}

1. **PARTE I**

| **1. Antecedentes Personales** |
| --- |
| A continuación, se presenta una tabla en la que debes completar la información solicitada. |

| Nombre estudiantes | **Moises Sanchez - Denisse Coilla - Nayerly Yáñez** |
| --- | --- |
| Rut | **23.648.469-6 - 20.190.158-8 - 21.376.769-0** |
| Carrera | **Ingeniería en Informática** |
| Sede | **Padre Alonso de Ovalle** |

| **2. Descripción Proyecto APT** |
| --- |
| En la descripción debes señalar brevemente el nombre de tu proyecto APT y las competencias del perfil de egreso que vas a poner en práctica. Si en tu carrera están definidas las áreas de desempeño, también menciona a qué áreas de desempeño está vinculado el proyecto. |

| Nombre del proyecto | SOC Report Manager |
| --- | --- |
| Área (s) de desempeño(s) | * Ciberseguridad * Gestión y Análisis de Datos * Automatización de Procesos de TI * Gestión de Proyectos Informáticos |
| Competencias | * Ofrecer propuestas de solución informática analizando de forma integral los procesos de acuerdo a los requerimientos de la organización. * Desarrollar una solución de software utilizando técnicas que permitan sistematizar el proceso de desarrollo y mantenimiento, asegurando el logro de los objetivos. * Construir modelos de datos para soportar los requerimientos de la organización de acuerdo a un diseño definido y escalable en el tiempo. * Construir el modelo arquitectónico de una solución sistémica que soporte los procesos de negocio de acuerdo los requerimientos de la organización y estándares industria. * Implementar soluciones sistémicas integrales para automatizar y optimizar procesos de negocio de acuerdo con las necesidades de la organización. * Desarrollar la transformación de grandes volúmenes de datos para la obtención de información y conocimiento de la organización a fin de apoyar la toma de decisiones y la mejora de los procesos de negocio, de acuerdo a las necesidades de la organización * Gestionar proyectos informáticos, ofreciendo alternativas para la toma de decisiones de acuerdo con los requerimientos de la organización. * Comunicar en forma oral o escrita, aplicando herramientas lingüístico-pragmáticas y estrategias de comprensión que permiten la solución de problemas comunicativos en los contextos académicos, de acuerdo al marco común de referencia de las lenguas. * Elaborar proyectos innovadores que agreguen valor a contextos sociales y productivos, de acuerdo a las necesidades del entorno. |

| **3. Fundamentación Proyecto APT** |
| --- |
| A continuación, se presentan distintos campos que debes completar con la información solicitada. Esta sección busca que describas en detalle tu proyecto y justifiques su relevancia y pertinencia. |

