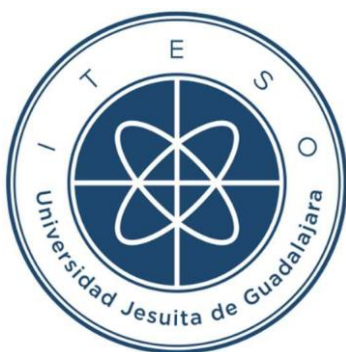


ITESO, UNIVERSIDAD JESUITA DE GUADALAJARA

Reconocimiento de validez oficial por acuerdo secretarial número 15018 publicado en el Diario Oficial de la Federación el 29 de noviembre de 1976.

MASTER IN BUSINESS ADMINISTRATION



INTELIGENCIA ARTIFICIAL GENERATIVA Y SUS OPORTUNIDADES PARA LOS
NEGOCIOS EN MÉXICO.

Informe del proceso de gestión del conocimiento para obtener el grado de

MAESTRO EN ADMINISTRACIÓN DE EMPRESAS

Presenta: Alberto de Jesús Zárate Gómez

Asesor: José Juan Calzada López

Tlaquepaque, Jalisco. 15 de junio de 2024

Resumen:

El documento elabora un análisis detallado sobre las inversiones y oportunidades relacionadas con la Inteligencia Artificial Generativa (*GenAI*) a nivel mundial, con un enfoque específico en el estado de Jalisco. El autor, con experiencia en empresas como Amazon y Google, destaca la necesidad de desarrollar un marco regulatorio para el uso responsable de *GenAI*, buscando mitigar riesgos potenciales y maximizar beneficios estratégicos para las empresas.

Se menciona que los clientes actualmente buscan herramientas intuitivas para implementar tecnologías de *GenAI*, con aplicaciones que van desde la optimización de ventas hasta mejoras en la seguridad del cliente. Sin embargo, se identifica que muchos clientes carecen del conocimiento técnico necesario para aprovechar al máximo estas herramientas.

Se identifican varios retos en la adopción de *GenAI*, incluyendo problemas éticos como *deepfakes* y discriminación por sesgos en datos. También se mencionan desafíos laborales debido a la automatización potencial de roles, como en el entretenimiento y la industria automotriz. Se propone que Jalisco aproveche *GenAI* en áreas como educación, industria, sociedad, gobierno y negocios para optimizar procesos y servicios, pero advierte que el retraso en la adopción podría llevar a una desventaja competitiva.

En resumen, la propuesta del autor es una llamada a la acción para que Jalisco explore y adopte proactivamente *GenAI*, identificando áreas clave de aplicación y promoviendo medidas regulatorias que faciliten su adopción segura y efectiva en diversas industrias locales, aprovechando así sus beneficios mientras mitiga sus riesgos potenciales.

Palabras Clave:

- GenAI
- Google
- Cloud
- Inteligencia Artificial
- Jalisco

ÍNDICE

INTRODUCCIÓN	6
¿Qué es la <i>GenAI</i> ?	6
Ejemplo 1: Diseño automotriz con ayuda de <i>GenAI</i>	6
Ejemplo 2: Predicción de cambios climáticos a través de <i>GenAI</i>	7
¿Qué son <i>OpenAI</i> y <i>ChatGPT</i> ?.....	7
FUNDAMENTACIÓN DEL TRABAJO	9
Objetivo	9
Alcance.....	9
Dentro del alcance del trabajo	9
Fuera del alcance del trabajo	10
PROBLEMÁTICA	11
Contexto actual de <i>Google</i>	11
Indexando las páginas de internet desde 1994.....	13
Fabricando chips especializados para la Inteligencia Artificial	16
¿Cuál es la diferencia entre las TPU y las GPU?	18
Análisis de mercado.....	19
MARCO CONCEPTUAL	21
Metodología.....	21
Herramientas de cómputo utilizadas en la elaboración del presente trabajo.....	23

¿Cómo funciona la Inteligencia Artificial Generativa (<i>Generative AI</i> o <i>Gen AI</i>)?.....	24
Contexto de <i>GenAI</i> como negocio	28
Roles y Responsabilidades.....	29
Diagrama de Flujo de datos como una herramienta fundamental	31
PROPUESTA DEL AUTOR.....	33
¿Qué se espera que suceda en base a esta propuesta?	33
¿En cuánto tiempo se tienen que tomar las acciones?	33
Análisis de la problemática en Jalisco	34
Retos Morales y Éticos	34
Retos Laborales	36
Retos de Adopción.....	38
Áreas de oportunidad para Jalisco	39
Riesgos de no aplicar la <i>GenAI</i> en Jalisco	42
Propuesta de un marco de regulación para una <i>GenAI</i> responsable.....	45
Transparencia en las fuentes de origen	46
Controles de seguridad	46
Apego a los valores humanos	46
Democratización	47
CONCLUSIONES.....	48
Bibliografía.....	50

No le estamos dejando un mundo maltrecho a aquellos que siguen nuestros pasos, les seguimos construyendo un mundo mejor que el que nos heredaron nuestros antecesores.

-Alberto de Jesús Zárate Gómez

INTRODUCCIÓN

¿Qué es la GenAI?

“La IA generativa (GenAI en inglés) se refiere a la rama de la inteligencia artificial que se enfoca en la creación de contenido nuevo y único—como texto, arte visual, música, código de software y más. A diferencia de la IA predictiva (o analítica), que utiliza el aprendizaje automático para analizar datos históricos, identificar patrones o tendencias y luego hacer predicciones, la IA generativa va un paso más allá de simplemente analizar y predecir: la IA generativa crea” (MongoDB, 2024).

GenAI se apoya en modelos de Inteligencia Artificial que son básicamente programas de aprendizaje automático y especializado que toman como base una serie de datos preexistentes para de ahí hacer un análisis, procesamiento y por último generar nuevos contenidos. Para ello, existen modelos especializados para analizar texto -comúnmente conocidos como LLM o modelos de lenguaje largo por sus siglas en inglés -, otros que son especializados en video, audio, etcétera. Un modelo que es especializado en el análisis de texto no será apto para analizar un video, así que cada modelo es creado para un fin específico y dependerá del contexto y resultados esperados del usuario.

Ejemplo 1: Diseño automotriz con ayuda de GenAI

¿Te imaginas poder diseñar un coche familiar que sea sumamente eficiente en el consumo de combustible con tan solo unas pocas palabras?

Suponga que tiene varias fotografías, películas o libros que describen las características de los automóviles en general (camionetas, autos compactos, deportivos, de lujo, monoplazas, SUV, etc.). GenAI podría analizar todos esos objetos y con su tecnología de Inteligencia Artificial generar nuevos diseños de automóviles instantáneamente en base a las especificaciones (en texto, imagen, video, etc.) que se le digan al modelo. Por ejemplo, *“un auto familiar con la mayor eficiencia aerodinámica y con el menor consumo posible de gasolina”*.

Entre más grande sea el número de objetos (fotos, películas, etc.), mayor será la fiabilidad en el diseño de los automóviles que te presentará el modelo. Ahora imagina que la base de datos que contiene los objetos o información a la que se le da acceso al modelo de GenAI es algo tan vasto como lo

almacenado en Internet (por ejemplo, lo indexado por los motores de búsqueda como *Google*), entonces tenemos una fuente en la que se basarían los modelos de *GenAI* que contienen un sinnúmero de imágenes, objetos, libros, audios, etc. que se han ido recopilando a lo largo de las últimas décadas. Casos como este podrán ayudar a que las empresas automotrices aceleren el desarrollo de sus autos de forma rápida, sencilla, económica, eficiente y responsable.

Ejemplo 2: Predicción de cambios climáticos a través de *GenAI*

Con modelos de *GenAI*, se podrían analizar 50 años de datos climáticos (p.ej., barimetría, temperatura, flora, fauna, contaminación, densidad poblacional, etc.) para pronosticar los cambios climáticos. E incluso se pueden obtener distintos resultados si se le introducen parámetros como por ejemplo ¿cómo se modificaría el clima de una región ante una hipotética explosión demográfica de X% anual durante los próximos 10 años?

GenAI simula escenarios hipotéticos al instante y puede ayudar a la toma de decisiones en base a información real que pueden coadyuvar a implementar medidas para mitigar impactos ambientales como sequías o inundaciones de forma proactiva y preventiva.

¿Qué son *OpenAI* y *ChatGPT*?

Desde finales del 2022, la *GenAI* ha experimentado un auge sin precedentes y en gran parte ha sido impulsada por el lanzamiento de *chatbots* como *ChatGPT*, desarrollado por la empresa *OpenAI*. Estas tecnologías analizan conjuntos de datos masivos para sintetizarlos y producir contenido nuevo en forma de, imágenes, videos, sonidos o texto.

ChatGPT es lo que comúnmente se denomina un “*chatbot*” que es una interfaz mediante la cual el usuario puede hacer peticiones (*prompts*, en inglés) para obtener una respuesta, típicamente mediante la generación de contenido nuevo en base a información ya existente. Estas peticiones pueden ser en forma de preguntas, como, por ejemplo, “¿Por qué el cielo es azul?” o ejecutar un comando como “Elabora una gráfica de tipo histograma mediante la cual se represente la densidad poblacional del municipio de Tlaquepaque, Jalisco, México en los últimos 20 años”.

El *chatbot* de *ChatGPT* está conectado a modelos de *GenAI* específicos que analizarán la pregunta, harán una búsqueda en las bases de datos relevantes y mediante un modelo de Inteligencia Artificial

conocido como “Lenguaje Natural” (*Natural Language*), brindará una respuesta a la petición del usuario articulada de tal forma que simule provenir de otra persona que habla el mismo idioma en el que fue hecha la petición. *ChatGPT* es el *chatbot* que fue desarrollado por la empresa *OpenAI* junto con los modelos de *GenAI* con los que interactúa para generar sus respuestas. Existen muchos *chatbots* al día de hoy y cada uno tiene un nombre distinto, dependiendo de la empresa que lo desarrolla y quien tiene los derechos sobre el mismo, así como los modelos de *GenAI* con los que interactúa. Otros *chatbots* populares son *Gemini* de *Google*, *Meta AI* de *Facebook* o *Copilot* de *Microsoft*, por nombrar algunos.

Dichos modelos de inteligencia artificial proponen una alternativa para generar avances en cualquier industria o campo de especialización de forma acelerada.

FUNDAMENTACIÓN DEL TRABAJO

Objetivo

Este trabajo de búsqueda y análisis está elaborado para lograr i) dar a conocer las inversiones que se están teniendo alrededor de *GenAI* a nivel mundial; ii) explorar las áreas de oportunidad en los negocios en Jalisco para procurar -en la medida de lo posible- mitigar los efectos adversos de un rezago en la adopción de tecnologías de *GenAI*, así como el riesgo en el uso de estas tecnologías para fines nocivos; iii) y proponer un marco de regulación para el uso de *GenAI* de forma responsable.

