fallos y aborda desequilibrios sistémicos en los procesos, la gobernanza y la cultura. Por ello, resulta esencial un enfoque de refuerzo positivo y no punitivo: reconocer conductas deseables, brindar retroalimentación constructiva y establecer mecanismos de acompañamiento que afiancen la confianza del personal. Así, un entorno

exento de reproches fomenta la notificación temprana y el autoexamen honesto [35], condiciones clave para un aprendizaje genuino. Así, la postura de seguridad centrada en el factor humano permanece robusta y pertinente ante amenazas emergentes, convirtiendo al banco en una organización verdaderamente resiliente.

4 Procedimientos y controles contra ingeniería social en banca central

4.1 Enfoque multicapa contra el phishing

Aludiendo a la guía del NCSC [17], la defensa frente al phishing debe articularse como un enfoque multicapa que combine lineamientos técnicos, de proceso y de cultura organizacional para crear múltiples puntos de detección, contención y aprendizaje con mínima fricción para el personal: primero, dificultar el contacto de los atacantes con los usuarios; después, favorecer el reconocimiento y el reporte oportuno mediante canales claros y retroalimentación útil; luego, amortiguar el impacto cuando un mensaje supera los filtros, reforzando la autenticación, la navegación segura y la protección de dispositivos; y, finalmente, responder con celeridad y de modo coordinado, estandarizando *playbooks* y formatos de documentación del incidente.

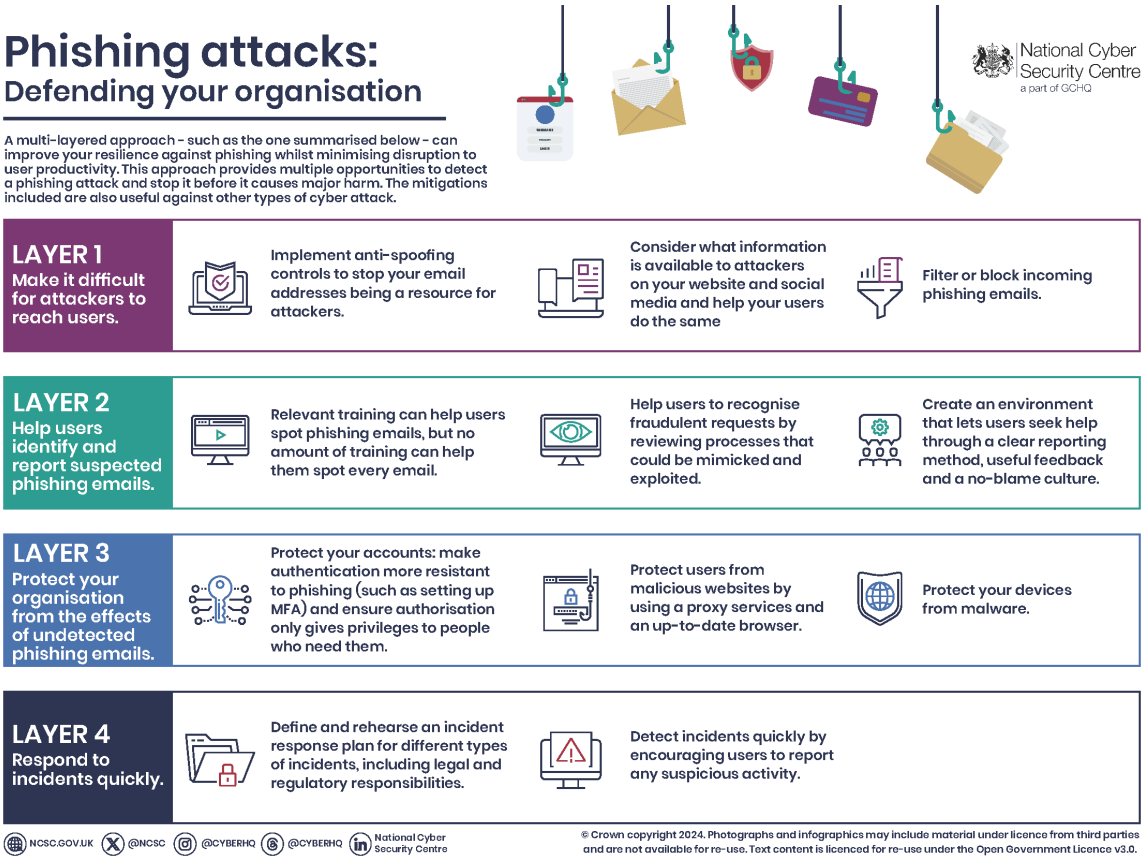


Figura 4.1: Resumen de las defensas por capas ante ataques de ingeniería social. Fuente: *Phishing attacks: defending your organisation* [17].

4.1.1 Capa 1: bloqueo del contacto con usuarios

Se busca aminorar drásticamente el volumen y la verosimilitud de las cargas fraudulentas antes de que lleguen a la bandeja de entrada, combinando autenticación del correo, higiene de la huella pública y filtrado preventivo [17].

En concreto, la guía del NCSC recomienda desplegar DMARC (*Domain-based Message Authentication, Reporting, and Conformance*), SPF (*Sender Policy Framework*) y DKIM (*DomainKeys Identified Mail*) para frustrar la suplantación de dominios —idealmente con políticas en *enforcement (p=quarantine/reject)*—, tal que los mensajes maliciosos sean bloqueados o puestos en cuarentena por los receptores. Respecto a la funcionalidad, DKIM añade una firma digital a los correos transmitidos por la organización y con SPF se indica qué servidores se hallan autorizados a enviar correos en nombre de su dominio, mientras que DMARC integra ambos recursos para proporcionar mecanismos de notificación y cumplimiento [37].

En paralelo, conviene acotar la exposición de datos accesibles en sitios corporativos o en perfiles públicos, que posibilitan la fabricación de pretextos personalizados, y aplicar filtrado o bloqueo del tráfico entrante para neutralizar adjuntos, URLs y patrones cono-

cidos antes del despacho al buzón. En conjunto, tales medidas interrumpen campañas a gran escala y disminuyen el tiempo que el personal dedica a comprobar y reportar la correspondencia sospechosa [17].

4.1.2 Capa 2: formación y cultura de reporte

El NCSC subraya que la capacitación es valiosa, pero insuficiente si no se acompaña de procedimientos robustos y una cultura de reporte no punitiva que incentive la participación sostenida: conviene revisar flujos susceptibles a la suplantación (como aprobaciones, cambios de cuenta o solicitudes de datos), incorporar verificaciones por un segundo canal en interacciones sensibles, habilitar un mecanismo de reporte de «un clic» en el cliente de correo y disponer de un canal directo al SOC con retroalimentación oportuna [17].

El NCSC desaconseja instaurar una cultura punitiva o de culpabilización en torno al phishing, puesto que esto inhibe el aviso temprano y reduce la colaboración. En su lugar, se recomienda promover refuerzos positivos, colaborar con el individuo cuando se enfrente a un intento de ingeniería social y suministrar soporte práctico, todo enmarcado en un entorno sin reproches que preserve la confianza y valore la contribución individual al aprendizaje colectivo y a la caracterización inicial de eventos [17].

4.1.3 Capa 3: protección ante mensajes no detectados

En virtud de que ningún filtro resulta infalible, esta fase procura minimizar el radio de explosión cuando un señuelo supera las barreras iniciales y el usuario interactúa con él, conjugando controles de identidad, navegación y endpoint.

En primer término, conviene endurecer la identidad con autenticación multifactor resistente al phishing (por ejemplo, mediante *FIDO2/passkeys*) y acceso condicional que eleve las exigencias conforme al riesgo, aplicando además el principio de mínimo privilegio para reducir las superficies de abuso sobre cuentas sensibles. Estos ajustes, recomendados para roles administrativos y usuarios de alto impacto, atenúan los secuestros de sesión y dificultan el uso fraudulento de credenciales [38].

Luego, es prudente proteger la navegación hacia destinos potencialmente maliciosos mediante servicios de *proxy*, inspección de URLs en tiempo de clic y clientes actualizados [17], de modo que un ingreso inadvertido no desemboque en una activación de código o una captura silenciosa de secretos. Lo anterior se complementa con la incorporación del *Protective Domain Name Service* (PDNS), el cual previene la consulta de dominios o direcciones IP conocidos por exponer material perjudicial [39].

Por último, las barreras de protección de los equipos frente al código malicioso se describen en la *Guía de seguridad de dispositivos* del NCSC [40], la cual brinda criterios normativos y parámetros de referencia para endurecer las estaciones de trabajo (computadores de escritorio y portátiles corporativos) y las terminales móviles. Dicha guía traduce las cláusulas generales en controles verificables y auditables.

En primer lugar, el documento explica la parametrización segura de las plataformas predominantes (Android, Windows, iOS, macOS y ChromeOS) tanto para flotas corporativas como para esquemas *bring your own device* (BYOD). En adición a las directrices por sistema, el NCSC suministra paquetes de configuración en su repositorio oficial de GitHub que sirven como línea base en soluciones *Mobile Device Management (MDM) / Unified Endpoint Managment (UEM)*, tal que las organizaciones instauran políticas coherentes desde el primer encendido y mantienen una postura homogénea durante todo el ciclo de vida del equipo [40].

En Android, se recomienda la administración centralizada mediante un servicio de MDM con capacidad para forzar el cumplimiento de controles técnicos, así como la adopción de un *Enterprise Mobility Management* compatible con OEMConfig, estándar que autoriza a los fabricantes a exponer ajustes avanzados a través de aplicaciones publicadas en Google Play y consumidas desde la consola corporativa. Igualmente, conviene habilitar *logging* y monitoreo, así como acoplarse a una arquitectura de red alineada con las opciones propuestas por el NCSC para ingreso remoto (desde esquemas tradicionales sustentados en VPN hasta diseños *zero trust*), reforzando autenticación y segmentación [40].

La política de dispositivo en Android debe abarcar, como mínimo, consideraciones sobre las interfaces externas y periféricos cableados o inalámbricos; biometría combinada con códigos de acceso y directrices de autenticación empresarial; gobierno de servicios en la nube de Google conforme a criterios internos; actualizaciones del sistema y del catálogo de aplicaciones a través de una opción automática; y alineación explícita entre la exigencia de contraseñas en el terminal y lo dispuesto en el manual corporativo [40].

En Windows, además de aquellos elementos que son transversales a los descritos para Android, Windows Autopilot habilita el *zero-touch enrollment* a través de una imagen de Windows confiable, la asignación inicial de cuentas sin privilegios locales y la eliminación de administradores innecesarios. En materia de productividad, se aconseja deshabilitar macros de Office [40]. No obstante, si resultan imprescindibles, deben ser restringidas a aplicaciones o usuarios concretos bajo criterios estrictos.