| Relevancia del proyecto APT | El proyecto es relevante porque aborda una problemática crítica en los Centros de Operaciones de Ciberseguridad (SOC): la gestión manual de reportes de turno y correos de alertas que contienen las observaciones de amenazas reales confirmadas por los analistas. Actualmente, los sistemas SIEM generan miles de alertas diariamente, pero gran parte de ellas corresponden a falsos positivos o eventos de baja relevancia. Son los analistas SOC quienes, a través de la investigación y el análisis, consignan en las observaciones aquellas alertas que representan incidentes verídicos y que requieren seguimiento.  Hoy, estas observaciones y alertas validadas se encuentran dispersas en reportes de turno en formato PDF y correos electrónicos, lo que genera dificultad para consolidar la información, analizarla de manera histórica o detectar patrones de ataque. La falta de un sistema automatizado implica riesgo de pérdida de datos, baja trazabilidad y mayor esfuerzo operativo.  La propuesta del proyecto **SOC Report Manager** responde directamente a esta necesidad: automatiza la ingesta y almacenamiento de reportes y correos del SOC, interpreta el contenido mediante IA (GPT-4o) para estandarizar observaciones heterogéneas y consolida los datos estructurados en BigQuery. Esto permite habilitar dashboards en Looker Studio donde los analistas y supervisores pueden:   * Filtrar incidentes por **IP atacante**, **cliente**, **agente afectado** o **analista que reportó**. * Revisar historiales de ataques y correlacionar eventos entre distintos turnos. * Identificar patrones de comportamiento, como una misma IP atacando varios agentes o repeticiones de un mismo tipo de ataque. * Contar con predicciones generadas en BigQuery ML, que proyectan tendencias de incidentes y apoyan la anticipación frente a amenazas recurrentes.   De esta forma, los principales beneficiados son los **analistas SOC**, que disponen de un repositorio único y consultable de las amenazas reales que ellos mismos confirmaron, reduciendo la carga manual y potenciando sus capacidades de análisis. Asimismo, el **COO** obtiene información consolidada y trazable, con indicadores confiables y capacidades predictivas que fortalecen la toma de decisiones y la comunicación con los clientes de la empresa.  En términos de aporte laboral y social, el proyecto contribuye a aumentar la eficiencia operativa del SOC, a mejorar la detección temprana de amenazas y a reforzar la seguridad de las organizaciones que dependen de sus servicios.  . |
| --- | --- |
| Descripción del Proyecto APT | El proyecto **SOC Report Manager** tiene como objetivo automatizar la gestión de reportes de turno en formato PDF y correos de alertas de seguridad que actualmente se reciben de manera manual en el centro de operaciones de ciberseguridad SC2. La solución busca garantizar trazabilidad, integridad y disponibilidad de la información, reduciendo riesgos de pérdida de datos, mejorando la eficiencia operativa y habilitando la explotación analítica y predictiva de los registros.  La propuesta consiste en diseñar e implementar una arquitectura de automatización y análisis que contemple el siguiente flujo:   * **Conexión al correo electrónico (Gmail API):** detección y descarga automática de correos y adjuntos en formato PDF que llegan cada 6 horas, junto con mensajes de alertas generados por los analistas SOC. * **Almacenamiento crudo en Google Cloud Storage:** resguardo inmutable de PDF y evidencias originales, aplicando versionado e idempotencia mediante hash para evitar duplicados y asegurar trazabilidad. * **Bitácora operacional en Cloud SQL:** registro del estado de procesamiento de cada mensaje (nuevo, procesado, error), número de intentos, latencias y causas de fallo, permitiendo auditoría y reintentos controlados. * **Interpretación con IA (GPT-4o):** estandarización de observaciones y alertas heterogéneas, devolviendo información estructurada en formato JSON con campos como fecha, cliente, tipo de alerta, severidad, IPs y descripción del evento. * **Carga analítica en BigQuery:** inserción de datos estructurados en tablas particionadas y clusterizadas, optimizadas para consultas masivas y visualización. * **Orquestación de procesos con n8n (Docker):** automatización del flujo completo de ingesta, transformación y carga de datos (ETL), garantizando portabilidad y control. * **Dashboards en Looker Studio:** visualización en tiempo real de métricas clave del SOC, como distribución de alertas, tendencias de criticidad, top IPs y clientes afectados. * **Predicciones con BigQuery ML:** análisis de series de tiempo y modelos predictivos sobre el volumen de alertas, habilitando detección temprana de patrones de ataque y posibles incidentes futuros.   De esta forma, el proyecto no solo resuelve el problema de la gestión manual de reportes y alertas, sino que incorpora un enfoque innovador al aprovechar inteligencia artificial para estructurar datos no estandarizados, garantizando trazabilidad con GCS y Cloud SQL, y potenciando la analítica y predicción en BigQuery y Looker Studio. |
| Pertinencia del proyecto con el perfil de egreso | El proyecto SOC Report Manager se vincula directamente con el perfil de egreso de Ingeniería en Informática, ya que integra competencias claves de la carrera en un contexto real de ciberseguridad. La problemática abordada es la gestión manual y dispersa de reportes de turno y alertas en formato PDF que requieren aplicar conocimientos en automatización de procesos, gestión de datos, ciberseguridad y desarrollo de soluciones tecnológicas escalables, áreas centrales de la formación profesional.  Para resolverla, resulta esencial la competencia de desarrollar soluciones de software que sistematicen procesos, evidenciada en la construcción de un flujo automatizado de extracción y procesamiento de reportes. Asimismo, la competencia de construir modelos de datos escalables se refleja en el diseño de la base en BigQuery, que permite organizar y analizar grandes volúmenes de registros de turnos y observaciones.  De igual forma, la competencia de implementar soluciones sistémicas integrales para automatizar y optimizar procesos de negocio se materializa en la orquestación con n8n y el despliegue en contenedores Docker. Finalmente, la competencia de gestionar proyectos informáticos y comunicar de forma clara información crítica se refleja en la planificación bajo metodología ágil y en la estandarización de reportes heterogéneos para transformarlos en información útil para la toma de decisiones.  En conjunto, estas competencias permiten dar solución a la problemática planteada, fortaleciendo la pertinencia del proyecto dentro del campo laboral de la Ingeniería en Informática y asegurando coherencia con el perfil de egreso. |
| Relación con los intereses profesionales | Nuestros intereses profesionales están orientados a la ciberseguridad, administración de sistemas en Linux, automatización de procesos y análisis de datos. Este proyecto refleja esos intereses, nos permite:   * Aplicar en un caso real el uso de tecnologías de automatización (n8n, Docker) y plataformas cloud (GCP). * Reforzar conocimientos en ciberseguridad y gestión de información crítica, propios de un entorno SOC. * Incorporar inteligencia artificial para mejorar la eficiencia en la lectura y procesamiento de información no estandarizada.   Realizar este proyecto contribuirá directamente a nuestro desarrollo profesional, ya que nos permitirá adquirir experiencia práctica en herramientas y metodologías que son altamente demandadas en el mercado laboral, potenciando nuestro perfil como especialista en ciberseguridad y automatización. |
| Factibilidad de desarrollo del Proyecto APT | El proyecto es factible de desarrollar dentro del semestre, ya que la duración de 10 semanas permite organizar un plan de trabajo con etapas definidas de diseño, implementación, pruebas y documentación. Las horas asignadas a la asignatura, junto con el trabajo autónomo del equipo, aseguran tiempo suficiente para cumplir los hitos propuestos.  En cuanto a materiales, se utilizarán principalmente recursos disponibles en Google Cloud Platform, el motor de orquestación n8n y contenedores Docker, lo que evita la necesidad de infraestructura física adicional.  Entre los factores que facilitan el desarrollo se encuentran el interés y compromiso de los integrantes, la experiencia previa obtenida en la práctica profesional en un entorno SOC, y los conocimientos adquiridos a lo largo de la carrera. Como posible dificultad, se considera la heterogeneidad de los reportes y observaciones, que presentan estructuras distintas según el analista. Este desafío se abordará mediante estandarización, uso de IA para interpretación de texto y pruebas iterativas.  En conclusión, el proyecto es viable en el tiempo establecido, con los recursos disponibles y con un nivel de complejidad adecuado para ser desarrollado por el equipo dentro del semestre académico. |