Las opiniones expresadas en este trabajo son únicamente del autor y no expresan los puntos de vista ni las opiniones de *Google*.

Alcance

El presente trabajo describe el conocimiento y experiencia adquirida por el autor a través de su trabajo en los departamentos de Inteligencia Artificial de Amazon (2019-2021 en Alexa) pero principalmente en *Google* (2021 a la fecha) en donde ha trabajado en la implementación de procesos de desarrollo de software y regulaciones para proveerle a los clientes las herramientas de *GenAI* que necesitan, particularmente en las plataformas de *Vertex AI* integradas con *Gemini* todo esto bajo el puesto de Gerente de Proyectos en los departamentos y empresas anteriormente mencionadas.

En Amazon, particularmente en el equipo de *Fundamentos de Seguridad de IA* (SAIF, por sus siglas en inglés) dentro de la organización de *Alexa*, el autor coordinó proyectos para prevenir la exfiltración de datos de los usuarios por parte de los modeladores que desarrollan los modelos de Inteligencia Artificial. Mientras que, en *Google*, ha coordinado el lanzamiento de varias herramientas como *Managed Notebooks*, *Colab Enterprise* y *Workbench Instances* que son interfaces de desarrollo de software para escribir y ejecutar código fuente de software dentro de la plataforma de la nube de *Google* (*Google Cloud Platform* también conocida como GCP).

Dentro del alcance del trabajo:

- Descripción general de clientes (anónimos) y la forma en la que aplican modelos de *GenAI*.

- Metodologías y herramientas aprendidas por parte del presentador que han sido útiles para el desarrollo de este trabajo y de sus funciones como Gerente de Proyectos en los departamentos de Inteligencia Artificial y en el desarrollo de software en general.
- Aplicación de técnicas y herramientas para satisfacer las necesidades de negocio, seguridad y privacidad que salvaguardan los intereses propios de los clientes, así como los de *Google*.
- Señalar áreas de oportunidad para la aplicabilidad de las tecnologías de *GenAI* dentro de industrias específicas en el estado de Jalisco.
- Conclusiones personales acerca de las previsiones que se esperan en la evolución y aplicación de *GenAI*.

Fuera del alcance del trabajo:

- Nombres específicos de los clientes, así como información confidencial (datos económicos, problemáticas específicas de los clientes, etc.)
- Técnicas, código fuente y herramientas específicas para el desarrollo de modelos de Inteligencia Artificial.
- Planes y estrategias actuales de *Google* o sus clientes que aún no son del dominio público.
- Información de Identificación Personal (*PII* por sus siglas en inglés), así como diagramas organizacionales, estados financieros o análisis financieros de los clientes o de la empresa.

PROBLEMÁTICA

La Inteligencia Artificial es un campo sumamente vasto, complejo y que apenas comienza a despegar, los clientes de *Vertex AI* buscan principalmente un ecosistema de herramientas intuitivas que les ayuden a implementar estas tecnologías en sus empresas. Ya sea para aumentar el número de sus ventas, disminuir el costo de sus operaciones, mejorar la seguridad de sus clientes, etcétera.

Algunos clientes desconocen cómo beneficiarse de la Inteligencia Artificial desde un punto de vista estratégico. Otros carecen del conocimiento técnico para desarrollar sus propios modelos, entrenarlos, evaluarlos o beneficiarse de los mismos. Es por ello que *Vertex AI* les brinda a los clientes un sinnúmero de modelos de *GenAI* (Google, Versiones del modelo y ciclo de vida, 2024) para que hagan uso de los mismos o elijan si prefieren hacer uso de las diferentes herramientas para el desarrollo, ejecución, almacenamiento, evaluación o entrenamiento para sus modelos propios.

Contexto actual de Google

Google es una empresa con presencia global y con sede en Mountain View, California, Estados Unidos. Fue fundada en 1998 por Larry Page y Sergey Brin, dos estudiantes de Doctorado de la Universidad de Stanford y cuyo propósito era organizar la información de todo el mundo para ponerla al alcance de todos de forma sencilla y gratuita.

Desde sus inicios, uno de los ejes fundamentales de la compañía fue el desarrollo de su buscador de internet (www.Google.com), pues conforme fue adquiriendo popularidad, dio pie a su monetización a través de anuncios (lanzados en el año 2000) para el posicionamiento de empresas, marcas, servicios e individuos en los resultados de búsqueda de sus usuarios.

En 2004, *Google* se preparaba para dar un paso crucial en su historia: su salida a bolsa (Comúnmente conocido en inglés como *IPO - Initial Public Offering*), en ese momento, los cofundadores de la compañía, Larry Page y Sergey Brin escribieron una carta que ahora se considera un documento fundamental para entender la esencia y los valores de la compañía (Brin, 2004):

Google no es una empresa convencional. No procuramos que lo sea. A través de la evolución de *Google* como una empresa privada, hemos manejado *Google* de formas distintas. También

hemos enfatizado una atmósfera de creatividad y retos que nos ha ayudado a brindar acceso a la información para todos aquellos que dependen de nosotros para hacerlo de forma objetiva, veraz y gratuita. (Brin, 2004)

A continuación, se muestra la misión, visión (Pereira, 2023) y los valores de la organización (Google, Community Guidelines, s.f.):

- Misión: Organizar la información del mundo y hacerla universalmente accesible y útil.
- Visión: Brindar el acceso a la información del mundo en un solo *clic*.
- Valores para sus empleados:
 - Respeta al usuario
 - Respeta la oportunidad
 - Respétense los unos a los otros

Pero contrario a lo que se piensa, *Google* no es únicamente un motor de búsqueda de páginas de internet, sino que se ha posicionado como un gigante tecnológico con un ecosistema de productos y servicios que abarca distintos sectores.

Más allá del buscador, genera utilidades en base a la venta de anuncios que le ayudan a las empresas a posicionar sus marcas dentro de los resultados de búsqueda (*Google Ads*) o productos como *Gmail* para correos electrónicos.

Google forma parte del conglomerado llamado *Alphabet*, que incorpora muchos otros productos y servicios como el sistema operativo para dispositivos y computadoras conocido como *Android*, *YouTube*, *Waymo* para la investigación y desarrollo (R&D) de automóviles autónomos o *Google Cloud Platform* – Que de acuerdo a los últimos informes de Gartner (David Wright, 2023) conocidos como “*Gartner Magic Quadrant*”, es uno de los 3 principales líderes en cuanto a proveer servicios en la nube (*Cloud*) a nivel global.

Ilustración 1, análisis PESTEL de *Google* de elaboración propia.



Como se muestra en la ilustración 1, haber realizado el análisis PESTEL le brindó al autor una mayor claridad y conciencia acerca de los diversos factores que intervienen en el éxito de una empresa como *Google*. Antes de dicho análisis, el autor tenía una visión parcial y subjetiva que estaba limitada por las experiencias en los departamentos específicos en los que ha laborado, sin tener una visión más panorámica.

Indexando las páginas de internet desde 1994

Gran parte del éxito del buscador de *Google* se ha debido a que, desde sus inicios, se ha automatizado la indexación de páginas de internet. Tal como lo describe en su página para desarrolladores (Google, Guía detallada sobre cómo funciona la Búsqueda de Google, 2024), una vez que se “rastrea una página, Google intenta comprender de qué se trata.” (la misma) “Esta etapa se denomina indexación y, además, incluye el procesamiento y el análisis del contenido del texto y las etiquetas y los atributos del contenido clave” como pueden ser los metadatos (p.ej. las palabras clave que describen la página web, las imágenes, videos, tablas, etc.),

Durante el proceso de indexación, Google determina si una página es el duplicado de otra página en Internet o es canónica. La página canónica es la que puede aparecer en los resultados de la búsqueda. Para seleccionar la página canónica, primero agrupamos (que se conoce como agrupamiento en clústeres) las páginas con contenido similar que encontramos en Internet y, luego, seleccionamos la que sea más representativa del grupo. Las otras páginas del grupo son versiones alternativas que pueden publicarse en diferentes contextos, como cuando el usuario realiza una búsqueda desde un dispositivo móvil o busca una página muy específica de ese clúster.

Google también recopila indicadores sobre la página canónica y su contenido, que se pueden usar en la siguiente etapa, en la que publicamos la página en los resultados de la búsqueda. Algunos indicadores incluyen el idioma de la página, el país en el que es local el contenido, la usabilidad de la página, etcétera.

La información recopilada sobre la página canónica y su clúster se pueden almacenar en el índice de Google, una gran base de datos alojada en miles de computadoras. La indexación no está garantizada. No se indexan todas las páginas que procese Google (Google, Guía detallada sobre cómo funciona la Búsqueda de Google, 2024)

Otra acción que realiza *Google* en el proceso de indexación, es determinar si la página tiene contenido sensible (p.ej. pornografía, ilegal, falso, etc.) para marcarla en una categoría específica y que dicho contenido no aparezca en búsquedas por parte de otros usuarios. Esto con el fin de evitar diseminar contenido ilegal, falso, etc. Un claro ejemplo fue que durante la pandemia ocasionada por el COVID-19, *Google* se dio a la tarea de luchar contra la desinformación en línea a través de filtrar contenido que era detectado como falso, subjetivo o sin fundamentación. Tal como lo explica en su artículo (El enfoque de Google para luchar contra la desinformación online, s.f.) a continuación.

A medida que la pandemia evolucionaba, también lo hacían las búsquedas online sobre temas relacionados con la salud, como tests en casa, vacunas y mascarillas, lo que a su vez ofrecía a los agentes perniciosos la oportunidad de sacar partido de la situación.

Desde hace tiempo tenemos políticas en todos nuestros productos que prohíben el contenido médico o relacionado con la salud que sea dañino o engañoso. En respuesta a la pandemia del COVID-19, nuestros equipos trabajaron para acabar con diversos abusos nuevos, como intentos de phishing, malware, teorías conspirativas peligrosas y fraudes.

Abordamos esta cuestión en todos nuestros procesos y servicios, desde la actualización de las políticas hasta su implementación obligatoria.

En YouTube, por ejemplo, este reto se afrontó con una política específica sobre la desinformación médica respecto al COVID-19, basada en las recomendaciones de las autoridades sanitarias. Esta política específica deja claro que no se permitirá en la plataforma contenido que suponga un grave riesgo de perjuicio significativo para los usuarios, como contenido que contradiga la información sanitaria de alta calidad de las autoridades sanitarias locales y de la Organización Mundial de la Salud (OMS). La adopción de esta política permitió a los equipos de cumplimiento tomar medidas para ayudar a proteger a la comunidad de YouTube. Por ejemplo, en los primeros 18 meses de la pandemia, YouTube eliminó más de un millón de vídeos relacionados con información peligrosa sobre el coronavirus, como falsas curas o afirmaciones de que la enfermedad no existe. (El enfoque de Google para luchar contra la desinformación online, s.f.)