Las recomendaciones generales para Windows abarcan autenticación biométrica (Windows Hello for Business) y contraseñas acordes a la norma de la entidad; cifrado con BitLocker respaldado por un módulo de plataforma confiable (*Trusted Platform Modu-*

le, TPM) y un PIN para neutralizar la extracción de datos por medio de ataques físicos; protección de interfaces externas y defensa frente a intrusiones directas en memoria (*Direct Memory Access*, DMA) al impedir la enumeración de dispositivos cuando el equipo se halla bloqueado; actualizaciones automáticas del sistema, *firmware* y controladores mediante Windows Update for Business; control de ejecución con AppLocker para reducir superficie frente a *malware* y *ransomware*, disponer de un catálogo corporativo para aplicaciones de terceros distribuido por MDM a través de una tienda privada; resguardo de cuentas en la nube con acceso condicional aplicado a funcionalidades sensibles; y ajustes del firewall de Windows para impedir conexiones no deseadas en redes públicas o privadas [40].

El documento también recopila criterios previos al despliegue: elección de dispositivos (características de seguridad, soporte del proveedor y compatibilidad con MDM), adquisición con visión de cadena de suministro (evaluación de riesgos del fabricante y del canal), aprovisionamiento y distribución (procedimientos para entregar teléfonos, tabletas y portátiles en estado conocido) y *zero-touch enrollment* a fin de automatizar la incorporación sin intervención manual [40]. La sección dedicada a MDM expone la integración de las aplicaciones empresariales, las facultades nativas del sistema y los servicios de infraestructura que, en conjunto, facilitan el control remoto de los dispositivos, la visibilidad permanente de su estado y la definición coherente de salvaguardas de seguridad. Entre sus usos más frecuentes, se destacan la obligatoriedad de establecer conexiones a través de VPN, la ejecución de un borrado remoto en caso de extravío y la comprobación constante del cumplimiento normativo, que contempla desde la instalación de actualizaciones hasta la detección de *jailbreak* o *root* y la supervisión de desviaciones. El NSCS subraya, adicionalmente, un aspecto esencial: el entorno MDM constituye un recurso altamente sensible que exige barreras de contención robustas, dado que centraliza las credenciales y privilegios requeridos para desbloquear, descifrar o remover toda la flota de equipos [40].

En cuanto a las políticas, el NCSC aborda el uso de biometría, la selección y configuración de antivirus/*Endpoint Detection and Response* (EDR), la seguridad del navegador (parches, bloqueo de extensiones, manejo de contraseñas, etc.), los servicios en la nube integrados (sincronización de datos y ajustes, bloqueo remoto, etc.) y los periféricos (criterios de conexión segura y restricciones por contexto). Respecto a los recursos de terceros, se detallan estrategias para disminuir la probabilidad y el impacto: listas de permitidos/denegados, análisis formal, reputación del desarrollador, inspecciones en tiendas, política de soporte y actualizaciones, separación entre el espacio laboral y el personal, diseños de red que acotan la propagación de un ataque y control estricto de aplicaciones con alto privilegio. Asimismo, la guía provee criterios para plataformas de mensajería corporativa, examinando propiedades criptográficas, gestión administrativa y exposición del dato [40].

Por último, se describen las prácticas que sostienen la operación diaria y la vigencia de la postura defensiva [40]. Se enfatiza el borrado seguro de equipos como un lineamiento operativo indispensable para erradicar información confidencial o código hostil, así como en la planificación de ciclos de renovación y la implementación de instrumentos de registro y supervisión que sean capaces de emitir alertas tempranas ante patrones de actividad anómalos. A ello se suma el gobierno integral del firmware, sustentado en inventarios minuciosos, la aplicación de parches y la revisión de procedencia confiable. Finalmente, se alude a la comunicación de conductas aceptables para el uso de activos corporativos y el retiro ordenado de productos obsoletos, con el fin de mitigar las vulnerabilidades asociadas al software descontinuado [40].

4.1.4 Capa 4: respuesta expedita al incidente

El NCSC indica que se debe establecer y someter a pruebas un plan de respuesta que delimite las responsabilidades jurídicas y regulatorias. Además, sugiere favorecer el descubrimiento temprano, promoviendo el reporte inmediato de actividades sospechosas por parte del personal [17].

4.2 Detección y respuesta basadas en comportamiento

4.2.1 Herramientas tecnológicas y criterios de selección

Las plataformas *anti-phishing* impulsadas por IA constituyen una barrera de protección en tiempo real frente a correos y enlaces maliciosos, al combinar señales de reputación del remitente, análisis sintáctico-semántico del contenido, inspección dinámica de adjuntos y verificación de URLs en el momento del clic [41]. Soluciones consolidadas como Microsoft Defender for Office 365, Mimecast, Proofpoint o Barracuda Sentinel incorporan mecanismos de detección de suplantación de identidad e intentos de BEC —fraude por compromiso del correo corporativo— mediante modelos de aprendizaje automático (ML) que perfilan patrones de comunicación, relaciones habituales y anomalías en el flujo de mensajes [41].

Respecto a la detección avanzada, proveedores como SlashNext combinan procesamiento de lenguaje natural (NLP) —técnicas que permiten a los sistemas interpretar texto humano—, visión por computador —modelos que inspeccionan la apariencia visual de páginas—, y análisis de comportamiento dinámico de sitios para caracterizar amenazas desconocidas y superar tácticas de evasión [42].

Finalmente, este ecosistema ofrece funciones de investigación automatizada y respuesta a incidentes, como clasificación asistida de reportes de usuarios, retiro masivo de mensajes similares en todos los buzones y contención de cuentas o dispositivos comprometidos, integrándose con SIEM y SOAR para orquestar tareas de remediación con mínima intervención humana.

La Tabla 4.1 presenta una categorización de herramientas recomendadas para cubrir requerimientos clave en incidentes de ingeniería social, tanto del lado del usuario como del equipo de seguridad.

Tabla 4.1. Productos en el mercado para mitigar ataques de ingeniería social.

Categoría (capa)	Canal / vector	Capacidades clave	Ejemplos de producto
Seguridad de correo nativa	Email y colaboración.	Safe Links, *sandbox* de adjuntos y detección de intentos de suplantación de identidad (en usuario o dominio) [43]. Asimismo, filtrado y bloqueo con modelos de IA [44].	Microsoft Defender for Office 365; Google Workspace Security.
Seguridad por API / Gateway con IA	Email (BEC, apropiación de cuentas o *spear-phishing*).	Análisis dinámico, puntaje de riesgo adaptativo y detección basada en IA, inteligencia de amenazas y reglas de expertos [45]. Cuando es un caso de falso negativo, se puede subir un tiquete para reentrenar.	Abnormal; Proofpoint Nexus; Cloudflare Email Security.
Reporte del usuario y orquestación de respuesta	Email y suites colaborativas.	Botón integrado para reportar, tiquete automático y eliminación controlada de correos (*soft delete*).	Abnormal; Outlook Report; KnowBe4 (Phishing Alert Button); Hoxhunt; Proofpoint.
Aislamiento de navegador y pasarela web segura (*Remote Browser Isolation*)	Enlaces de email o web.	Renderizado remoto, aislamiento de links desde correo y desarme de contenido *(Content Disarm and Reconstruction)*, con el cual se eliminan macros, *scripts* y objetos incrustados en una versión limpia y funcional del archivo) [46].	Cloudflare Browser Isolation; Menlo Security; Zscaler.

Continúa en la página siguiente

Tabla 4.1. Productos en el mercado para mitigar ataques de ingeniería social. (Continuación)

Protección móvil *(Mobile Threat Defense)*	SMS, mensajería y QR.	Detección de amenazas *on-device* con ML, filtrado de URL y bloqueo de *phishing* o dominios maliciosos en tiempo real [47].	Lookout MTD; Zimperium; Jamf Protect/-Defense.
Autenticación de dominio y protección de marca	Correo y dominios.	DMARC/SPF/DKIM en estado *enforcement*, detección y retirada de dominios o URLs de suplantación.	Valimail; dmarcian; Fortra DMARC Protection; Netcraft.
Inteligencia de amenazas e investigación de URLs.	Multicanal	Análisis de reputación y *sandbox* para URLs sospechosas. Estas URLs, combinadas con artefactos de fraude a *endpoints* o indicadores de abuso de marca, forman parte de los IoC que se comparten y reciben en plataformas de CTI [48].	MISP; urlscan.io; VirusTotal; PhishTank.
Voz y centro de contacto.	Telefónico	Biometría de voz para autenticación, detección de *deepfakes* de voz y fraude telefónico.	Pindrop; NICE

4.2.2 Niveles de respuesta según riesgo calculado

Las políticas adaptativas sustentadas en una métrica de riesgo concreta habilitan respuestas adicionales ante autenticaciones sospechosas, examinando señales contextuales —por ejemplo, credenciales filtradas, uso de redes anónimas, desplazamientos atípicos en intervalos imposibles, dispositivos infectados con malware o propiedades de inicio de sesión no habituales— y aplicando controles proporcionales al nivel de exposición calculado. En Microsoft Entra ID Protection, dichos indicios alimentan dos políticas de acceso (riesgo del inicio de sesión y riesgo del usuario) con umbrales graduados (bajo, medio o alto), los cuales disparan salvaguardas como la imposición de MFA o cambio de contraseña obligatorio, así como bloqueos selectivos, cuando la probabilidad de compromiso supera el umbral definido [49]. Tal lógica se parametriza mediante *Conditional Access* y permite que la respuesta varíe si el comportamiento se desvía del patrón histórico del sujeto, minimizando la fricción para actividades legítimas y endureciendo el control cuando la evidencia lo exige [50].

Las cuentas con privilegios elevados representan objetivos preferentes para el esca-

lamiento de privilegios y el abuso de credenciales [10]. Por tal motivo, demandan una observabilidad reforzada y una activación *just-in-time* complementada con procesos de aprobación y auditoría continuos. Microsoft Entra Privileged Identity Management (PIM) y las soluciones de *Privileged Access Management* (PAM) aportan gobernanza sobre la concesión temporal de roles; la supervisión del uso de permisos, inspeccionando si son excesivos o innecesarios; y la revocación automática tras la tarea, reduciendo la ventana de exposición asociada a entidades administrativas [51].