1. **PARTE II**

| **4. Objetivos** |
| --- |
| En este apartado debes definir objetivos generales y específicos del Proyecto APT. Es importante aclarar que los objetivos se deben plantear en forma clara, concisa y sin dar mayores explicaciones, es decir, deben entenderse por sí solos. Se sugiere redactarlos utilizando un verbo en infinitivo, pues ello obliga a precisar acciones concretas. |

| Objetivo general | Automatizar la gestión de reportes de turno y alertas de seguridad generados en el centro de operaciones de ciberseguridad SC2, mediante una solución que permita organizar la información, reducir procesos manuales y facilitar análisis avanzados para la toma de decisiones. |
| --- | --- |
| Objetivos específicos | * Conectar el sistema de correo electrónico para descargar automáticamente los reportes de turno y alertas de seguridad. * Implementar un proceso de estandarización y deduplicación de la información recibida desde distintos analistas. * Diseñar y desplegar una base de datos en Google Cloud Platform para almacenar y organizar reportes y alertas de forma segura y escalable. * Configurar flujos de ingesta y transformación de datos mediante n8n y procesamiento con IA (GPT). * Asegurar la portabilidad del sistema utilizando contenedores Docker. * Desarrollar dashboards interactivos en Looker Studio que permitan visualizar métricas clave y correlacionar reportes con alertas. * Integrar las observaciones de turno y las alertas procesadas como insumo para generar indicadores y habilitar análisis predictivos con BigQuery ML. |

| **5. Metodología** |
| --- |
| En el siguiente apartado deberás describir la metodología, propia de tu disciplina, que utilizarás para resolver el proyecto APT antes descrito, incluyendo las etapas y métodos de trabajo. |