Estos procesos automatizados desde los inicios de *Google* en 1994 han coadyuvado a brindar una mayor certeza, objetividad y claridad en los contenidos que son indexados como positivos y que por lo tanto son expuestos a los usuarios que realizan búsquedas a través de su buscador en línea (www.Google.com).

Asimismo, la indexación de las páginas de internet incluyendo videos de *YouTube* le han permitido a *Google* generar un acervo de datos de magnitudes sin precedentes. Esto contribuye directamente al entrenamiento de los modelos de *GenAI* que necesitan grandes cantidades de datos (idealmente previamente filtrados o curados de contenido falso, nocivo, ilegal, inmoral, etc.) para el correcto entrenamiento de sus modelos y tener una mayor certeza en los nuevos contenidos producidos. Esto le da a *Google* una ventaja competitiva enorme en comparación de otras empresas que ofrecen servicios de *GenAI* para empresas.

Entonces se puede entender que los acervos de datos son una de las claves para poder seguir desarrollando y entrenando modelos de *GenAI*, tal como lo menciona Aarón Mok en su artículo llamado “*Gigantes de la IA como OpenAI y Anthropic están luchando por conseguir suficientes datos para entrenar modelos*” (Mok, 2024),

Empresas como *OpenAI* y *Anthropic* están luchando por conseguir uno de los recursos más valiosos de la IA: datos confiables. Ese déficit podría obstaculizar el desarrollo de grandes modelos de lenguaje que impulsen sus chatbots a medida que se intensifica la carrera por crear los mejores productos en el sector en crecimiento. (Mok, 2024)

Es por ello que empresas que no tienen un acervo de datos con un volumen considerable, con datos históricos, recientes y fiables se encuentran en una clara desventaja para innovar y brindarle a sus clientes modelos cada vez más potentes, competitivos y sin alucinación. La “alucinación” es un efecto que se da cuando un modelo de *GenAI* típicamente de texto o de video, hace suposiciones carentes de fundamentos o datos duros. IBM lo explica de la siguiente forma en su artículo:

La alucinación por IA es un fenómeno en el que un modelo de lenguaje grande (LLM, por sus siglas en inglés), a menudo un chatbot de IA generativo o una herramienta de visión por computadora, percibe patrones u objetos que son inexistentes o imperceptibles para los observadores humanos, creando resultados que no tienen sentido o son completamente inexactos. (¿Qué son las alucinaciones de la IA?, 2024)

Por poner un ejemplo, si un usuario le preguntara a un *chatbot* con un bajo nivel de confiabilidad ¿Quién ganó las elecciones presidenciales en México? El mismo día de las elecciones y antes que se conozcan los resultados oficiales, el modelo de *GenAI* puede inferir que en base a que una candidata o candidato encabezaba las encuestas, seguramente habría sido la ganadora o ganador y por ende podría dar un resultado subjetivo o carente de fundamentos reales. Como se puede imaginar, esto representa un problema grave para la *GenAI* pues reduce la fiabilidad de los resultados obtenidos.

Fabricando chips especializados para la Inteligencia Artificial

Nvidia es una empresa que se ha caracterizado por desarrollar y producir chips de video (conocidos como GPU – *Graphic Processing Unit*) para equipos de cómputo que van desde equipos celulares, tabletas, computadoras y servidores a gran escala. Algunos de estos chips tienen aceleración gráfica que contienen algoritmos que favorecen a la optimización de las operaciones matemáticas necesarias para realizar tareas o procesos de *GenAI*. Es por ello que desde el auge de *GenAI* a finales del 2022, Nvidia ha estado posicionado como el principal proveedor de estos chips hacia todas las empresas que ofrezcan servicios de *GenAI* para la venta y consumo de sus clientes. Tal como lo explica Rob Starks Jr. en su artículo (Jr, 2023):

Las acciones de *Nvidia* (NVDA 2,57%) han estado en alza en los últimos meses, recuperándose de un terrible 2022. Las acciones de la compañía han subido más del 194% desde principios de año, superando recientemente una capitalización de mercado de 1 billón de dólares y cotizando en su nivel más alto jamás visto, impulsado por un auge de la inteligencia artificial (IA).

Wall Street se ha dado cuenta de que la sociedad está entrando rápidamente en la era de la IA, ya que esta tecnología se está volviendo más sofisticada y potencialmente puede revolucionar muchas industrias. Dado que *Nvidia* diseña un componente esencial que impulsa el entrenamiento y el uso de la IA, la empresa está bien posicionada para ser uno de los primeros ganadores más destacados en el espacio de la IA. (Jr, 2023)

En la Ilustración 2 se puede ver como al cierre de la bolsa el 24 de mayo del 2024, la acción de *Nvidia* ya cotizaba en \$1,064.69dls. Tomando como punto de referencia el 30 de septiembre del 2022 - antes del anuncio de *ChatGPT*, la acción de *Nvidia* cotizaba en \$121.39dls. Esto nos da una referencia de lo valiosos que han sido estos GPUs para el crecimiento de *Nvidia* pues ha sido el principal motor del crecimiento de casi 900% a la fecha del 24 de mayo de 2024.

Ilustración 2 en donde se ve el comportamiento histórico de la acción de Nvidia (NASDAQ, 2024)



Sin embargo, *Google* ha comenzado a diseñar y fabricar chips especializados para el uso de IA desde ya hace algunos años y en el año 2024 ya ofrecían la 5ta versión de estos chips. *Google* ha acuñado el término “Unidad de Procesamiento Tensorial” (o TPU por sus siglas en inglés). *Google* los describe como “aceleradoras de IA con diseño personalizado que están optimizadas para el entrenamiento y la inferencia de grandes modelos de IA. Son ideales para una variedad de casos de uso, como chatbots, generación de código, generación de contenido multimedia, voz sintética, servicios de visión, motores de recomendaciones y modelos de personalización, entre otros.” (Cloud, Acelera el desarrollo de IA con Google Cloud TPU, 2024).

¿Cuál es la diferencia entre las TPU y las GPU?

Cabe resaltar que las TPU no están disponibles para su venta comercial. Es decir, se emplean en servidores de la nube de *Google* (Plataforma de la nube de *Google* o GCP por sus siglas en inglés) y los clientes pueden pagar por su uso (similar a pagar por una renta del hardware en base a su tiempo de uso). *Google* establece que la diferencia entre las TPU y las GPU es que “Las Cloud TPU están

optimizadas para el entrenamiento de modelos de aprendizaje profundo grandes y complejos que presentan muchos cálculos de matrices, por ejemplo, la creación de modelos de lenguaje extenso (LLM). Las Cloud TPU también tienen SparseCores, que son procesadores de Dataflow que aceleran los modelos que se basan en incorporaciones que se encuentran en los modelos de recomendación. Otros casos de uso incluyen la atención médica, como el modelado de plegamiento de proteínas y el descubrimiento de fármacos.”

Por lo anterior considera entonces que *Google* tiene la ventaja de los acervos de datos de un tamaño significativo debido a la indexación de páginas de internet y videos de *YouTube* desde hace ya varios años, así como también el diseño y fabricación de los chips de hardware necesarios para correr procesos de GenAI en su plataforma de la nube para la venta de servicios. A esto se le conoce como “tener una solución vertical completa (*full stack*)” que va desde el hardware, los datos, los servicios y los modelos de GenAI lo cual lo pone como uno de los principales competidores en la oferta de servicios de GenAI en el mundo.

Estas dos ventajas competitivas fueron mencionadas por un vocero de Microsoft en el mes de marzo del 2024 (Revell, 2024)

Hoy en día, sólo una empresa -Google- está integrada verticalmente de una manera que le proporciona fuerza e independencia en cada capa de IA, desde chips hasta una próspera tienda de aplicaciones móviles. Todos deben confiar en las asociaciones para innovar y competir.
(Revell, 2024)

Análisis de mercado

A pesar de que la GenAI ha ido evolucionando desde hace varios años, las oportunidades de negocio comenzaron a ganar mayor interés a nivel mundial a raíz de que una empresa llamada “OpenAI” puso a disposición del público un chat interactivo basado en un modelo de lenguaje conocido como “ChatGPT” el 30 de noviembre de 2022.

“Seis días después del lanzamiento, Sam Altman comunica que ChatGPT alcanzó el millón de usuarios.” (Edwards, 2023) y *“ChatGPT se convierte en la aplicación de consumo de más rápido crecimiento de todos los tiempos, logrando los 100 millones de usuarios mensuales.”* (Edwards, 2023)

Poco después, en enero del 2023, Microsoft invirtió \$10Mmdd (Forbes, 2023) para asociarse con *OpenAI* e integrar este modelo de *GenAI* en muchos de sus productos (Bing, Office, etc.) en la versión empresarial que ahora se conoce como *Copilot*.

Y esto es solo un ejemplo del auge que se ha generado por uno los principales competidores dentro de la industria de *GenAI*. Sin embargo, otros competidores como *Google* (Dastin, 2023), *Meta* (Izacard, 2023), *HuggingFace*, *DeepMind* y muchas otras empresas han desarrollado sus propios modelos y los han puesto a disposición del público ya sea de forma gratuita o con un costo adicional en el último año.

En diciembre del 2023, IDC (Shirer, 2023) estimó que las empresas de IT habrían invertido más de \$19.4Mmdd en soluciones de *GenAI* para finales del 2023 y se proyecta que estas inversiones podrían alcanzar hasta \$151.1Mmdd para el 2027. Este crecimiento se debería a la creciente demanda de modelos, herramientas y datos a través de un sinnúmero de industrias.

Así también lo pone de manifiesto la “*Propuesta de agenda nacional de Inteligencia Artificial para México (2024-2030)*” (Lagunes A., 2023, p. 15), pues menciona que “*De acuerdo con el Dr. Ryan Abbot, se espera que para 2030, la tecnología basada en la IA contribuya con hasta 15.7 billones de dólares a la economía global*” 2023), una cifra que supera la producción actual combinada de China e India, resaltando el inmenso potencial de esta tecnología para transformar las economías a nivel mundial.”

MARCO CONCEPTUAL

Metodología

En el marco del presente análisis, se utilizaron las herramientas para el análisis de los indicadores micro y macroeconómicos, así como sus efectos en la organización, sociedad y economía aprendidas en la materia “Entorno Económico” impartida por el Mtro. Francisco Núñez en el ITESO para llevar a cabo el trabajo de búsqueda y análisis presentado en la sección del “Análisis de Mercado” del presente documento.