En suma a lo anterior, la detección asistida por IA conforma un mecanismo de defensa inteligente que trasciende las reglas estáticas. En particular, tal instrumento brinda una protección dinámica y consciente del contexto, a través de la cual es factible ajustar la exigencia de factores adicionales de autenticación, activar el bloqueo de sesiones anómalas o imponer medidas de contención en tiempo real, sin perturbar flujos de bajo riesgo [49]. En un banco central, el resultado es una priorización eficaz del trabajo analítico, la reducción de falsos positivos y la protección reforzada sobre los usuarios sensibles, con métricas técnicas más favorables.

4.3 Estándares documentales mínimos para incidentes de phishing

La documentación estandarizada de incidentes, anclada en MITRE ATT&CK y acompañada de una categorización de la gravedad, contribuye no solo al archivo histórico, sino que constituye un habilitador crítico de inteligencia de amenazas accionable y de asignación eficiente de recursos. En el ámbito de los bancos centrales, este enfoque promueve una priorización rápida y consistente de la respuesta, afinando la planificación defensiva a escala institucional. Al exigir el etiquetado conforme a ATT&CK para eventos de phishing, las entidades convergen en un vocabulario común y neutro sobre el comportamiento adversario (tácticas, técnicas y subtécnicas) que reduce la ambigüedad analítica, habilita la caracterización de tendencias y acelera la comprensión de cómo se ejecutó el ataque y con qué medios concretos [37]. Cuando dicha taxonomía se compagina con esquemas de ponderación de severidad, cada registro adquiere un contexto operativo inmediato para activar niveles de servicio, escalamientos y uso racional de capacidades [52]. Además, dicho etiquetado viaja de forma íntegra a través de formatos y protocolos de intercambio (STIX/TAXII) y de plataformas colaborativas como MISP, convirtiendo descripciones humanas en artefactos legibles por una máquina que alimentan SIEM/SOAR y comunidades del foro sin pérdida de semántica [48].

MITRE ATT&CK T1566 (phishing) describe una familia de técnicas de acceso inicial fundamentadas en la persuasión electrónica (correo, mensajería, servicios de terceros o

llamadas), con las cuales el adversario induce a la víctima a inaugurar una vía de compromiso al revelar credenciales o iniciar código malicioso. Puntualmente, la técnica T1566 se desagrega en cuatro variantes que precisan el vector concreto [37]:

- **T1566.001 *Spearphishing Attachment***: archivos adjuntos maliciosos enviados a destinatarios concretos, como ficheros ejecutables o documentos con macros.
- **T1566.002 *Spearphishing Link***: enlaces que, mediante persuasión e interacción con el usuario, capturan datos o inducen descargas de malware.
- **T1566.003 *Spearphishing via Service***: contacto con la víctima a través de servicios de terceros como redes sociales o servicios de nube, sorteando así los controles del correo empresarial.
- **T1566.004 *Spearphishing Voice*:** el atacante emplea llamadas u otras comunicaciones de voz para manipular a la víctima y forzar la entrega de códigos, acciones de aprobación, instalación de software, etc.

Respecto a las categorías de severidad [52]:

- **Crítica (nivel 1)**: Interrupción significativa de operaciones esenciales, pérdida masiva de datos, impacto financiero severo o daño reputacional catastrófico. Requiere acción inmediata y movilización de la totalidad de los recursos.
- **Alta (nivel 2)**: Interrupción importante de funciones específicas del negocio, pérdida financiera significativa o tiempo de inactividad operativo considerable. Requiere acción rápida y prioritaria.
- **Moderada (nivel 3)**: Interrupción notable de operaciones, inconvenientes para un número moderado de usuarios o con impacto financiero limitado. Requiere atención y resolución planificada.
- **Baja (nivel 4)**: Interrupciones menores o inconvenientes que no afectan funciones críticas ni causan pérdidas sustanciales. Puede resolverse en el marco de las operaciones diarias.
- **Informativa (nivel 5)**: Sin impacto inmediato en los procesos, utilizada para fines de monitoreo o inteligencia de amenazas.

La Tabla 4.2 describe algunos eventos de ingeniería social relevantes, mapeándolos a MITRE ATT&CK T1566 y sus subtécnicas (adjunto, enlace, servicio de terceros y voz), a escenarios del ecosistema financiero y a criterios correspondientes a la escala de severidad.

Tabla 4.2. Etiquetado de incidentes de *phishing* según MITRE.

Etiqueta ATT&CK	Ejemplos	Asignación de severidad y condiciones orientativas
T1566.001 – *Spearphishing Attachment*	Envío a áreas de pagos o tesorería mensajes con instrucciones urgentes que contienen planillas o PDFs manipulados.	**Crítica**: ejecución confirmada en puestos con privilegios o conector SWIFT/ISO 20022; credenciales o llaves expuestas. **Alta**: descarga o ejecución en estaciones de operación con contención oportuna. **Moderada**: apertura del archivo sin activación o cuando esta es mitigada por EDR. **Baja**: adjunto bloqueado por *gateway/sandbox*. **Informativa**: señuelo detectado en campaña externa sin entrega.
T1566.002 – *Spearphishing Link*	Links a portales falsos de SSO, OWA, VPN o portal de liquidador con páginas clonadas.	**Crítica**: robo de credenciales privilegiadas y sesión activa; movimiento lateral detectado. **Alta**: ingreso de un usuario de alto valor en portal fraudulento sin MFA efectivo. **Moderada**: clic con reporte inmediato y restablecimiento sin evidencia de exfiltración. **Baja**: clic aislado en usuario no crítico, sin autenticación ni descarga. **Informativa**: URL maliciosa bloqueada por Safe Links*/proxy* antes del clic.
T1566.003 – *Spearphishing via Service*	Notificaciones ilegítimas desde supuestos proveedores a responsables de pagos.	**Crítica**: cambios de cuenta beneficiaria o autorización de pago fraudulenta (*Vendor Email Compromise* o VEC) completados. **Alta**: intento de modificación bancaria detenido posaprobación inicial. **Moderada**: solicitud de actualización de datos detectada y revertida. **Baja**: invitación sospechosa anulada; sin interacción sensible. **Informativa**: plantilla de estafa mapeado en CTI sin impacto.
T1566.004 – *Spearphishing Voice (vishing)*	Suplantación de directivo o regulador solicitando *one-time passcodes* o aprobación de transferencia urgente.	**Crítica**: divulgación de OTP/token o autorización de operación crítica; acceso remoto concedido. **Alta**: entrega de datos internos no públicos con riesgo relevante. **Moderada**: interacción prolongada sin cesión de secretos, pero con exposición de procesos. **Baja**: intento breve superado por el usuario sin verificación . **Informativa**: número/cadena identificado en listas de vigilancia; sin contacto.

5 Gestión del factor humano durante incidentes de ingeniería social

5.1 Identificación y entrenamiento del personal clave

La capacitación equilibrada en destrezas técnicas y competencias interpersonales [53], sumada a la participación temprana de RR. HH., Asuntos Legales y Comunicaciones, evidencia una comprensión madura de que un incidente cibernético trasciende el plano puramente tecnológico. En el ámbito de los bancos centrales, una respuesta verdaderamente exhaustiva atiende no solo la brecha operativa, sino también las derivaciones jurídicas, reputacionales y humanas: cumplimiento normativo y eventuales litigios; bienestar del personal ante situaciones de estrés, confusión o presión; y conducción de narrativas internas y externas con mensajes consistentes y temporalidad adecuada ante los *stakeholders* oportunos.

Cuando se alude a la ingeniería social, el detonante se origina en las personas —persuasión, urgencia fabricada, suplantación, entre otros— y sus efectos alcanzan la identidad, la confianza pública y el clima laboral. En consecuencia, la arquitectura de respuesta debe considerar asesoría jurídica para delimitar responsabilidades, recursos humanos para abordar la praxis u ofrecer apoyo psicosocial y comunicaciones para coordinar vocerías, alinear mensajes y asegurar la trazabilidad documental [54].

Conviene subrayar, además, la distinción con aquellos episodios originados por adversarios altamente sofisticados que operan sin interacción humana (automatización de explotación, cadenas de malware autónomas, abuso de integraciones con terceros o campañas de botnets) y que, por su naturaleza, permanecen confinados al dominio tecnológico. En tales escenarios, la respuesta se concentra en la contención, erradicación, parcheo, telemetría de alto volumen y restauración de servicios dentro de entornos y plataformas. Así, el componente humano participa principalmente como operador de procedimientos técnicos y custodio de evidencias.

Por el contrario, en incidentes de ingeniería social, la variable crítica es el comportamiento: la percepción del riesgo, la toma de decisiones bajo presión, el uso de canales

de reporte no punitivos y la recolección de lecciones aprendidas que retroalimenten la cultura y la gobernanza. De ahí que un equipo multidisciplinario, con líneas de decisión claras y protocolos de escalamiento, resulte imprescindible para mitigar impactos sistémicos y preservar la legitimidad institucional, al tiempo que los especialistas en ciberseguridad sostienen la defensa técnica y documentan los hechos con rigor.

En consonancia con lo anterior, la labor del equipo de respuesta es esencial para detectar, investigar y atender ciberamenazas sin interrumpir la continuidad operativa. A continuación, se describen los roles del equipo de respuesta a incidentes (*Incident Response Tean*, IRT) que intervienen durante una crisis organizacional y qué destrezas específicas deberían dominar [4] [53] [55].

- **Líder del IRT**: dirige la estrategia y la coordinación interáreas, prioriza objetivos, decide interrupciones controladas y fija los hitos del ciclo —contención, erradicación y recuperación—. Requiere juicio situacional, templanza bajo presión y conocimiento de guiones operativos.
- **Investigador principal/analista de seguridad**: recaba y examina material probatorio, formula y contrasta hipótesis y dirige el diagnóstico de causas raíz. En ese sentido, domina SIEM (*Security Information and Event Management*), IDS/IPS (*Intrusion Detection/Prevention Systems*) y EDR en puestos de trabajo; correlación de eventos; informática forense; sandboxing; scripting y automatización, de modo que despliega medidas de recuperación con fricción mínima.
- **Líder de comunicaciones**: segmenta audiencias, redacta mensajes coherentes y cronometra anuncios. Además, armoniza vocerías internas/externas y mantiene trazabilidad documental de lo comunicado.
- **Responsable de documentación y cronología**: construye la línea de tiempo única, registra decisiones y resultados, y custodia referencias, evidencia y artefactos para auditoría.
- **Representantes de RR.HH. y legales**: encauzan el tratamiento de datos personales, valoran la exposición regulatoria, explican las decisiones disciplinarias y llevan a cabo las notificaciones obligatorias cuando proceda.
- **Otros perfiles (TI, cumplimiento y terceros)**: aportan conocimiento del entorno, validan los controles, ejecutan las tareas de restauración y respaldan la revisión posterior.