| Descripción de la Metodología |
| --- |
| El desarrollo del proyecto se llevará a cabo bajo un enfoque ágil, utilizando la metodología Scrum como marco principal de trabajo. Esta elección se justifica porque permite organizar las tareas en ciclos cortos e iterativos, entregando avances funcionales en cada sprint y asegurando la adaptación a cambios o imprevistos que puedan surgir durante el semestre.  El proyecto se dividirá en sprints de 2 semanas, cada uno enfocado en un componente clave:   * **Sprint 1:** Configuración de n8n y conexión con la API de Gmail. * **Sprint 2:** Procesamiento de archivos PDF y alertas, estandarización con GPT. * **Sprint 3:** Integración con BigQuery y pruebas de consultas. * **Sprint 4:** Implementación de dashboards en Looker Studio. * **Sprint 5:** Ajustes finales, validación con escenarios reales y documentación.   Las funciones del equipo se organizarán de la siguiente manera:   * **Product Owner:** Definir prioridades, validar entregables y mantener la visión general del proyecto. * **Scrum Master:** Coordinar las actividades del equipo, remover obstáculos y garantizar el cumplimiento de la metodología. * **Equipo de Desarrollo:** Implementar técnicamente cada sprint, incluyendo la orquestación con n8n, despliegue en contenedores Docker, almacenamiento en GCP y construcción de dashboards en Looker Studio.   Este enfoque ágil fomenta la colaboración constante, la entrega continua de valor y asegura que el producto final cumpla con los objetivos definidos, dentro del plazo establecido para la asignatura. |

| **6. Evidencias** |
| --- |
| A continuación, describe qué evidencias serán evaluadas en el informe de avance y en el informe final de tu proyecto APT. Estas evidencias deben ser acordadas con tu docente. Se entenderá por evidencia los productos que se desarrollen durante el proyecto y cuyo propósito sea visibilizar o documentar cómo se ha implementado el trabajo. |

| **Tipo de evidencia**  **(avance o final)** | **Nombre de la evidencia** | **Descripción** | **Justificación** |
| --- | --- | --- | --- |
| **Avance** | **Documento de diseño de arquitectura** | **Representación de la arquitectura del proyecto, mostrando los componentes (n8n, Gmail API, GPT, BigQuery, Looker Studio) y su interacción.** | **Permite validar que la solución planteada responde al problema y que es técnicamente viable.** |
| **Avance** | **Flujo automatizado en n8n** | **Capturas y pruebas iniciales de la configuración de n8n para la conexión con Gmail y extracción de correos.** | **Evidencia el inicio de la automatización, asegurando la conexión con la fuente de datos.** |
| **Avance** | **Prototipo de procesamiento de datos con GPT** | **Resultados de pruebas de extracción y estandarización de observaciones de turno y alertas usando GPT.** | **Muestra que el sistema puede interpretar información no estandarizada y transformarla en datos estructurados.** |
| **Final** | **Base de datos en BigQuery** | **Dataset con los reportes procesados, listos para consultas.** | **Garantiza que la información fue organizada y centralizada para análisis.** |
| **Final** | **Dashboard en Looker Studio** | **Visualización interactiva de métricas, indicadores y predicciones.** | **Permite a los usuarios (SOC y COO) acceder fácilmente a información relevante para la toma de decisiones.** |
| **Final** | **Informe de pruebas y validación** | **Documento con los resultados de las pruebas realizadas al sistema, incluyendo limitaciones y mejoras.** | **Asegura que el sistema fue validado y cumple con los objetivos planteados.** |

| **7. Plan de Trabajo** |
| --- |
| En la siguiente tabla define la planificación de tu Proyecto APT de acuerdo a lo requerido. |