- Estrategia de búsqueda

Para obtener información precisa y actualizada, se priorizó la búsqueda en medios electrónicos. Esto permitió enfocarse en artículos relevantes publicados en los últimos 12-15 meses, período en el que se han desarrollado las tecnologías objeto de estudio.

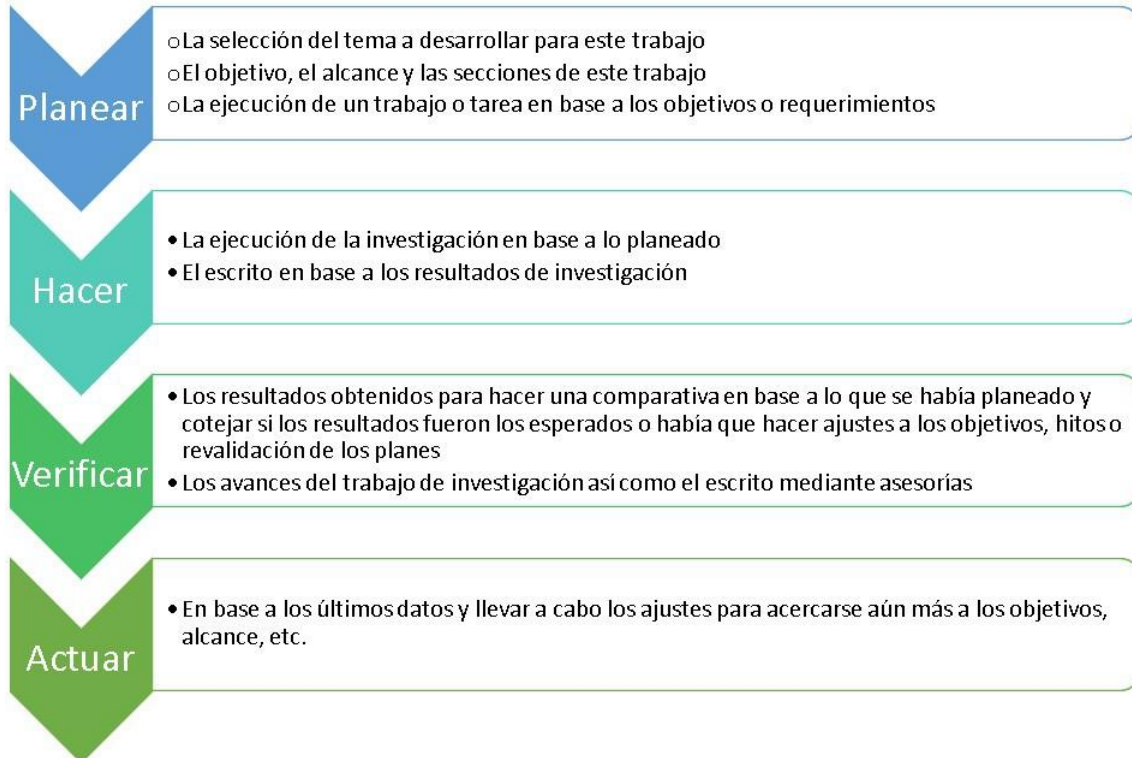
- Recursos utilizados:
 - Bases de datos académicas - Acceso a publicaciones científicas y estudios especializados.
 - Sitios web de empresas e instituciones - Información sobre tendencias del mercado y casos de éxito.
 - Revistas y blogs especializados - Análisis y opiniones de expertos en el sector.
 - Limitaciones de las publicaciones impresas - Si bien se consultaron libros y publicaciones impresas, estas no resultaron de gran utilidad debido a la naturaleza reciente de las tecnologías analizadas. La rápida evolución de este campo hace que la información impresa se torne obsoleta con relativa rapidez.
 - Resultados y beneficios - El enfoque en medios electrónicos permitió obtener una visión actualizada y completa del mercado, incluyendo:
 - Tendencias emergentes - Identificar las últimas tecnologías y sus aplicaciones.
 - Competencia en el mercado - Conocer los principales actores y sus estrategias.

- Casos de éxito - Aprender de las experiencias de empresas que han implementado estas tecnologías.

Tanto en la carrera de Ingeniería en Sistemas Computacionales como en la Maestría en Administración de Empresas (ambas en el ITESO), particularmente en la materia “Modelos de Calidad y Productividad” impartida por el Mtro. Jaime Borrego Navarro, el autor de este trabajo aprendió y mejoró su entendimiento de:

- El ciclo de Deming (Deming, s.f.) - Una de las metodologías fundamentales con las que se apoyó el autor del presente trabajo para sobresalir en el ámbito profesional (especialmente como Gerente de Proyectos, rol en el que se ha desempeñado desde el 2007 al 2024). El ciclo de Deming es una metodología creada por Peter Deming y también es conocida como el ciclo PHVA (o PDCA por sus siglas en inglés *Plan, Do, Check, Act*) que se refiere al ciclo constante de planeación, ejecución, validación y corrección a lo largo de un proyecto.

En la ilustración 3 se muestra el proceso de Deming de forma secuencial y a mayor detalle.



El beneficio de la aplicación de este ciclo constantemente lleva a una mejora continua y a una predictibilidad de los resultados obtenidos de forma eficiente, así como una comunicación más efectiva y el establecimiento de un plan y avances de trabajo en todo momento, brindando mayor claridad y entendimiento para el asesor del presente trabajo – el Mtro. José Juan Calzada - así como posteriormente para las partes interesadas en este trabajo.

Para elaborar el presente proyecto, primero se generó un anteproyecto en base a un tema de relevancia actualmente como lo es *GenAI*, se estableció un Plan de trabajo con la supervisión del asesor con fechas de entrega detalladas así como el índice del proyecto para después ejecutar el desarrollo (Hacer) en base al plan de trabajo y así obtener retroalimentación (Verificar) y Actuar en base a la retroalimentación obtenida durante nuestras sesiones de asesoría para después hacer las correcciones tanto al plan de trabajo así como al documento y al trabajo en sí.

- Diagramas de flujo. Esta herramienta gráfica ayuda a comprender el flujo de información, de procesos y de las partes involucradas para brindar de forma rápida, clara y concisa para ayudar al lector a tener un entendimiento sólido acerca de los procesos, trabajos o tareas. Esto ha resultado fundamental para definir procesos nuevos dentro de los proyectos que han estado a cargo del autor, en algunos casos involucrando varios millones de dólares en inversión y en remuneración para las empresas o clientes con los que ha laborado.

El autor aprendió el valor de analizar Casos Prácticos de Éxito en la materia “Visión Estratégica” impartida por el Mtro. Luis Edmundo Garrido Sánchez en el ITESO, lo que le ayudó en la elaboración del presente trabajo particularmente para la selección de los casos de éxito más relevantes, claros y que se pudieran ligar con los objetivos de este trabajo.

Herramientas de cómputo utilizadas en la elaboración del presente trabajo

- Microsoft Word para la elaboración del presente trabajo, así como las revisiones y discusiones del mismo.
- Microsoft Power Point para la elaboración de algunas ilustraciones como el análisis PESTEL de *Google*, el círculo de Deming, diagrama de flujo, etc.
- Navegador de Internet llamado *Chrome* (Google, Chrome, s.f.) para la búsqueda y análisis de publicaciones relacionadas al presente trabajo.

- El uso del *chatbot Gemini* (Google, Gemini, s.f.) de *Google* para identificar mejoras en la redacción de las palabras del autor de este trabajo.
- El uso de correo electrónico para mantener la comunicación con el asesor del presente trabajo y sus revisiones.

¿Cómo funciona la Inteligencia Artificial Generativa (*Generative AI* o *Gen AI*)?

Como se ha evidenciado en el análisis de mercado, *GenAI* se ha convertido en una nueva rama de la Inteligencia Artificial que está revolucionando varias industrias. ¿Pero cómo funciona exactamente? A continuación, se presenta el proceso mediante el cual *Google* crea nuevo contenido a través de *GenAI* tomando como base un acervo de datos preexistente.

1. *Adquisición de datos y preprocesamiento*

Primero, un modelo de lenguaje analiza un conjunto de datos (*data set*) para deducir y sintetizar la información solicitada en cuestión de segundos o fracciones de segundo y entregar un resultado similar al que pudiera producir otro ser humano. Estos datos deben de estar en formato digital (texto, audio, video, código, etc.). Esto es fundamentalmente distinto de lo que hacen buscadores de internet que entregan miles (o millones) de resultados en base a una búsqueda sin llevar a cabo un procesamiento y síntesis de dichos resultados para su revisión y aplicación o uso.

2. *El procesamiento de la información en base a la composición de los Modelos*

Una vez obtenidos los datos y habiendo hecho un preprocesamiento, se usa alguna de las diversas arquitecturas de *GenAI* (algunas de las más comunes se listan en los siguientes párrafos) y cada una está diseñada para un objetivo del usuario en específico, por lo que tendrá ventajas y desventajas dependiendo de cómo se emplee dicho modelo. Por ejemplo, un modelo de lenguaje largo (*Large Language Model* – LLM) sirve para análisis, procesamiento y resumen de textos muy largos como un libro. Mientras que un modelo desarrollado específicamente para el análisis de videos, no podría analizar el texto en un libro de forma efectiva, lo cual es una desventaja si se quiere emplear el modelo incorrecto para una finalidad distinta para la cual fue creado.

A continuación, un listado de dos de los modelos más populares y sus arquitecturas:

- Redes Generativas Adversas (comúnmente referidas como *Generative Adversarial Networks* o *GAN* por sus siglas en inglés):

Imagina que hay dos modelos jugando un videojuego. Un modelo (generador) intenta crear nuevos datos que sean indistinguibles de datos reales. El otro modelo (discriminador) trata de identificar los datos que son falsos. A través de esta especie de competencia entre ambos modelos se van mejorando el uno al otro y esto lleva a obtener datos más refinados y veraces.

- Encodificadores Automáticos Variables (*Variational Autoencoders* – *VAEs* por sus siglas en inglés):

Estos modelos codifican los datos para generar una representación resumida de las funciones principales de los datos. El decodificador utiliza estos datos para generar nueva información que capture la esencia de los datos originales y después lo compara contra la información original para hacer el refinamiento de la información.

3. *Entrenar un Modelo*

Una vez que la información ha sido procesada por alguno de los Modelos mencionados en los pasos anteriores, se procede al entrenamiento del modelo. Esto quiere decir que mediante un proceso iterativo (cíclico o repetitivo), el modelo ajusta sus parámetros internos para poder generar información más refinada y certera. Mediante este proceso el modelo se vuelve cada vez más preciso y confiable. Este es el paso más laborioso, tardado y costoso para las compañías, pues entrenar modelos puede tomar desde varios minutos hasta varias horas o días.