La distribución de responsabilidades cobra materialidad operativa al mapear cada rol sobre la tabla RACI establecida en el Capítulo 1. El Líder del IRT asume típicamente el aprobador (A) durante la emergencia, mientras que el SOC o el investigador principal actúan como responsables (R) en triaje, contención y análisis. Por su lado, el personal de legal, RR. HH., comunicaciones y auditoría intervienen como consultados (C) en de-

cisiones con impacto jurídico, laboral o reputacional, al tiempo que la alta dirección y los jefes de área son informados (I) mediante resúmenes ejecutivos y marcadores de avance. Esta alineación con RACI clarifica el escalamiento, delimita las aprobaciones apremiantes (por ejemplo, para la retirada masiva de correos o para la revocación de sesiones) y reduce las ambigüedades durante el relevo de turnos.

5.2 Integración operativa de equipos multidisciplinarios

La concurrencia de disciplinas influye sobre la contención técnica, la gobernanza, el cumplimiento regulatorio, la ética del cuidado y la comunicación estratégica. Por consiguiente, para coordinar áreas internas en respuestas cohesionadas, se deben contemplar los siguientes elementos [55]:

- **Roles y responsabilidades claros**. Antes de cualquier contingencia, es perentorio definir, documentar y difundir atribuciones, límites de decisión, suplencias y rutas de transferencia [53]. Lo anterior en armonía con la matriz RACI previamente aprobada, de modo que la titularidad y las labores no queden en disputa cuando el tiempo apremia.
- **Patrocinio ejecutivo y vínculo interdepartamental**. La gestión operativa recae en TI, mientras que los representantes de cada unidad de negocio principal —en particular, jurídico, cumplimiento y RR.HH.— aportan visión institucional, criterio normativo y resguardo del capital humano.
- **Canales de comunicación preparados de antemano**. Disponer de directorios con teléfonos y puntos de contacto elimina dependencias frágiles durante una contingencia. Asimismo, contemplar interlocutores externos —fuerzas del orden, medios, ISAC (*Information Sharing and Analysis Centers*) o proveedores críticos— acorta tiempos y reduce incertidumbres [56].
- **ChatOps y sala de crisis (virtual o presencial)**. Un canal único por incidente, hilos por cada línea de acción, registro automático de hitos y reglas de relevo evitan conversaciones paralelas y pérdida de contexto.
- **Escenarios de simulación y ejercicios integrales**. Diseñar ejercicios de mesa con casos verosímiles permite corroborar la pertinencia de los playbooks, supervisar los tiempos de reacción y afinar vocerías. Resulta crucial realizar simulacros de brechas para todo el equipo, con énfasis en comunicación de crisis [56].
- **Interoperabilidad de herramientas de seguridad**. Favorecer las integraciones entre SIEM, IDS/IPS, EDR, casillas de reporte y plataformas de *ticketing*, conservando métricas desde la alerta hasta el cierre y habilitando trazabilidad rigurosa.

- **Controles técnicos de base**. Segmentación, parcheo con ventanas definidas, MFA resistente a phishing y políticas restrictivas de macros acotan el radio de impacto y agilizan la contención tras la entrega.
- **Estructura de mando unificada**. Conviene fijar la taxonomía de severidades [53], los criterios de priorización y las funciones de cada célula, desde la investigación forense hasta la comunicación ejecutiva, con participación de jurídico y RR.HH.
- **Privacidad, ámbito laboral y cumplimiento**. La interacción con las áreas jurídicas y la figura del delegado de protección de datos regula acceso a información personal [53], preservación probatoria y notificaciones obligatorias, conciliando diligencia operativa con restricciones normativas.

La insistencia en roles debidamente delineados, canales de comunicación establecidos y simulacros interfuncionales evidencia que la respuesta ante incidentes en bancos centrales obedece menos a la improvisación y más a una preparación concertada. La planificación minuciosa y la práctica periódica de acciones multidisciplinares convierten un episodio potencialmente caótico en un esfuerzo estructurado y coherente. Al trazar compromisos, estandarizar protocolos de intercambio de información y conservar copias de manuales y *runbooks* con anterioridad [55], la organización reduce de forma notable los tiempos de reacción y limita errores bajo presión. Procediendo así, cada parte interesada conoce con precisión su función y cómo llevarla a cabo, impidiendo la aparición de malentendidos y demoras que podrían amplificar el impacto sobre los mercados financieros o la confianza pública.

5.2.1 Comunicación eficaz bajo presión

Una comunicación en contingencias exige una arquitectura informativa robusta, disciplina procedimental y previsión. Antes de que se materialice un incidente, conviene construir un mapa de audiencias con niveles de exposición, necesidades de información, canales formales y ventanas temporales predefinidos para cada flujo [54]: áreas operativas que requieren instrucciones inmediatas; alta dirección que demanda panoramas ejecutivos con cursos de acción; y funciones de control que precisan hechos verificados y trazabilidad. Tal segmentación, traducida en mensajes nítidos y adaptados a cada equipo, disminuye desacoples e interpretaciones erróneas cuando la presión se incrementa.

Sobre esta base, algunas de las estrategias clave son [54] [56] :

- Un comité de comunicación de crisis con representación de ciberseguridad, asesoría jurídica, liderazgo institucional y relaciones públicas elabora reglas de redacción y aprobación acelerada a través de glosarios comunes para nombrar

eventos sin ambigüedad y de listas de verificación que imponen mínimos no negociables (hechos confirmados, alcance, medidas adoptadas, siguiente actualización, etc.).

- La temporalidad importa tanto como el núcleo temático. Es indispensable pautar hitos: comunicado inicial con hechos corroborados y compromisos de investigación, ciclos regulares de actualización con horario fijo —aunque no haya novedades sustantivas, se confirma continuidad del trabajo— y cierre provisional con lecciones preliminares y próximos pasos.
- La coherencia se preserva con una única fuente de verdad, la cual garantice una mensajería congruente y coordinada en todas las plataformas.
- El reconocimiento de responsabilidad cuando procede, pero sin especular sobre causas todavía en análisis, preserva credibilidad y reduce conflictos posteriores con supervisores y opinión pública.
- La protección de la información exige clasificación previa de lo que puede divulgarse y lo que debe reservarse —por confidencialidad, cadena de custodia o riesgo operativo—, y filtros de revisión que prevengan deslices en datos personales, indicadores técnicos sensibles o detalles forenses aún inmaduros.
- Incorporar la comunicación en crisis dentro de los simulacros de brechas dirigidos a todo el equipo, de modo que cada integrante interiorice los protocolos, los estilos de mensaje y los criterios de coordinación.
- Medición y retroalimentación a través de tasas de lectura por colectivo, puntualidad en los envíos, consistencia entre canales, volumen de consultas derivadas, contención de rumores y tiempos de latencia entre hitos.

Esta sección se enlaza de manera orgánica con el plan de gestión de incidentes y resiliencia cibernética, subrayando cómo la comunicación de crisis vinculada al factor humano se integra en el ciclo integral de respuesta. En este terreno, la comunicación no se reduce a transmitir datos, sino que constituye un instrumento estratégico de preservación de la confianza institucional y de contención del daño reputacional, aspectos que resultan fundamentales para los bancos centrales. Los ataques de ingeniería social explotan precisamente la confianza como vector de entrada y, cuando la narrativa se conduce de forma deficiente, pueden erosionarla en cuestión de horas, con consecuencias que van desde la inestabilidad de los mercados hasta episodios de pánico social y deterioro de la credibilidad institucional. En este sentido, la comunicación de crisis se convierte en una pieza clave, al otorgar claridad, solidez y celeridad en los mensajes.

5.2.2 ChatOps y sala de crisis

Comunicación operativa disciplinada. Ante contingencias resulta clave una arquitectura comunicacional rigurosa. La práctica de ChatOps, extendida en DevOps y respuesta a incidentes, concentra la colaboración en un canal único con todas las partes pertinentes, conformando una sala de crisis virtual con intercambio inmediato. Conviene crear un espacio dedicado por cada evento (`incidente-XYZ`) y menejar allí toda la conversación, evitando la dispersión en hilos privados o correo [57]. En la práctica, organizaciones como Shopify generan de forma automática un canal en Slack al declararse el incidente, de modo que todos los intervinientes comparten contexto, cronología y maniobras en tiempo real [58]. Así, previene la duplicación de esfuerzos y la pérdida de información crítica.

Canal maestro e hilos por línea de acción. Una convención recomendable consiste en mantener un espacio principal para la colaboración general y, en paralelo, hilos para subtareas [59]. Si coinciden, por ejemplo, una investigación de malware y comunicaciones a clientes, conviene abrir hilos separados dentro del mismo entorno para cada frente de trabajo. Así, el tablero central conserva legibilidad (anuncios y decisiones clave), mientras el detalle técnico ocurre en conversaciones subordinadas. Al cierre, un resumen ejecutivo de cada hilo se publica en el espacio maestro para registro [58]. Además, las plataformas de chat registran automáticamente interacciones y eventos, creando un histórico cronológico valioso para auditoría y *post mortem*. Resulta prudente, por tanto, parametrizar integraciones que conserven esta bitácora [59]. De hecho, numerosas instituciones archivan dichos registros en sus sistemas de gestión de servicios de TI o como tiquetes para consulta futura.

Registro automático de eventos. La documentación en tiempo real durante una contingencia es crucial. En entornos con ChatOps, las integraciones pueden anotar de forma automática acciones operativas en el canal central: «se ejecutó un *rollback* en el servidor X» o «alerta contenida por el firewall» [58]. Ello no solo informa de inmediato a toda el equipo de respuesta, sino que también construye una bitácora cronológica de pasos, decisiones y responsables. Al concluir el episodio, tales datos se compilan en un informe formal. Esta trazabilidad automatizada reduce la carga del coordinador, preserva evidencia con valor pericial y ofrece material verificable ante auditorías o requerimientos regulatorios.

Reglas de relevo (handover). En incidentes prolongados (interrupciones severas o campañas sostenidas), el cambio de turno exige protocolos rigurosos que aseguren transferencia íntegra de contexto sin pérdidas de información, lo cual se ve favorecido por ChatOps [60]. Lo idóneo consiste en realizar un relevo sincrónico (presencial o por videoconferencia) entre el equipo saliente y el entrante, repasando estado actual, iniciativas pendientes, riesgos abiertos y recomendaciones operativas. Complementariamen-

te, resulta conveniente fijar en el canal un resumen escrito de *handover* para referencia posterior [58]. El objetivo es mantener la continuidad de la respuesta sin duplicidades ni omisiones. Asimismo, es recomendable disponer de un formato normalizado (*checklist* de traspaso) que ambas partes completen y validen.