| **Plan de Trabajo Proyecto APT** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Competencia o unidades de competencias | Nombre de Actividades/Tareas | Descripción Actividades/Tareas | Recursos | Duración de la actividad | Responsable[[1]](#footnote-0) | Observaciones |
| *Ofrecer propuestas de solución informática / Construir el modelo arquitectónico de una solución sistémica / Gestionar proyectos informáticos* | Diseño de la arquitectura | Definir los componentes del sistema (n8n, Gmail API, GCP, GPT, BigQuery, Looker Studio) y el flujo de integración. | *Documentación, herramientas de diagramación (Draw.io, Lucidchart).* | *1 semana* | Todo el equipo | Puede requerir ajustes según observaciones del docente. |
| *Desarrollar soluciones de software / Implementar soluciones sistémicas integrales para automatizar y optimizar procesos de negocio* | *Configuración de n8n en Docker* | *Implementar n8n en un contenedor Docker en GCP para orquestar el flujo ETL de correos (ingesta, control de errores, reintentos).* | *Docker, VM en GCP, créditos académicos.* | *1 semanas* | *Nayerly* | *Posibles dificultades en permisos de red y credenciales.* |
| *Desarrollar la transformación de grandes volúmenes de datos* | *Procesamiento de correos y archivos PDF* | *Conectar Gmail API, descargar reportes de turno y alertas de seguridad; envío a GPT para estandarizar campos (fecha, IP, agente, tipo de evento, severidad, etc.).* | *Gmail API, GPT API, plantillas de extracción.* | *3 semanas* | *Denisse* | *Riesgo de incompatibilidad en formatos de PDF o estructura de alertas.* |
| *Construir modelos de datos para soportar los requerimientos de la organización* | *Integración con bases de datos* | *Estructurar y almacenar los datos procesados en Cloud SQL (operacional) y BigQuery (analítico) con claves, particiones y políticas de acceso.* | *Cloud SQL, BigQuery.* | *1 semana* | *Moisés* | *Posibles retrasos en la configuración de permisos IAM.* |
| *Desarrollar la transformación de grandes volúmenes de datos para la obtención de información y conocimiento de la organización.* | *Creación de dashboards en Looker Studio* | *Diseñar paneles con métricas relevantes: frecuencia de alertas, IPs recurrentes, agentes afectados, tendencias, responsables que notifican, etc.* | *BigQuery, Looker Studio* | *2 semanas* | *Nayerly* | *Puede requerir iteraciones para definir métricas relevantes al SOC.* |
| *Elaborar proyectos innovadores que agreguen valor a contextos sociales y productivos* | *Implementación de análisis predictivo* | *Configurar BigQuery ML para modelos simples (tendencias, probabilidad de repetición por IP/agente/severidad) y exponer resultados en el dashboard.* | *BigQuery ML* | *2 semanas* | *Moisés* | *Puede requerir limpieza de datos para entrenar modelos.* |

| **8. Carta Gantt** |
| --- |
| En esté apartado se visualiza las actividades en general que realizaremos en todo este proceso de desarrollo del proyecto APT, además con el tiempo cambiaremos estas actividades generales por actividades más específicas. |

| **Actividad** | **Fase 1** | | | | **Fase 2** | | | | | | | | | | | | **Fase 3** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **S 1** | **S 2** | **S 3** | **S 4** | **S 5** | **S 6** | **S 7** | **S 8** | **S 9** | **S 10** | **S 11** | **S 12** | **S 13** | **S 14** | **S 15** | **S 16** | | **S 17** | **S 18** |
| **Rellenar documentación individual** | **x** | **x** |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| **Rellenar documentación Grupal y Presentación** |  |  | **x** |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| Presentar el proyecto |  |  |  | **x** |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| Rellenar documento de reflexión y charlar posibles cambios al proyecto. |  |  |  |  | **x** |  |  |  |  |  |  |  |  |  |  |  | |  |  |
| Rellenar documentos grupales |  |  |  |  |  | **x** |  |  |  |  |  |  |  |  |  |  | |  |  |
| Desarrollo de proyecto |  |  |  |  |  |  | **x** | **x** | **x** | **x** | **x** | **x** | **x** | **x** |  |  | |  |  |
| Rellenar documentación de evidencias del proyecto y presentación |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **x** |  | |  |  |
| Charlar las posibles mejoras y rellenar documentación individual y grupal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **x** | |  |  |
| Realizar última etapa de presentación |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | **x** | **x** |

1. En caso de que el Proyecto APT sea grupal, en esta columna deben indicar el nombre de los responsables de cada tarea o actividad. Esto posteriormente permitirá diferenciar la evaluación por cada integrante. [↑](#footnote-ref-0)