4. *Generación de nuevos datos*

Una vez que se entrena el modelo de *GenAI*, puede generar nuevos datos basado en la información que ha aprendido. El tipo de información que genere dependerá de la composición o arquitectura de los modelos y de la calidad de la información que se le suministre en el conjunto de datos. De esta forma podemos concluir que la fiabilidad de la información que generen los modelos, dependerá

directamente de la cantidad de iteraciones de entrenamiento por las cuales fue sometido dicho modelo, la calidad y la cantidad de los datos proporcionados para su análisis inicial.

El proceso mencionado anteriormente es generalizado. En *Google Vertex AI* (AI, Descripción general de la IA generativa en Vertex AI, 2024) se agregan algunos pasos y componentes adicionales para llevar a cabo un saneamiento de la información. Esto se hace principalmente para ofrecer una solución a nivel empresarial – es decir, que las empresas y compañías que deseen hacer uso de estas tecnologías pueden estar seguras de que se toman las debidas precauciones de seguridad, privacidad, apego a regulaciones en las distintas regiones del mundo, etc. para ofrecer un servicio de mayor calidad. Estos pasos o componentes son (AI, Descripción general de los fundamentos, 2024):

- Fundamentación (*Grounding*)

Se refiere al proceso de habilitar a los modelos de lenguaje (texto) de información específica del caso de uso o del contexto. Por ejemplo, un modelo de lenguaje sin fundamentación podría generar la descripción general de un gato, pero no podría describir las características específicas de un gato en específico. La descripción de “fundamentación” según *Google* es (AI, Descripción general de la IA generativa en Vertex AI, 2024):

Si necesitas que las respuestas de los modelos se basen en una fuente de información, como tu propio corpus de datos, puedes usar la fundamentación en la IA generativa en Vertex AI. La fundamentación ayuda a reducir las alucinaciones del modelo, en especial en temas desconocidos, y también le proporciona al modelo acceso a información nueva. (AI, Descripción general de la IA generativa en Vertex AI, 2024)

- Validación de Citas

Una de las claves para tener veracidad en los resultados de los nuevos contenidos generados, es conocer la fuente de la información, además de que es necesario verificar que las fuentes de donde viene la información sean debidamente acreditadas a los autores de estas.

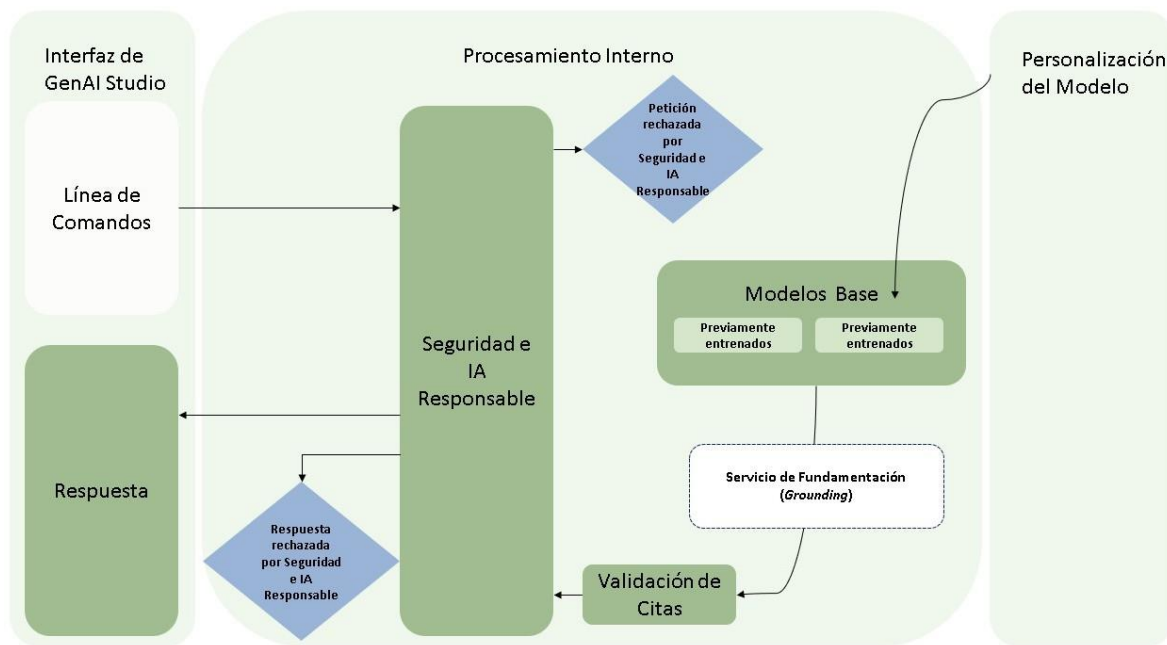
La descripción del servicio de “Validación de Citas” según *Google* es (AI, Descripción general de la IA generativa en Vertex AI, 2024):

Después de que se genera la respuesta, la IA generativa en Vertex AI verifica si las citas deben incluirse en la respuesta. Si una cantidad significativa del texto en la respuesta proviene de una fuente en particular, esa fuente se agrega a los metadatos de cita en la respuesta. (AI, Descripción general de la IA generativa en Vertex AI, 2024)

Existen otras opciones que incluso se ofrecen sin costo (Por ejemplo, HuggingFace, ChatGPT, entre otras), pero que no garantizan la seguridad, el servicio de soporte o la privacidad de la información del cliente al mismo nivel que lo hacen otras compañías, por lo que quizá no sean recomendables para usos corporativos debido a los riesgos inherentes que esto puede ocasionar.

A continuación, se muestran los componentes y el flujo de datos para el procesamiento de un modelo de *GenAI*, particularmente por el *chatbot Gemini* de *Google* (AI, Descripción general de la IA generativa en Vertex AI, 2024). Es importante resaltar que en el diagrama se ven tres componentes que no se utilizan en muchos de los *chatbots*, son “Seguridad e IA Responsable”, “Servicio de Fundamentación” y “Validación de Citas”. Estos elementos son críticos porque sin ellos, se omite cualquier relación entre el contenido entregado por la respuesta (procesamiento) del modelo y las fuentes bibliográficas de donde se obtuvo el contenido. Esto es grave, ya que ha derivado en varias demandas en contra de *ChatGPT* por omisión de los derechos de autor. En uno de los casos más sonados, el demandante por este motivo es el New York Times, en donde “*millones de artículos del New York Times fueron utilizados para entrenar chatbots que ahora compiten con el diario*” (Mac, 2023)

En el diagrama de la Ilustración 4 se muestran los diferentes componentes que conforman una solución de *GenAI* para un nivel Corporativo, específicamente ofrecido por *Vertex AI*.



Como se detalla en la sección del Análisis de Mercado, empresas como *Google*, *Microsoft*, *Meta* (antes conocida como *Facebook*) o *Nvidia* -por nombrar algunas- han apostado por invertir grandes cantidades de recursos a la investigación y desarrollo de modelos que ponen a disposición de sus clientes mediante sus plataformas, principalmente en la nube.

Contexto de *GenAI* como negocio

Los proveedores de servicios en la nube como *Google*, *Microsoft*, *Amazon*, entre otros, están apostando por atraer a los clientes que buscan beneficiarse de la incorporación de modelos y herramientas de *GenAI* en sus empresas para aumentar sus ventas, rentabilidad, seguridad, eficiencias, etcétera.

Vertex AI de *Google* desarrolla herramientas, aplicaciones y servicios que ponen a disposición del cliente desarrollar sus propios modelos de Inteligencia Artificial en la nube, entrenar y ejecutar sus propios modelos, así como los que ofrece *Google* y otros repositorios (*HuggingFace*, *GitHub*, etc.) además de poder realizar la evaluación de los resultados de sus modelos para determinar la eficacia de sus resultados y poderles hacer los ajustes que el cliente determine necesarios.

La plataforma de *Vertex AI* fue lanzada al mercado el 18 de mayo del 2021 (Cloud, Google Cloud Launches Vertex AI, Making Machine Learning More Accessible and Useful For Developers and Businesses, 2021) para poner a disposición de los clientes todo el potencial de los modelos de Inteligencia Artificial que ha desarrollado *Google* a lo largo de los años. Tal ha sido el éxito de *Vertex AI*, que ha alcanzado a crecer en uso “*casi 6 veces entre el primer semestre del año (2023) en comparación con el segundo semestre*” dijo el CEO de *Google*, Sundar Pichai (Dignan, 2024).

Específicamente, los clientes de *Vertex AI* buscan hacer uso de las tecnologías que ofrece la Inteligencia Artificial para ser implementadas en sus empresas. Los clientes van desde usuarios independientes, hasta empresas de talla mundial que van desde el ramo financiero, ventas (*retail*), mercadotecnia, seguridad, logística, hardware, etcétera.

Roles y Responsabilidades

Dentro de *Vertex AI*, existen distintos roles para satisfacer las necesidades de los clientes, a continuación, el autor explica a grandes rasgos en qué consiste cada uno de estos roles y cómo es la cadena de responsabilidades en los proyectos para el desarrollo de software, incluyendo *GenAI*:

Gerente de Producto (Product Manager - PM)

Es responsable del acercamiento a los clientes y al constante análisis de la industria para determinar los requerimientos y estrategias de los próximos servicios, productos o funcionalidad que se deberán desarrollar para satisfacer las necesidades de los clientes o innovar en la industria. Una de las principales tareas de los Gerentes de Producto después del análisis de los requerimientos del cliente o las tendencias de la industria, es la creación de un documento de requerimientos en donde se establecen detalladamente las experiencias del usuario que se quieren satisfacer. Una experiencia del usuario puede ser, por ejemplo: “*Al abrir un documento de texto, el usuario podrá editarlo de forma simultánea y colaborativa con otros usuarios a los cuales se les haya dado acceso a dicho documento.*”

Diseñador de XU (UX Designer - UXD)

En dicho caso, XU se refiere a la eXperiencia del Usuario. Los diseñadores de experiencias del usuario establecen los elementos visuales, interactivos, y la secuencia de acciones que deberá realizar

el usuario en una interfaz gráfica para que sea de forma intuitiva, eficiente y amigable. Esto ayuda a reducir las frustraciones o fricciones de parte del usuario, aumenta la satisfacción del cliente (*Customer Satisfaction*) hacia el servicio recibido, lo cual suele redituarse en mayor inversión y uso del bien o servicio.

Gerente de Proyectos (Program Manager - PgM)

Encargado de establecer un plan de proyecto, un equipo de trabajo, realizar un plan de gestión de riesgos, monitorear el progreso del proyecto y mantener una comunicación clara, precisa y oportuna sobre el proyecto hacia todas las partes involucradas e interesadas en que el proyecto llegue a su culminación satisfactoria. Básicamente, coordinar que se ejecuten las acciones del círculo de Deming.