Roles y moderación de la sala: En un *war room* administrado de manera apropiada, suelen asignarse roles explícitos al equipo de respuesta. Siguiendo el modelo del Incident Command System (ICS) adaptado a TI, muchas empresas asignan un comandante del incidente que modera la sala de crisis y dirige la respuesta [58]. Concretamente, se asegura de que se sigan las reglas de comunicación instauradas, impone orden cuando varias personas participan a la vez y garantiza que todas las decisiones queden claras para el grupo.

En concordancia con lo anterior, un manual de sala de crisis para la institución financiera debe contemplar: cómo crear y nombrar canales de incidente, el formato de comunicación (uso de hilos, códigos o etiquetas de prioridad), los deberes que asume cada rol en la crisis y los procedimientos de escalamiento y relevo.

5.3 Privacidad laboral y comunicación regulatoria

En instituciones financieras, la administración de incidentes o investigaciones internas exige un delicado equilibrio entre la protección del patrimonio de información y los derechos fundamentales de empleados y clientes, el cual se ciña a las obligaciones impuestas por las autoridades supervisoras. Este marco requiere una arquitectura normativa que armonice seguridad, dignidad y proporcionalidad, junto con procedimientos de aprobación ágiles y trazables. En términos jurídicos, la inspección de cuentas corporativas y recursos asignados al personal (correo electrónico, estaciones de trabajo o aplicaciones de negocio) solo procede cuando concurren finalidades legítimas vinculadas al desempeño profesional y siempre bajo límites claros: conocimiento previo por parte del trabajador, criterios de mínima intrusión, alcance acotado a la sospecha razonable y respeto estricto por la intimidad.

Así lo refleja el control del correo corporativo en Europa y América Latina, el cual se rige por un marco que equilibra interés empresarial y derechos fundamentales: el artículo 8 del CEDH y la doctrina del TEDH (caso Bărbulescu vs. Rumanía) protegen la vida privada y la correspondencia y las guías de EDPS desaconsejan lecturas sistemáticas del contenido sin motivo legítimo [61], mientras que, en España, el ET artículo 20.3 y la jurisprudencia del Tribunal Supremo fijan criterios de idoneidad, necesidad y proporcionalidad estricta [62]. En Latinoamérica, la Ley Federal del Trabajo en México permite que el empleador revise correos o mensajes que ha provisionado si existe un

propósito válido (cumplimiento de obligaciones laborales, prevención de uso indebido, protección de información, entre otros) [63].

En la práctica, lo anterior significa:

- **Política interna transparente y consentimiento informado**. La organización debe contar con una directriz inequívoca sobre el empleo de medios digitales corporativos y sobre la posibilidad de monitorización de tales recursos [61]. El personal ha de ser informado de manera expresa y, preferentemente, prestar su conformidad por escrito, de modo que desaparezca la expectativa de privacidad en esos entornos [61]. De tal forma, al prohibir el uso personal del correo institucional, se refuerza la legitimidad de inspeccionarlo conforme a la política vigente [61].
- **Inviolabilidad de comunicaciones personales**. No procede acceder a mensajes ni a cuentas privadas del trabajador. En otras palabras, la entidad puede revisar el correo corporativo —propiedad de la empresa cedida temporalmente al empleado—, pero no una cuenta personal (por ejemplo, Gmail), aun cuando se haya ingresado desde un equipo corporativo [62]. En Europa y Latinoamérica, la doctrina judicial declara que la persona conserva un ámbito de privacidad sobre las comunicaciones personales en equipos de la empresa, salvo información previa y reglas internas explícitas que señalen lo contrario [62].
- **Condiciones para una inspección lícita**. Cualquier supervisión o inspección del tráfico del empleado debe responder a un motivo legítimo y determinado —por ejemplo, sospecha fundada de filtración de datos, fraude interno, incumplimiento grave— y resultar proporcionado en alcance y duración [64]. Queda excluida la vigilancia generalizada sin justificación. Ante indicios razonables, conviene recabar aprobación previa de la dirección y del área jurídica/privacidad antes de revisar correos o chats. Asimismo, es recomendable —y, en algunas jurisdicciones, obligatorio— otorgar garantías de transparencia, tales como realizar la revisión en presencia del afectado o de un tercero imparcial (notario o representante legal) que de fe del procedimiento [64]. Esta cautela protege los derechos del trabajador y robustece la validez probatoria de la evidencia obtenida.
- **Tratamiento de datos personales del personal**: En entornos financieros, este manejo exige un andamiaje normativo y operativo que combine legalidad, proporcionalidad y transparencia. Conforme al GDPR, las actuaciones han de regirse por los principios de finalidad, minimización de datos y limitación del periodo de conservación. En suma a ello, en España, la Ley Orgánica 3 (2018) reconoce derechos digitales del trabajador, entre ellos, intimidad en el uso de medios corporativos, desconexión y límites a la videovigilancia o geolocalización [65].

  Antes de consultar correos, historiales o logs vinculados a personas identificadas, resulta recomendable un flujo de aprobación con validación escrita de aseso-

ría jurídica y del DPD, acotando alcance, finalidad, categorías consultadas, plazos de retención y, cuando proceda, técnicas de anonimización o seudonimización. En escenarios de riesgo elevado, conviene realizar una evaluación de impacto para la protección de datos (EIPD) siguiendo guías de la autoridad de control [66].

5.3.1 Notificaciones obligatorias

Cuando ocurre una brecha de datos personales —por ejemplo, exposición inadvertida de información de clientes o empleados— el GDPR exige comunicarla a la autoridad de control pertinente (como la AEPD) en un plazo máximo de 72 horas luego de tener conocimiento del incidente [66], salvo que sea improbable que el evento suponga un riesgo para los derechos y libertades de los individuos. Por el contrario, si la filtración implica perjuicios graves, la comunicación hacia los afectados debe producirse sin demora indebida [66], expresando claramente lo sucedido, las consecuencias probables y los pasos recomendados para mitigar impactos.

En virtud de lo anterior, el marco operativo interno debe introducir un protocolo de respuesta ante brechas a fin de sopesar si aplica notificación regulatoria y coordinar con los equipos jurídicos o de protección de datos la elaboración del reporte oficial que incluirá: resumen del incidente, categorías y número estimado de personas afectadas, naturaleza de los datos comprometidos y medidas correctivas tomadas [66].

En paralelo, la preservación probatoria adquiere un relieve singular en incidentes con vector humano, puesto que el itinerario de decisiones y la trazabilidad del intercambio comunicacional resultan determinantes. La recolección de evidencia digital (registros de sistemas, correos, artefactos adjuntos, capturas o terminales) ha de seguir técnicas forenses que garanticen integridad y autenticidad [64]. En particular, el procedimiento operativo debe describir con precisión: identificación de elementos relevantes, aislamiento del material comprometido, documentación exhaustiva de cada intervención y custodia bajo control. Si el tratamiento implica datos personales, corresponde la validación legal apropiada, asegurando que se lleve a cabo en la medida necesaria y conforme a los lineamientos vigentes [64].

6 Cultura y concientización frente a la ingeniería social

6.1 Estrategias para promover una cultura organizacional segura

Forjar una cultura de ciberseguridad robusta exige incorporar la seguridad de la información en el comportamiento cotidiano hasta volverla un hábito automático, no un añadido tardío. Para lograrlo, conviene articular mensajes claros, refuerzos accionales y retroalimentación periódica que conecten decisiones individuales con la exposición al riesgo institucional [67].

6.1.1 Campañas internas con enfoque conductual y segmentado

De acuerdo con los principios del NCSC para crear condiciones culturales favorables [67], las guías de ENISA sobre cambio en el comportamiento [68] y el ciclo de aprendizaje continuo descrito por NIST SP 800-50r1 alrededor de la concientización y la privacidad, se describen las siguientes consideraciones.

1. **Arquitectura de mensajes orientada a conducta (*behavioral design*)**:
    - Enmarcar cada pieza con el «por qué»: propósito, impacto personal (protección de nómina o reputación profesional) y beneficio colectivo.
    - Aplicar microrecordatorios contextuales (*nudges*): *banners* en el correo al detectar adjuntos sensibles, avisos en el navegador ante dominios dudosos y mensajes breves en aplicaciones de mensajería corporativa.
    - Preferir formatos cortos —infografías, tarjetas de 60–90 segundos o historias de incidentes «sanitizadas»— y llamados a la acción concretos.

2. **Segmentación por función y exposición al riesgo**
    - Tesorería/pagos, compras, jurídico/contratos, mesa de ayuda, desarrollo/-DevOps y directivos requieren énfasis distintos: BEC y confirmación de beneficiario en tesorería; gestión de proveedores en compras; e higiene de secretos en tecnología.
    - Plasmar las narrativas en el contexto latinoamericano en cuanto a idioma, referencias regulatorias y tipologías de fraude locales.
3. **Temario recomendado**: En concordancia con el listado en la Tabla 2.1.
4. **Canales y frecuencia de comunicación**
    - **Mensual**: Cápsulas temáticas por perfil.
    - **Trimestral**: Relato breve de lecciones aprendidas a partir de ejercicios o incidentes reales, con política de aprendizaje sin reproche.
    - **Eventos puntuales**: Fichas informativas ante olas de fraude (campañas BEC regionales o picos de smishing bancario), reforzando criterios de OOB.
5. **Transparencia y rendición de cuentas**: Informes regulares que compartan hallazgos de auditoría técnica y resultados de ejercicios de respuesta en lenguaje comprensible, con métricas de avance y próximos hitos.

6.1.2 Motivadores y reconocimiento de conductas seguras

- **Compromiso del liderazgo**: La alta dirección debe participar de forma activa, compartir aprendizajes derivados de incidentes y abrir conversaciones periódicas con los equipos, sentando un ejemplo positivo. En este sentido, el patrocinio ejecutivo es crucial y debe ser permanente [68].
- **Gamificación**: La incorporación de mecánicas lúdicas (retos breves, puntos, niveles y recompensas simbólicas) incrementa la participación y favorece la retención de conocimientos [23]. Asimismo, tablas de clasificación por áreas, insignias digitales y trofeos con foco en prácticas acertadas (por ejemplo, «reporte oportuno de phishing») refuerzan hábitos sin apelar a la sanción.
- **Incentivos y recompensas coherentes con el riesgo humano**: El reconocimiento periódico de dinámicas seguras fortalece normas sociales positivas [22] [26]. La integración de métricas de concientización en los sistemas internos de desempeño y promoción —por ejemplo, metas de reporte temprano, culminación de módulos o superación de evaluaciones— aporta consistencia con la gobernanza del riesgo humano.