Ingeniero de software (Software Engineer - SWE)

Encargados de codificar las soluciones de software para satisfacer los requerimientos establecidos por el Gerente de Producto. Tienen un amplio conocimiento de las herramientas de software para desarrollar las nuevas funciones, bienes o servicios de la forma más eficiente, segura, de calidad y en el tiempo establecido.

En la Ilustración 5 se muestra el flujo secuencial típico durante el curso de un proyecto de desarrollo de software, incluyendo el lanzamiento de nuevos modelos y aplicaciones referentes a *GenAI* en base a la experiencia profesional del autor de los últimos 18 años.

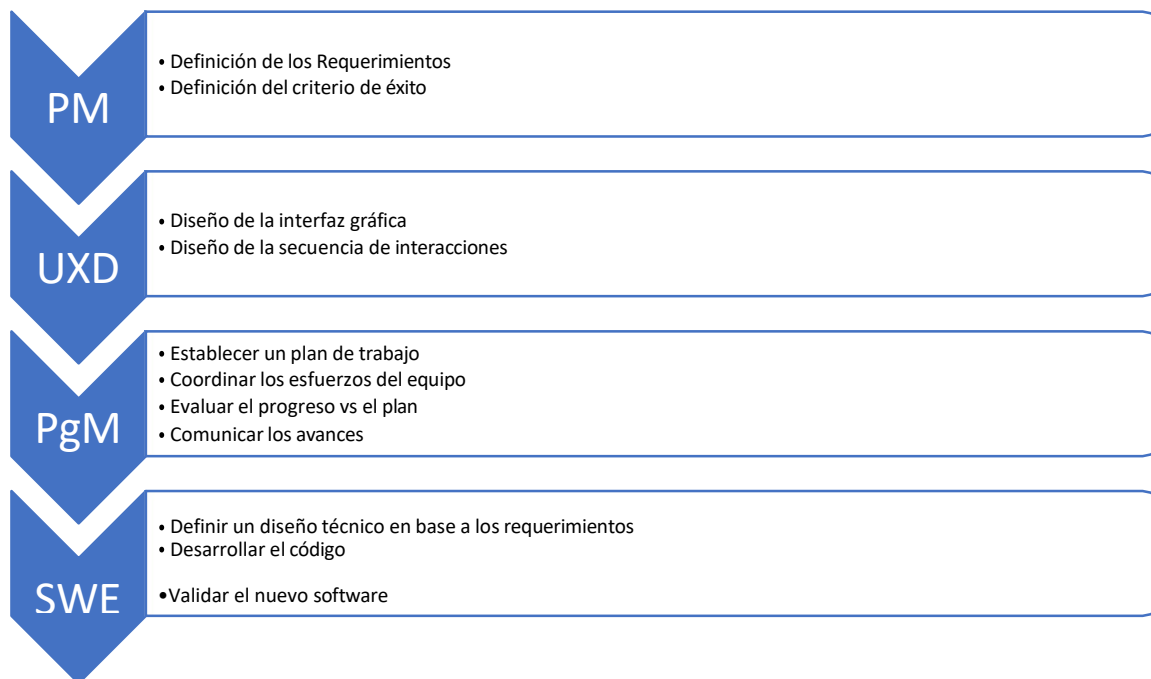


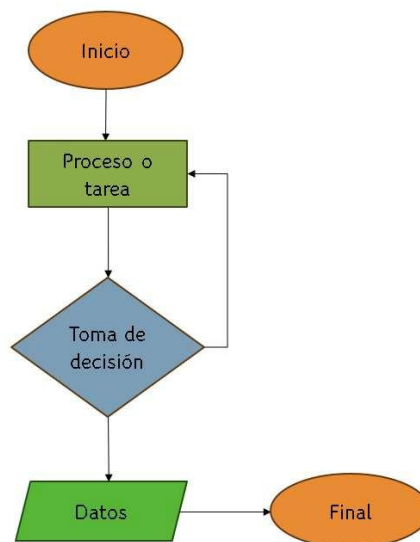
Diagrama de Flujo de datos como una herramienta fundamental

Esta herramienta ofrece un “*Método estructurado para documentar gráficamente un proceso como un flujo de pasos sucesivos y alternativos, el "proceso de diagrama de flujo", fue expuesto por Frank Gilbreth, en la Sociedad Americana de Ingenieros Mecánicos (ASME), en 1921*” (Varios, 2024)

Los diagramas de flujo son utilizados comúnmente dentro de los proyectos, específicamente para entender los Diseños de XU y los documentos de definición técnicos en donde se establece el flujo de información desde el punto de origen (por ejemplo, el usuario navegando dentro de una página), cómo fluyen los datos a través de los distintos componentes (red, bases de datos, cortafuegos, balanceador de cargas, etc.) hasta llegar al fin de su flujo, típicamente otorgando una respuesta al usuario o modificando los datos mediante su procesamiento y por último almacenamiento del resultado.

Estos diagramas ayudan a dar mayor claridad al desarrollador y a los gerentes del proyecto, acerca de los diferentes componentes que van a manipular la información y así buscar eficiencias como por ejemplo, reducir la redundancia de los elementos como tener dos bases de datos que almacenan la misma información pero que se encuentran en secuencia (primero se almacena la información en la base de datos 1, después en la base de datos 2) y esta redundancia genera ineficiencia en el flujo de datos debida la demora y el costo financiero y operativo que conlleva la gestión de dos bases de datos.

En la Ilustración 6 se muestra un ejemplo general de esta herramienta para brindar una mayor claridad acerca del flujo del proceso o la información. Existe un mayor número de símbolos utilizados para representar diagramas de flujo, pero los que se presentan a continuación son los más comunes.



PROPUESTA DEL AUTOR

Esta propuesta se fundamenta en una sólida base de investigación y análisis exhaustivo por parte del autor, realizada tanto en medios digitales como a través de su amplia experiencia profesional. Su trayectoria de más de 5 años en el campo de la Inteligencia Artificial le ha permitido desarrollar un profundo conocimiento y habilidades en el desarrollo e implementación de soluciones de IA.

Adicionalmente, cuenta con más de 18 años de experiencia como Gerente de Proyectos, lo que le proporciona una capacidad excepcional para liderar y gestionar proyectos de manera efectiva, asegurando su éxito y cumpliendo con los objetivos establecidos.

¿Qué se espera que suceda en base a esta propuesta?

Dar a conocer las tendencias de oportunidades de negocio en el uso de la *GenAI*, así como los riesgos existentes y potenciales sobre un rezago tecnológico y de especialización laboral en Jalisco, México. Particularmente, este trabajo se concluyó el 31 de mayo del 2024 y se prevé presentar dicha propuesta al Dr. Luis Fernando Gutiérrez Preciado coordinador de Ingeniería en Sistemas Computacionales (ISC) en el ITESO y al Dr. José Francisco Cervantes Álvarez, quien se ha dado a la tarea de investigar los impactos y áreas de oportunidad en el uso de tecnologías de *GenAI*.

A través de compartirle el presente trabajo por la vía de correo electrónico a sistemas@iteso.mx y a fcervantes@iteso.mx, se espera influir en el currículo de estudios de la carrera universitaria de ISC y de otras licenciaturas o posgrados que resulten apropiados.

¿En cuánto tiempo se tienen que tomar las acciones?

Debido a que *GenAI* está avanzando a un ritmo veloz, es conveniente empezar a tomar acciones a la brevedad. Por lo que se ha compartido el presente trabajo en el mes de junio del 2024 y el autor se mantuvo a disposición de los Dres. Luis Fernando Gutiérrez y Francisco Cervantes para mantener la comunicación abierta.

Análisis de la problemática en Jalisco

Tomando en consideración que los avances en *GenAI* apenas están siendo explorados, ya existen diversos retos que se deben considerar. En particular, se considera que la siguiente lista de problemáticas es de preocupación dentro del contexto actual por el que atraviesa México y en particular, Jalisco:

Retos Morales y Éticos

Comúnmente conocidos como *Deepfakes*, las alteraciones de imágenes, videos y audio pueden contribuir a la alarmante tendencia del robo de identidades, extorsiones y a la desinformación. Esto es aún más preocupante ya que *GenAI* puede generar medios audiovisuales sumamente realistas en cuestión de segundos y que pueden ser utilizados con fines dañinos o ilegales.

“Se viraliza una foto del Papa Francisco creada con una herramienta de inteligencia artificial.” (Scheffler, 2023), misma que se puede observar en la Ilustración 7, en donde, mediante el uso de *GenAI*, se generó una imagen falsa en donde aparece el papa Francisco vistiendo una chamarra de la exclusiva marca Balenciaga, pero a pesar del realismo, la imagen es falsa.



Debido a que los modelos de *GenAI* se basan en el análisis de datos generados a base de artículos, información, consultas, etc. generados en ciertos sectores o regiones de la población, los resultados pueden estar sesgados y pueden llevar a la discriminación de ciertos grupos de personas, países,

regiones y perpetuar estereotipos. Como lo escribe Zeeshan Aleem (Aleem, 2023) en su artículo de MSNBC, *“Los datos más recientes del organismo de control de los medios y la desinformación NewsGuard identifican más de 600 “sitios de información y noticias poco confiables generados por IA” en Internet. Según el Washington Post, se trata de un aumento de más del 1000% desde mayo”* el artículo continúa y da un ejemplo de una noticia falsa generada por GenAI:

Un video afirma ridículamente que el expresidente Barack Obama asesinó a uno de sus ex chefs de la Casa Blanca, y el video utiliza tecnología de inteligencia artificial para evocar una declaración de "Obama" generada por computadora en respuesta a la historia falsa. (Aleem, 2023)

Esto se puede prestar a esparcir noticias falsas, rumores o modificar la percepción de la gente en entornos electorales como en el actual año en donde coinciden las elecciones presidenciales en México y Estados Unidos, ya sea mediante encuestas falseadas o truqueadas utilizando GenAI para generar una percepción errónea de que sus candidatos aventajan en los resultados de las encuestas o generando noticias falsas sobre sus partidos o candidatos opositores para confundir al padrón electoral.

En la última década se habían estado utilizando los denominados “bots” (proveniente de la palabra *robot*) en redes sociales para repetir noticias falsas o tomar acciones específicas y predeterminadas (por ejemplo, dar “me gusta” a ciertas publicaciones o videos para tener injerencia en la viralización de los mismos) o esparcir un mensaje generalmente falso dentro de las redes sociales que pudieran desfavorecer la imagen de una persona, institución o ideología.

Dichos *bots* han sido controlados por individuos que ofrecen estos servicios a cambio de una remuneración monetaria y aunque su alcance sea considerable, han tenido limitantes pues suelen ser detectados con relativa facilidad debido a las tendencias de sus participaciones en redes sociales o que sus conversaciones fueran limitadas para responder interactivamente y poder engañar a otros seres humanos. Sin embargo, como lo detalla esta infografía (Agency, s.f.) de la *Cybersecurity and Infrastructure Security Agency* (CISA, por sus siglas en inglés), estos *bots* pueden ser controlados por GenAI y así poder enmascarar mejor su identidad y tener interacciones muy parecidas a las de un ser humano, lo que dificultaría identificarlos como *bots*.

Retos Laborales

El ser humano tendrá una desventaja si se le compara con la eficiencia del análisis de la información que el mismo ser humano ha desarrollado en su historia para producir resultados más refinados y eficientes. Ningún ser humano es capaz de memorizar, analizar y discernir toda esta información en fracciones de segundos para generar una recomendación o resolución objetiva. Por lo que puede parecer una amenaza a profesionistas en varias industrias.

Inclusive en el 2023 se suscitó una huelga de 118 días de actores y escritores (*Screen Actors Guild y Writers Guild of America*) en Hollywood, California en donde parte de las demandas de los huelguistas, eran las garantías de que la Inteligencia Artificial no haría uso de las imágenes, videos y grabaciones de voz de los actores y actrices para generar películas con alto grado de realismo sin tener que pagar por los derechos de uso de las personas.

Esta huelga prácticamente paralizó a la industria cinematográfica de Hollywood de más de \$1'300,000 millones de dólares por 118 días (Angela Watercutter, 2023) hasta que lograron llegar a un acuerdo con los estudios cinematográficos para establecer un marco legal bajo el cual se regirán los límites del uso de *GenAI* para mitigar el impacto a los escritores y actores.

Otro ejemplo es el de la conducción autónoma de vehículos en donde en años recientes, se ha generalizado el temor de que los vehículos autónomos podrían sustituir a los choferes de transportes de carga debido a que es un trabajo extenuante físicamente y que puede generar altos riesgos debido al cansancio de los conductores ya que son expuestos a largas jornadas de manejo y al estrés para hacer sus entregas en tiempo.

En la Ilustración 8 se puede apreciar un vehículo desarrollado por *Waymo*, propiedad de *Alphabet*, que habilita la conducción de forma autónoma. (Doubling down on Waymo One, 2023). En el año 2024 se les puede ver circulando por las calles en algunas de las ciudades de Estados Unidos, pero únicamente realizan labores de pruebas para su investigación y desarrollo, no se encuentran a la venta al público.



Los automóviles que ofrecen un manejo completamente autónomo aún no se encuentran a la venta para el público en general. A pesar de que empresas automotrices y del ramo de la tecnología como son *Waymo*, *Tesla*, *Volvo*, etc. han impulsado el desarrollo de vehículos autónomos para el transporte privado, comercial y de carga con la idea de que eventualmente pudieran desplazar o reemplazar a los conductores de camiones de carga, de transporte de personas, taxistas, etc. el *National Highway Traffic Safety Administration (NHTSA)*, por sus siglas en inglés), confirma que hoy en día solo existen funciones que coadyuvan en la conducción de los vehículos para brindar una mayor seguridad al conductor y los pasajeros (NHTSA, s.f.). Funciones como:

- Frenado automático cuando se detecte una posible colisión con un objeto cercano mediante los sensores de proximidad o las cámaras del vehículo.
- Mantener una velocidad de cruce para mitigar la fatiga de los conductores y disminuir el consumo excesivo de combustible generado por una conducción más errática en cuanto a las aceleraciones y desaceleraciones de los conductores.
- Mantener el vehículo dentro del mismo carril a menos que se accione manualmente la direccional para el cambio de carril.

- Detectar los objetos en la periferia del vehículo para alertar a los conductores en caso de que se quiera hacer un cambio de carril, eliminando así el riesgo de que existan objetos en el punto ciego de los conductores y provocando así colisiones en el camino.

Pero como tal, aún no existen vehículos completamente autónomos para la venta al público, pues deben pasar un sinnúmero de pruebas rigurosas y reparar defectos mecánicos, de logística y de cómputo para poder considerar que serán puestos a disposición del público en general en las próximas décadas.

A pesar de estos retos laborales, la *“Propuesta de agenda nacional de Inteligencia Artificial para México (2024-2030)”* (Lagunes A., 2023, p. 15), aborda que *“resulta relevante destacar que la IA no eliminará empleos per se; más bien, los transformará y podrá traer beneficios a aquellas personas que puedan acceder a esta y otras tecnologías emergentes.”* continúa diciendo que *“de acuerdo con un análisis del Foro Económico Mundial se estima que para el año 2025, 85 millones de empleos podrían ser reemplazados por la automatización pero se crearán 97 millones de nuevas oportunidades laborales.”* Así mismo, el estudio menciona que *“En el ámbito del talento, México se destaca en alfabetización en IA; sin embargo, esto ha tenido un mayor impacto al centro del país (CDMX y EDOMEX) - lo cual representa una brecha geográfica de habilidades tecnológicas de todo tipo- licenciados en computación y habilidades disruptivas.”*

Esto es motivo de alerta pues Jalisco se considera una de las regiones más prolíferas en cuanto a tecnologías de información, adquiriendo incluso el mote de *“Valle del Silicio mexicano”* desde hace ya varias décadas y sin embargo esto no se ha visto reflejado en la preparación de las personas, industrias e infraestructura hacia las tecnologías de Inteligencia Artificial.

Retos de Adopción

Dados los rezagos en la alfabetización y adopción de las tecnologías de IA en la región de Jalisco (Lagunes A., 2023, p. 15), existe el riesgo de que no existan suficientes profesionistas capacitados para establecer una mesa de diálogo y así coadyuvar a la creación de un marco legal.

Asimismo, debido a la época de elecciones y a la transición de gobiernos tanto a nivel Federal como Estatal a finales de Mayo y principios de Junio del 2024, las propuestas de cada candidato tomarían precedencia sobre nuevas propuestas como la actual, lo cual es un riesgo latente.

Áreas de oportunidad para Jalisco

Como se ha mencionado anteriormente, los servicios de *GenAI* están generando interés a través de distintos ámbitos en diversos países y esto es solo el comienzo. En base a que la experiencia laboral del autor de este trabajo ha estado íntimamente ligada a la Inteligencia Artificial (*Amazon* 2017-2021 y *Google Vertex AI* 2021 a la fecha), se ha dado a la tarea de realizar un análisis con áreas de oportunidad en las que algunas industrias en México podrían beneficiarse de esta nueva tendencia. Estas propuestas nacen también con base a la problemática mencionada en la sección anterior, en donde empresas de talla mundial se han enfrentado al reto y la complejidad de la Inteligencia Artificial y cómo pueden desarrollar planes específicos para que sus empresas se beneficien de ella.

Algunos ejemplos ilustrativos de cómo los modelos de *GenAI* pueden ayudar desde individuos, sociedades, industrias, gobiernos o empresas en México y particularmente en Jalisco actualmente:

- Individuos:
 - Mitigar la segregación educativa
 - Personas que no tengan el tiempo para aprender las habilidades de programación de computadoras pueden recurrir a los *chatbots* de *GenAI* para generar código en el lenguaje de su preferencia (*Java*, *Python*, *Go*, etc.) prácticamente de forma instantánea en base a una petición del usuario e incluso darles recomendaciones, observaciones u optimizaciones de código preexistente para una mayor eficiencia.
 - Planes de estudio que estén a la vanguardia y normalicen el nivel educativo
 - Será necesario modificar los planes educativos especializados y adaptarlos a las necesidades y tendencias económicas de Jalisco. Esto es fundamental para preparar a los estudiantes para ser más competitivos en el mundo laboral. Estos planes educativos deberían basarse en los modelos, materias y metodologías más innovadores y efectivos que se están aplicando en otras regiones del mundo. Esta estrategia contribuiría a cerrar las brechas de rendimiento educativo que existen entre países y comunidades con diferentes niveles socioeconómicos, tal como lo reflejan los resultados de la Organización para la Cooperación y el Desarrollo Económicos (OCDE).

- En este contexto, el autor considera que la inclusión de materias relacionadas con la Inteligencia Artificial Generativa (*GenAI*) en los planes de estudio de carreras y posgrados relevantes en el Instituto Tecnológico de Estudios Superiores de Occidente (ITESO), como Ingeniería en Sistemas Computacionales o la Maestría en Administración de Empresas, es de suma importancia. Estas materias permitirán a los estudiantes adquirir un conocimiento profundo y una preparación adecuada para aprovechar el potencial transformador de las tecnologías *GenAI* en diversos sectores económicos.
- Industrias:
 - Doblajes y asistencia auditiva o visual al instante
 - En la industria del entretenimiento ya se están haciendo uso de modelos de *GenAI* para traducir videos o películas en tiempo real a cualquier idioma. Por lo que ya no es necesario retrasar el estreno de películas en cines para poder presentarla subtitulada en un idioma distinto al original o brindarles asistencia visual o sonora a personas con limitantes visuales o auditivas en su lengua materna al instante para una mayor democratización de la información. Esto podría aplicarse, por ejemplo, en los diversos proyectos de la Ciudad Creativa Digital Guadalajara – <https://ciudadcreativadigital.mx/>
 - Análisis de fallos en la manufactura para la prevención de accidentes
 - La *GenAI* podría ser usada como herramienta para analizar datos de fallos y accidentes en la industria manufacturera en Jalisco, por ejemplo, en la elaboración del tequila. Al aplicar modelos de *GenAI* a estos datos, se pueden generar planes de mantenimiento y prevención de riesgos más efectivos. Estos planes se beneficiarán de una mejora continua gracias a la incorporación de datos agregados (sin identificadores de empresas, direcciones o sectores específicos) de las regiones con mayor concentración de empresas manufactureras. Esta estrategia colaborativa contribuirá a reducir la incidencia de fallos y accidentes, optimizando la calidad general de los procesos productivos.
- Sociedades:
 - Brindar consultas médicas a todos los mexicanos.