- **Capacitación interactiva**: Los formatos pasivos reducen el aprendizaje significativo. En su lugar, conviene priorizar cuestionarios adaptativos, simulaciones breves, estudios de caso y guiones situacionales que reproduzcan fricciones reales con retroalimentación inmediata [23].
- **Compromiso continuo y microaprendizaje**: Secuencias cortas y frecuentes sostienen la memoria a largo plazo, mientras que demostraciones en vivo, espacios *brown bag* (charlas informales en horario de almuerzo) y seminarios web con preguntas abiertas fomentan la participación transversal [10].

La transición desde campañas informativas esporádicas hacia una cultura de seguridad arraigada exige resignificar la ciberseguridad: dejar de concebirla como carga de cumplimiento y entenderla como valor compartido, impulsado por apoyo directivo y reforzado mediante incentivos positivos. El objetivo consiste en que la protección se vuelva instintiva y de alto valor personal, no una imposición formal. En ese marco, el personal deja de actuar como receptor pasivo de contenidos para convertirse en codefensor activo, consolidando una mentalidad de «seguridad primero» [10].Tal resiliencia cultural resulta decisiva: incluso frente a ataques avanzados de ingeniería social, predomina un escepticismo informado y OOB, transformando el componente humano en barrera proactiva antes que en vulnerabilidad persistente.

6.2 Mecanismos para impulsar la adopción de prácticas de seguridad

De acuerdo con [69] y [70], se sugiere:

- **Urgencia estratégica**: Comunicar con evidencia las consecuencias de ignorar la cultura de ciberseguridad: filtraciones de datos, sanciones regulatorias, interrupciones operativas, deterioro reputacional y fuga de clientes.
- **Visión y hoja de ruta**: Definir una visión comprensible —por ejemplo, defensa en profundidad— y articular el por qué, qué y cómo para todo el personal, con metas, responsables y plazos comprobables.
- **Incorporación temprana**: Incluir criterios de seguridad desde la fase inicial del proyecto: valoración de riesgos, mapeo de partes interesadas, plan de comunicaciones y requisitos de cumplimiento.
- **Comités transfuncionales**: Constituir grupos de trabajo con representantes de gestión de cambio, riesgos, ciberseguridad, TI, jurídico y comunicaciones, orientados a alineación táctica y responsabilidad compartida.

- **Seguridad como habilitador**: Atender inquietudes del usuario final, depurar protocolos complejos y posicionar la ciberseguridad como facilitador de objetivos de negocio, no solo como restricción.
- **Refuerzo continuo**: Más que reiterar recursos didácticos, el programa debe optimizar el aprendizaje en el tiempo mediante repetición espaciada, que distribuye la práctica para contrarrestar la curva del olvido; práctica de recuperación, que fomenta recordar sin apoyo para consolidar memoria a largo plazo; y entremezcla de temas, que alterna familias de amenazas para mejorar la distinción.
- **Presupuesto segregado**: Asignar una partida exclusiva para seguridad de la información, independiente del gasto general de TI, con trazabilidad de inversiones y evaluación del retorno esperado.
- **Victorias tempranas**: Desplegar controles de alto impacto —por ejemplo, MFA priorizada, OOB para pagos críticos, botón de reporte en el cliente de correo— que demuestren una disminución de riesgo cuantificable en los primeros meses y generen tracción institucional.
- **Supervisión y mejora continua**: Migrar hacia un modelo de gobernanza con planes plurianuales, revisiones periódicas, hitos precisos y actualización iterativa del programa formativo.
- **Participación de RR.HH.**: Trabajar con recursos humanos en la selección, inducción, formación recurrente, expectativas de comportamiento y régimen disciplinario, garantizando armonía con el código ético y el marco normativo interno.

6.3 Encuestas y métricas para supervisar el clima organizacional en seguridad

La medición periódica del clima organizacional en seguridad demanda un diseño de encuestas riguroso que capture percepciones, actitudes y experiencias del personal respecto a prácticas, políticas y procedimientos [71]. Un cuestionario estructurado —por ejemplo, un instrumento inspirado en los 67 ítems de DVMS-CAT— con reactivos en escala de Likert (escala ordinal de acuerdo/desacuerdo) facilita estimar tendencias y contrastes entre áreas [27]. La confidencialidad y el anonimato resultan esenciales para respuestas francas [71], por ello se sugiere agrupar puntajes y suprimir metadatos identificadores. Además, la participación aumenta cuando se comunica con transparencia el propósito, se garantiza accesibilidad, se ofrecen incentivos prudentes y se retroalimenta a la plantilla con resultados; la invitación debe abarcar múltiples dependencias, no exclusivamente funciones de ciberseguridad[27]. Respecto a la interpretación, esta

debe compaginar enfoques cuantitativos (medias, frecuencias, correlaciones, comparaciones interdepartamentales y puntuaciones alineadas con factores de cultura positiva) y cualitativos (análisis temático de respuestas abiertas), a fin de sintetizar los hallazgos en informes ejecutivos de utilidad.

Aludiendo al marco DVMS-CAT, este examina la cultura de ciberseguridad a través de seis ejes derivados de la *culture web* de Johnson & Scholes: símbolos (señales del compromiso directivo), estructuras de poder (capacidad decisoria), arquitectura organizacional (silos tecnológicos y de negocio), sistemas de control (prácticas y procesos), hábitos y rutinas del personal, y relatos que moldean la memoria institucional [27]. Las afirmaciones tipo Likert indagan, entre otros aspectos, claridad del canal de reporte, apoyo de jefaturas durante incidentes, colaboración interdepartamental y valoración de simulacros periódicos [27]. En función de dichas respuestas, el cuestionario produce un mapa de riesgo cultural que dirige las intervenciones focalizadas y el seguimiento de mejoras.

La Tabla 6.1 presenta las técnicas y preguntas clave para este tipo de encuestas [27].

Tabla 6.1. Preguntas y técnicas para encuestas de clima en seguridad de la información.

Categoría de factor cultural	Pregunta	Escala de respuesta	Interpretación del riesgo/fortaleza
Símbolos (compromiso del liderazgo)	La alta dirección demuestra activamente la importancia de la ciberseguridad en sus comunicaciones y acciones.	1 (totalmente en desacuerdo) a 5 (totalmente de acuerdo)	Puntuaciones bajas indican falta de visibilidad y de compromiso del liderazgo, lo que debilita la cultura.
Estructuras de poder (capacidades de gestión)	Mi supervisor directo tiene el conocimiento y los recursos para apoyar las prácticas de ciberseguridad de mi equipo.		Puntuaciones bajas sugieren brechas en la capacidad de gestión para implementar y reforzar la seguridad.
Estructuras organizacionales (silos)	Se evidencia una colaboración efectiva entre mi departamento y el equipo de ciberseguridad.		Puntuaciones bajas revelan silos organizacionales que impiden una respuesta de seguridad coordinada.
Sistemas de control (prácticas/procesos)	Los procedimientos de seguridad son claros, fáciles de entender y aplicar en mi trabajo diario.		Puntuaciones bajas indican procedimientos complejos o poco claros, lo que lleva a la resistencia o el incumplimiento.
Historias (narrativas organizacionales)	La organización comparte regularmente las historias de éxito o las lecciones aprendidas relacionadas con la ciberseguridad.		Puntuaciones bajas sugieren una falta de refuerzo positivo o de aprendizaje apoyado en experiencias pasadas.
Hábitos y rutinas (personas)	Reporto consistentemente correos electrónicos o actividades sospechosas, incluso si no estoy seguro de que sean una amenaza.	1 (nunca) a 5 (siempre)	Puntuaciones altas demuestran la adopción de hábitos seguros, mientras que las bajas indican falta de proactividad.

Conclusión

El documento traza una arquitectura integral para el factor humano en ciberseguridad, donde la gobernanza deja de ser un añadido procedimental y pasa a constituir una aptitud estratégica. La autoevaluación de madurez periódica —vinculada a marcos como CERT-RMM—, la matriz RACI y un tablero de indicadores clave de rendimiento orientados al ámbito ejecutivo construyen un ciclo de dirección claro, con trazabilidad, umbrales y priorización de esfuerzos sustentada en datos.

Respecto a la formación y el entrenamiento, la propuesta migra desde campañas esporádicas hacia un currículo modular con microaprendizaje (cápsulas breves y frecuentes), donde el diseño se centra en el pensamiento crítico frente a la ingeniería social. La evaluación de efectividad combina conocimiento inmediato, retención a mediano o largo plazo y transferencia al puesto mediante situaciones realistas. Esta tríada pedagógica, complementada con retroalimentación ágil, reduce sesgos, mejora la respuesta cotidiana y facilita decisiones acertadas bajo presión.

En términos operativos, los simulacros no anunciados suministran indicios del desempeño real. La simulación de phishing, los ejercicios de red team (equipo ofensivo simulado) y purple team (colaboración entre ofensiva y defensa) ofrecen una lectura de la agilidad conductual, los tiempos de reacción y la capacidad de contención. Un conjunto de métricas coherentes —tasa de clic, reporte inicial, TTR del usuario, MTTD/MTTC del equipo azul, entre otros— refina límites, señala cuellos de botella y acelera la remediación técnica y procedimental.

Aludiendo al phishing en banca central, el enfoque por capas estructura las barreras preventivas, la educación enfocada en el reporte temprano, la protección residual ante mensajes que evaden filtros y la respuesta expedita ante incidentes. A su vez, la detección adaptada al comportamiento en tiempo real, la instrumentación tecnológica con criterios claros y la documentación mínima exigible —guías, plantillas, bitácoras y cronologías— reducen la ambigüedad y favorecen la toma de decisiones consistentes en ventanas críticas.

Durante incidentes, la gestión del componente humano exige identificar al personal clave, integrar equipos multidisciplinares, y mantener una comunicación eficaz con au-

diencias internas y externas. La orquestación mediante ChatOps (canal único de coordinación) y salas de crisis (virtuales o presenciales), junto con lineamientos de privacidad, reglas laborales y notificaciones obligatorias, preserva los protocolos de acción, la recopilación automática de evidencias y la coherencia narrativa ante los supervisores y la sociedad.

En cuanto a la cultura organizacional, se evoluciona desde la concientización declarativa hacia prácticas internalizadas: campañas segmentadas, gamificación anclada en conductas críticas, reconocimiento de hábitos seguros y mecanismos para impulsar cambios sostenidos. Las encuestas de clima y los instrumentos psicométricos ofrecen lecturas de símbolos, rutinas y percepciones, con rutas de mejora que trascienden la capacitación y ahondan en procesos, incentivos y liderazgo.

En conjunto, la hoja de ruta resultante configura un sistema vivo que enlaza la dirección estratégica, la pedagogía basada en evidencia, el ensayo operativo y el aprendizaje institucional. El vector humano deja de representar un flanco débil para convertirse en una barrera activa, gracias a métricas comparables en el tiempo, procesos ensayados y una narrativa institucional que prioriza el escepticismo informado, la confirmación por canal independiente y el reporte ágil. Procediendo así, los bancos centrales elevan su resiliencia frente a campañas sofisticadas, reducen la superficie de ataque y fortalecen la confianza pública en infraestructuras esenciales.

Referencias

[1] Nojus Bendoraitis. *Free cybersecurity maturity assessment questionnaire for evaluating your security posture*. Copla. 2025. URL: https://copla.com/blog/cybersecurity/free-cybersecurity-maturity-assessment-questionnaire-for-evaluating-your-security-posture/ (vid. pág. 3).

[2] *Cyber Resilience Review (CRR) Question Set with Guidance*. U.S. Department of Homeland Security Cybersecurity e Infrastructure Security Agency (CISA), 2020. URL: https://www.cisa.gov/sites/default/files/c3vp/csc-crr-question-set-and-guidance.pdf (vid. pág. 3).

[3] Randolph S. Sergent y David Sharp. *Cybersecurity Preparedness*. Maryland Health Care Comission. 2025. URL: https://mhcc.maryland.gov/mhcc/pages/hit/hit_cybersecurity/documents/Cybersecurity_Self-Assessment_Tool.pdf (vid. pág. 3).

[4] Richard A. Caralli, Julia H. Allen, David W. White, Lisa R. Young, Nader Mehravari y Pamela D. Curtis. *CERT® Resilience Management Model, Version 1.2*. Carnegie Mellon University's Software Engineering Institute, 2016. URL: https://www.sei.cmu.edu/documents/1629/CERT_Resilience_Management_Model_Version_1_2.pdf (vid. págs. 8, 49).

[5] *RACI Matrix For Cybersecurity*. Meegle. 2025. URL: https://www.meegle.com/en_us/topics/raci-matrix/raci-matrix-for-cybersecurity (vid. pág. 11).

[6] Jon Waldman. *Building a Strong Cybersecurity Culture Within Your Organization*. SBS Cyber Security. 2025. URL: https://sbscyber.com/blog/cybersecurity-culture (vid. pág. 11).

[7] *How Executives Can Use Cybersecurity KPIs to Make Informed Decisions*. Peris.ai. 2025. URL: https://peris.ai/post/how-executives-can-use-cybersecurity-kpis-to-make-informed-decisions (vid. pág. 14).

[8] Tyas Tunggal y Kaushik Sen. *Top Cybersecurity Metrics and KPIs for 2025*. UpGuard. 2025. URL: https://www.upguard.com/blog/cybersecurity-metrics (vid. pág. 14).

[9] Geoff Hancock. *Cybersecurity Metrics And KPIs CISOs Use To Prove Business Value*. PurpleSec. 2025. URL: https://purplesec.us/learn/cybersecurity-metrics-kpis/ (vid. pág. 14).

[10] Max Shier. *Building a Strong Cyber-Aware Culture: Advanced Strategies for Cybersecurity Awareness Training*. Optiv. 2024. URL: https://www.optiv.com/insights/discover/blog/building-strong-cyber-aware-culture-advanced-strategies-cybersecurity (vid. págs. 17, 22-24, 45, 59).

[11] Tahir. *What is Phishing?: Attacks, Types, Defenses, and the Impact of AI*. 2025. URL: https://medium.com/@tahirbalarabe2/what-is-phishing-attacks-types-defenses-and-the-impact-of-ai-dec3b2c8d125 (vid. págs. 17, 18).

[12] Srini Tummalapenta. *How a new wave of deepfake-driven cyber crime targets businesses*. IBM. n.d. URL: https://www.ibm.com/think/insights/new-wave-deepfake-cybercrime (vid. pág. 17).

[13] Cory Marchand. *The Psychology of Social Engineering*. Coalition. 2025. URL: https://www.coalitioninc.com/blog/security-labs/the-psychology-of-social-engineering (vid. págs. 18, 27).

[14] *3 Social Engineering Tactics Targeting the Financial Services Industry*. ZeroFox. 2025. URL: https://www.zerofox.com/blog/3-social-engineering-tactics-targeting-the-financial-services-industry/ (vid. pág. 18).

[15] *Internet Crime Report 2024*. Federal Bureau of Investigation, 2024. URL: https://www.ic3.gov/AnnualReport/Reports/2024_IC3Report.pdf (vid. pág. 18).

[16] *Cybersecurity Best Practices*. Cybersecurity e Infrastructure Security Agency (CISA). n.d. URL: https://www.cisa.gov/topics/cybersecurity-best-practices (vid. pág. 18).

[17] *Phishing attacks: defending your organisation*. National Cyber Security Centre (NCSC). 2018. URL: https://www.ncsc.gov.uk/guidance/phishing (vid. págs. 18, 37-39, 42).

[18] Dave Cook y Tyler Johnson. *Cybersecurity Stop of the Month: Attack Sequence of TOAD Threats*. proofpoint. 2023. URL: https://www.proofpoint.com/us/blog/email-and-cloud-threats/cybersecurity-stop-month-attack-sequence-toad-threats (vid. pág. 20).

[19] *MFA Fatigue Attacks: Ultimate Prevention Guide*. Hoxhunt. 2024. URL: https://hoxhunt.com/blog/mfa-fatigue (vid. pág. 20).

[20] *Protect against consent phishing*. Microsoft. n.d. URL: https://learn.microsoft.com/en-us/entra/identity/enterprise-apps/protect-against-consent-phishing (vid. pág. 21).

[21] *Baiting in Cyber Security: What It Is and How to Avoid the Trap*. Threatcop. 2025. URL: https://threatcop.com/blog/baiting-in-cyber-security/ (vid. pág. 21).

[22] *Does Gamified Cyber Security Training Actually Work?* Hoxhunt. n.d. URL: https://hoxhunt.com/blog/gamified-cyber-security-training (vid. págs. 22, 30, 31, 58).

[23] Gyan Chawdhary. *Gamified Cybersecurity Training: Everything You Need to Know*

with 7 Ideas. Security Compass. 2025. URL: https://www.securitycompass.com/blog/gamified-cybersecurity-training/ (vid. págs. 23, 58, 59).

[24] *Is Your Social Engineering Training Actually Working?* Hoxhunt. 2024. URL: https://hoxhunt.com/blog/social-engineering-training (vid. págs. 23, 24, 31, 34).

[25] *How to Measure the Success of Your Security Awareness Program*. Fortra. 2023. URL: https://www.terranovasecurity.com/blog/measure-success-security-awareness-program (vid. págs. 24, 25).

[26] *Phishing Simulation*. proofpoint. 2025. URL: https://www.proofpoint.com/us/threat-reference/phishing-simulation (vid. págs. 25, 28, 29, 34, 58).

[27] *Cybersecurity culture assessment report for Sample Organization*. DVMS Institute. 2024. URL: https://dvmsinstitute.com/wp-content/uploads/2024/10/DVMS-Cybersecurity-Culture-Assessment-Tool-Leaflet-with-Sample-Report.pdf (vid. págs. 25, 60, 61).

[28] *Phishing Simulation: The How, What and Why*. Uniqkey. 2025. URL: https://blog.uniqkey.eu/phishing-simulation/ (vid. págs. 27-31).

[29] *The art of deception: social engineering in red teaming*. Claranet. 2024. URL: https://www.claranet.com/uk/blog/art-deception-social-engineering-red-teaming/ (vid. pág. 29).

[30] *TIBER-EU FRAMEWORK How to implement the European framework for Threat Intelligence-Based Ethical Red teaming*. European Central Bank (ECB), 2025. URL: https://www.ecb.europa.eu/pub/pdf/other/ecb.tiber_eu_framework_2025~b32eff9a10.en.pdf (vid. pág. 30).

[31] *2024 CBEST thematic*. Bank of England. 2024. URL: https://www.bankofengland.co.uk/financial-stability/operational-resilience-of-the-financial-sector/2024-cbest-thematic (vid. pág. 30).

[32] *Security Awareness Computer-Based Training Reviews and Ratings*. Gartner. n.d. URL: https://www.gartner.com/reviews/market/security-awareness-computer-based-training (vid. pág. 30).

[33] *Privacy and Security Compliance*. KnowBe4. n.d. URL: https://www.knowbe4.com/products/compliance-plus (vid. pág. 31).

[34] *Cyber Assessment Framework - Principle D2 Lessons Learned*. National Cyber Security Centre (NCSC). 2025. URL: https://www.ncsc.gov.uk/collection/cyber-assessment-framework/caf-objective-d/principle-d2-lessons-learned (vid. págs. 34, 35).

[35] *Incident Response: Lessons Learned Template*. Secureworks. n.d. URL: https://www.secureworks.com/-/media/files/us/white-papers/secureworks_nco_incidentresponselessonslearnedtemplate.pdf (vid. págs. 35, 36).

[36] *DevSecOps Fundamentals Guidebook: DevSecOps Tools and Activities*. United States Department of War. 2021. URL: https://dodcio.defense.gov/Portals/0/Documents/Library/DevSecOpsTools-ActivitiesGuidebook.pdf (vid. pág. 35).

[37] *Phishing*. MITRE. 2025. URL: https://attack.mitre.org/techniques/T1566/ (vid. págs. 38, 45, 46).
[38] *Require phishing-resistant multifactor authentication for administrators*. Microsoft. 2025. URL: https://learn.microsoft.com/en-us/entra/identity/conditional-access/policy-admin-phish-resistant-mfa (vid. pág. 39).
[39] *Protective Domain Name Service (PDNS)*. National Cyber Security Centre (NCSC). 2024. URL: https://www.ncsc.gov.uk/information/pdns (vid. pág. 39).
[40] *Device security guidance*. National Cyber Security Centre (NCSC). 2025. URL: https://www.ncsc.gov.uk/collection/device-security-guidance (vid. págs. 40-42).
[41] Ritesh Shetty. *7 Best Anti-Phishing Software Solutions for 2025 [AI-Powered Security]*. Arya. 2025. URL: https://arya.ai/blog/best-anti-phishing-software (vid. pág. 42).
[42] *SlashNext*. Varonis. n.d. URL: https://learn.varonis.com/slashnext/ (vid. pág. 42).
[43] *Safe Links in Microsoft Defender for Office 365*. Microsoft. 2025. URL: https://learn.microsoft.com/en-us/defender-office-365/safe-links-about (vid. pág. 43).
[44] *How-to guide: Defending against malware and phishing attacks*. Google. 2025. URL: https://workspace.google.com/blog/identity-and-security/how-guide-defending-against-malware-and-phishing-attacks (vid. pág. 43).
[45] *Email Security*. Cloudflare. 2025. URL: https://www.cloudflare.com/zero-trust/products/email-security/ (vid. pág. 43).
[46] Eric Avigdor y Amit Jain. *Zero Compromise with Content Disarm and Reconstruction (CDR), powered by Zscaler Browser Isolation and Votiro*. Zscaler. 2024. URL: https://www.zscaler.com/blogs/company-news/zero-compromise-content-disarm-and-reconstruction-cdr-powered-zscaler-browser (vid. pág. 43).
[47] *Jamf for Mobile: Reimagine los dispositivos móviles en el trabajo.* Jamf. n.d. URL: https://www.jamf.com/es/soluciones/jamf-for-mobile/ (vid. pág. 44).
[48] *Una introducción al Intercambio de Información de Ciberseguridad - MISP - Threat Sharing*. Computer Incident Response Center Luxembourg (CIRCL). 2024. URL: https://www.misp-project.org/misp-training/0-intro-shorter_es.pdf (vid. págs. 44, 45).
[49] *Tutorial: Use risk detections for user sign-ins to trigger Microsoft Entra multifactor authentication or password changes*. Microsoft. 2025. URL: https://learn.microsoft.com/en-us/entra/identity/authentication/tutorial-risk-based-sspr-mfa (vid. págs. 44, 45).
[50] *What is Conditional Access?* Microsoft. 2025. URL: https://learn.microsoft.com/en-us/entra/identity/conditional-access/overview (vid. pág. 44).
[51] *What is Microsoft Entra Privileged Identity Management?* Microsoft. 2025. URL:

https://learn.microsoft.com/en-us/entra/id-governance/privileged-identity-management/pim-configure (vid. pág. 45).

[52] *How the Probability and Impact Matrix Enhances Risk Management*. SearchInform. n.d. URL: https://searchinform.com/articles/risk-management/tools/probability-and-impact-matrix/ (vid. págs. 45, 46).

[53] *Guide to Cyber Security Incident Response*. Evalian. n.d. URL: https://evalian.co.uk/guide-to-incident-response/ (vid. págs. 48-51).

[54] *Mastering Cybersecurity: How Communicators Can Navigate Through Crises*. Peris.ai. 2025. URL: https://peris.ai/post/mastering-cybersecurity-how-communicators-can-navigate-through-crises (vid. págs. 48, 51).

[55] *Gestión de incidentes para equipos de alta velocidad*. Atlassian. n.d. URL: https://www.atlassian.com/es/incident-management (vid. págs. 49-51).

[56] Jennifer Gregory. *Cybersecurity crisis communication: What to do*. IBM. 2025. URL: https://www.ibm.com/think/topics/cybersecurity-crisis-communication-what-to-do (vid. págs. 50, 51).

[57] *Gestión de incidentes para equipos de alta velocidad*. Atlassian. n.d. URL: https://www.atlassian.com/es/incident-management/devops/chatops (vid. pág. 53).

[58] Daniella Niyonkuru. *Implementing ChatOps into our Incident Management Procedure*. Shopify. 2018. URL: https://shopify.engineering/implementing-chatops-into-our-incident-management-procedure (vid. págs. 53, 54).

[59] *ChatOps: potencia tu equipo técnico en Slack*. Slack. 2025. URL: https://slack.com/intl/es-es/blog/collaboration/chatops-potencia-equipo-tecnico-slack (vid. pág. 53).

[60] Vishal Biyani. *Building a Modern Chatops Platform*. InfraCloud. 2016. URL: https://www.infracloud.io/blogs/building-modern-chatops-platform/ (vid. pág. 53).

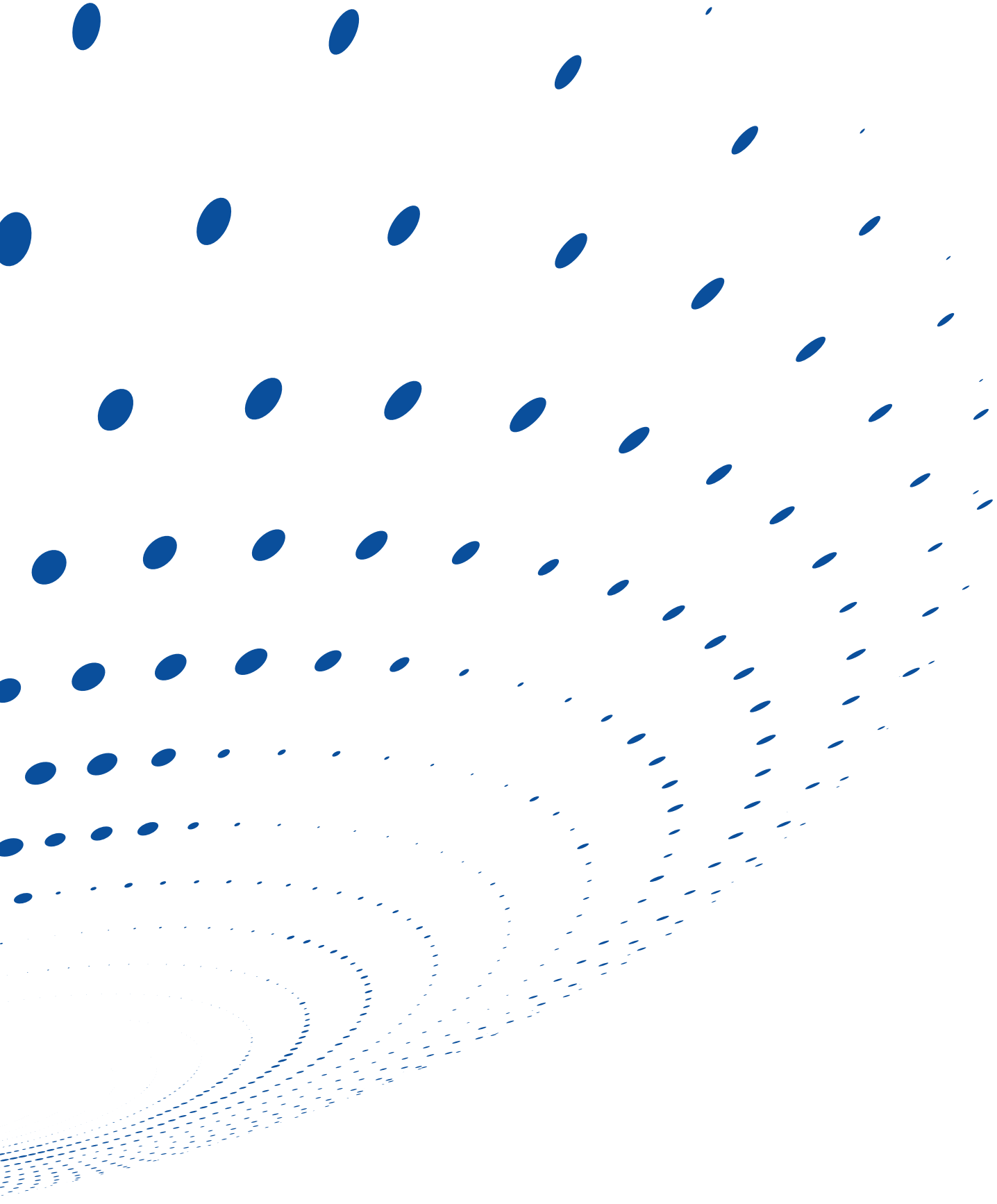
[61] *El Tribunal Supremo establece los requisitos para revisar el correo electrónico del trabajador*. Estudio Jurídico EJASO. 2018. URL: https://ejaso.com/conocimiento/el-tribunal-supremo-establece-los-requisitos-para-revisar-el-correo-electronico-del-trabajador (vid. págs. 54, 55).

[62] *¿El correo corporativo está protegido por el secreto de las comunicaciones?* ECIJA. 2014. URL: https://www.ecija.com/actualidad-insights/el-correo-electronico-corporativo-esta-protegido-por-el-secreto-de-las-comunicaciones/?utm_source=chatgpt.com (vid. págs. 54, 55).

[63] Carlos Requena. *¿Puede patrón auditar correos de empleados? (I)*. Forbes. 2019. URL: https://forbes.com.mx/puede-patron-auditar-correos-de-empleados-i/ (vid. pág. 55).

[64] Carlos Rozen y Martín Elizalde. *¿Puede una organización revisar los emails de sus empleados?* Asociación Argentina de Ética y Compliance. 2025. URL: https://eticaycompliance.org/puede-una-organizacion-revisar-los-emails-d

e-sus-empleados/ (vid. págs. 55, 56).
[65] Gobierno de España. *Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales*. 2018. URL: https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673 (vid. pág. 55).
[66] *Gestión del riesgo y evaluación de impacto en tratamientos de datos personale*. Agencia Española de Protección de Datos (AEPD), 2021. URL: https://www.aepd.es/guias/gestion-riesgo-y-evaluacion-impacto-en-tratamientos-datos-personales.pdf (vid. pág. 56).
[67] *Developing a positive cyber security culture*. National Cyber Security Centre (NCSC). 2025. URL: https://www.ncsc.gov.uk/collection/board-toolkit/principle-c-people/developing-a-positive-cyber-security-culture (vid. pág. 57).
[68] *Cybersecurity Culture Guidelines: Behavioural Aspects of Cybersecurity*. European Union Agency for Cybersecurity (ENISA), 2018. URL: https://www.enisa.europa.eu/sites/default/files/publications/WP2018%200.3.3.2.%20Review%20of%20Behavioural%20Sciences%20Research%20in%20the%20Field%20of%20Cybersecurity.pdf (vid. págs. 57, 58).
[69] *Leading Change: Creating a Culture of Cybersecurity*. Mainstay Technologies. n.d. URL: https://www.mstech.com/leading-change-culture-cybersecurity (vid. pág. 59).
[70] *Change management and cybersecurity: The importance of people in building a secure organisation*. ILX. 2025. URL: https://www.ilxgroup.com/usa/blog/change-management-and-cybersecurity-the-importance-of-people-in-building-a-secure-organisation (vid. pág. 59).
[71] *Security Culture Tool*. National Protective Security Authority (NPSA). 2025. URL: https://www.npsa.gov.uk/security-best-practices/security-culture/security-culture-tool (vid. pág. 60).



FLAR
FONDO LATINOAMERICANO DE RESERVAS

Fondo Latinoamericano de Reservas | FLAR
Calle 84A No. 12-18 Piso 7 | Bogotá, Colombia
Correo electrónico: flar@flar.net
Tel: (571) 634 4360