- A lo largo de su historia, la profesión médica ha exigido a los profesionales mantenerse a la vanguardia en el conocimiento de nuevas enfermedades, epidemias, medicamentos, métodos y tratamientos. Sin embargo, la eficacia de un médico tratante aún depende de diversos factores humanos, incluyendo la subjetividad basada en creencias o experiencias personales, la falta de información completa, los errores humanos, el cansancio o agotamiento del médico tratante en el momento de la consulta y las omisiones de los pacientes, lo que puede dificultar el establecimiento de un diagnóstico certero y eficaz. La *GenAI* podría ser utilizada como una herramienta con el potencial de transformar la práctica médica, permitiendo la generación de diagnósticos personalizados basados en la vasta información histórica y global acumulada. Estos diagnósticos podrían ser propuestos al médico tratante para su revisión y confirmación antes de iniciar el tratamiento o la cirugía del paciente.
 - Brindar asesorías legales sin importar el nivel socioeconómico, cultural o dialecto.
 - La instalación de quioscos interactivos en espacios públicos estratégicos de Jalisco, como oficinas gubernamentales, parques, hospitales, escuelas y áreas recreativas, podría brindar acceso universal a asesoría legal rápida, gratuita y en el idioma o dialecto nativo de cada persona. Estos quioscos estarían equipados con asistentes inteligentes como *Amazon Echo* ("*Alexa*") o *Google Mini*, con conexión a internet y acceso a bases de datos de regulaciones y leyes mexicanas y esto supondría un costo relativamente bajo ya que cada dispositivo tiene un costo aproximado de entre \$500-\$2,000 pesos. Además, integrarían un modelo de *GenAI* que funcione como intérprete-traductor, permitiendo atender a personas que hablan diferentes dialectos. Esta iniciativa reduciría significativamente la vulnerabilidad de grupos como personas de escasos recursos, comunidades marginadas y migrantes, quienes, por desconocimiento de sus derechos, son víctimas de abusos por parte de conocidos, familiares, autoridades o instituciones.
- Gobiernos:
 - Sugerir reformas innovadoras y personalizadas a cada sociedad, región y problemática

- *GenAI* puede utilizarse para desarrollar modelos capaces de generar regulaciones y leyes nuevas adaptadas a las necesidades actuales de manera más objetiva, evitando la politización o la búsqueda de beneficios personales. Estos modelos pueden tomar como base regulaciones que han tenido éxito en otras comunidades, países o regiones, y proponer soluciones innovadoras para abordar problemas contemporáneos. Las propuestas generadas por estos modelos podrían servir como referencia para que los legisladores las consideren, realicen los ajustes y cambios necesarios, y evalúen su implementación como leyes para enfrentar los desafíos actuales.
- Empresas:
 - Establecimiento de estrategias de mercado relevantes al instante
 - Analizar tendencias de mercado globales o locales para generar estrategias de mercado específicamente enfocadas al giro de una empresa y al perfil de sus clientes en el segmento de mercado o en la región de su interés.
 - Reducción de riesgos de ciberseguridad
 - La *GenAI* permite desarrollar modelos capaces de realizar un análisis exhaustivo de las vulnerabilidades de seguridad presentes en un sistema informático, incluyendo bases de datos, servidores, terminales, aplicaciones y usuarios. Estos modelos no solo pueden identificar las vulnerabilidades, sino también ofrecer sugerencias e incluso generar el código necesario para su rápida y eficiente resolución. Por ejemplo, estos modelos podrían detectar y neutralizar virus informáticos de forma automática y proveer a los administradores de sistemas un resumen detallado de las acciones que se tomaron de forma automatizada.

Riesgos de no aplicar la *GenAI* en Jalisco

Dan Ives realizó un análisis en *Markets Insider* (Ives, 2024) donde expresó optimismo con base en los resultados financieros compartidos en marzo por *Microsoft Corp.* y *Alphabet Inc.* pues reitera su creencia de que “la revolución de la IA ha llegado”. Microsoft reportó un incremento del 17% en

utilidades comparando al mismo periodo del año anterior para llegar a un total de \$61.9 mMds mientras que *Alphabet* reportó un incremento del 15% en sus utilidades en el mismo periodo para dar un total de \$80.539 mMds. Esto significa un parteaguas en la industria de la tecnología dado que las empresas capaces de ofrecer servicios de *GenAI* atraen interés e inversión de clientes para incursionar en este ámbito. Dan Ives incluso hace hincapié en que la revolución de la Inteligencia Artificial apenas comienza.

En un artículo de CNBC en el pasado mes de marzo de 2024, se hace mención de una encuesta (Colillas, 2024) realizada por el Instituto Tecnológico de Massachussets (*MIT*, por sus siglas en inglés) y Telestra arrojó que solamente el 9% de los líderes empresariales encuestados han implementado la IA generativa dentro de sus compañías. En el mismo artículo se menciona parte de la problemática para aumentar la adopción de estas tecnologías.

“Existe una idea errónea sobre lo fácil que es ejecutar una IA generativa, madura y lista para la empresa”, dijo Stela Solar, directora inaugural del Centro Nacional de Inteligencia Artificial de Australia en el informe de la encuesta, esto lo atribuye a que “Hay elementos circundantes como el diseño de la aplicación, la conexión a datos y procesos comerciales, políticas corporativas y más que aún se necesitan.” (Colillas, 2024)

El autor del presente trabajo advierte sobre el riesgo de que México, y particularmente Jalisco, se quede atrás en el ámbito de la Inteligencia Artificial. Esta situación podría generar una desventaja competitiva estratégica y un rezago en el mercado laboral en comparación con países que ya están a la vanguardia o implementando medidas para prevenir un impacto negativo en su fuerza laboral.

La teoría subyacente es que quienes adopten estas tecnologías en una etapa temprana y las integren en sus negocios, servicios o industrias tendrán una ventaja significativa sobre aquellos competidores que tardan en hacerlo. Los pioneros en adoptar estas tecnologías podrán comprender rápidamente sus beneficios y serán los primeros en contratar a profesionales capacitados para explotarlas al máximo. Esta ventaja se ampliará aún más en detrimento de los competidores que se retrasen.

Es natural que surjan preocupaciones sobre la responsabilidad y las implicaciones del desarrollo de estas tecnologías de IA. La más común es el temor a una amenaza para el mercado laboral y la posible

desaparición de profesiones como médicos, abogados, arquitectos, choferes, entre otras, al ser superadas y potencialmente reemplazadas por estas tecnologías.

Estas tendencias representan una amenaza real para la supervivencia de muchas profesiones, empresas y perfiles laborales. La rapidez, el bajo costo, la precisión y la mitigación de accidentes laborales que ofrecen estos modelos podrían llevar a su sustitución.

Así lo explica Marcus Lu en un artículo publicado en marzo del 2024 a través del portal (Lu, 2024), en donde realiza un análisis con base en más de 19,000 tareas ejecutadas por la *GenAI* y aplicadas en más de 867 ocupaciones. En la Tabla 1 se pueden ver los distintos grupos ocupacionales y el grado en el que la *GenAI* podría reemplazar las tareas realizadas por este tipo de grupos. Una tarea puede ser algo como “elaborar un análisis de mercado” o algo más simple como “resumir un escrito”.

Tabla 1 con los diversos grupos ocupacionales y su impacto. (Lu, 2024),

Departamento	Gran impacto (%)	Pequeño impacto (%)	Sin impacto (%)
Tecnologías de Información	73	26	1
Finanzas	70	21	9
Ventas	67	16	17
Operaciones	68	18	17
Recursos Humanos	57	41	2
Mercadotecnia	56	41	3
Leyes	46	50	4

Cadena de suministro	43	18	39
-----------------------------	----	----	----

En la tabla anterior, las actividades catalogadas como de “Gran impacto” significan todas aquellas tareas que pueden ser completamente automatizadas o considerablemente alteradas por tecnologías de *GenAI*, mientras que las de “Pequeño impacto” son aquellas tareas que tienen menor potencial para ser afectadas.

Con base en dicho análisis, Marcus Lu concluye que los ramos más afectados serían los de las Tecnologías de Información y Finanzas. Sin embargo, se debe considerar que con el paso del tiempo se habrán desarrollado nuevos modelos y aplicaciones de la *GenAI* para tener un mayor rango aplicativo en las diversas industrias con base en las necesidades con mayor eficacia, simplicidad y eficiencia.

Propuesta de un marco de regulación para una *GenAI* responsable

Un marco regulatorio define el conjunto de leyes, normas y directrices que rigen una industria o actividad específica. Su objetivo es establecer un ordenamiento jurídico claro y transparente que garantice la seguridad, la protección de los consumidores y la competencia justa.

Un ejemplo ilustrativo es el marco regulatorio de la Secretaría de Vialidad y Transporte. Este marco establece las normas de circulación, las señales de tránsito y las sanciones para quienes las infrinjan. Su propósito es garantizar la seguridad vial y proteger a los usuarios de las vías públicas.

El auge de *GenAI* exige la creación de un marco regulatorio que promueva su uso responsable. La ausencia de regulaciones podría exacerbar el uso indebido de estas tecnologías con fines irresponsables, ilegales o inmorales, sin la existencia de sanciones o un marco de referencia para determinarlas en función de la gravedad del acto. Un ejemplo preocupante es el uso de *deepfakes* para suplantar la identidad de personas o realizar engaños y estafas. El vacío legal que existe hoy en día supone entonces que las personas que incurran en hacer mal uso de tecnologías de *GenAI* no podrían ser detenidas ni penalizadas.

Google (IA Responsable, 2024), *Microsoft* (Inteligencia artificial responsable, 2024) y otras empresas que ofrecen servicios de *GenAI* hacen énfasis en una IA responsable desde un punto de vista comercial

con el uso de herramientas y procesos para el aseguramiento de la calidad en la información que se entrega a los usuarios, privacidad y seguridad de sus usuarios. En este caso, una propuesta desde un punto de vista más pragmático y general con base en la experiencia profesional del autor:

Transparencia en las fuentes de origen

- Se deberán desarrollar estándares para el desarrollo de los modelos de *GenAI* con un enfoque en la transparencia de cómo obtienen los datos y cómo generan los nuevos contenidos para certificar que no están creando un sesgo informativo (o desinformativo). Solo los modelos que hayan sido supervisados y garanticen esta transparencia deberían de ser certificados para el uso gubernamental, institucional y empresarial.
- Se deberán establecer los procesos de auditoría para la revisión constante de los modelos que operen en el país para asegurar que la implementación de los mismos no difiere con el diseño que habría sido aprobado antes de su utilización.

Controles de seguridad

- Establecer un marco legal administrativo y penal a quienes hagan uso de las tecnologías de *GenAI* para fines fraudulentos, difamatorios, desinformativos o coercitivos.
- Capacitar a las agencias de seguridad (En México, por ejemplo, la Unidad de Ciberseguridad – UNICIBER, dependiente del Estado Mayor General de la Armada, perteneciente a la Armada de México, como indica la (Secretaría de Marina, 2022)) para detectar a las organizaciones e individuos que hagan uso inapropiado de estas tecnologías.

Apego a los valores humanos

- Se deberán establecer normas para el diseño, desarrollo y uso de modelos de *GenAI* teniendo como prioridad la salvaguarda de los valores humanos en todo momento.
- Los modelos de *GenAI* deberán considerar un enfoque para el beneficio social y justo para evitar un sesgo en la visión de los usuarios en contra de un grupo de personas, organizaciones políticas, gubernamentales, Organizaciones No Gubernamentales (ONG) o grupos marginados.

Democratización

- Establecer marcos regulatorios que pongan un porcentaje de modelos de *GenAI* al beneficio de la población de todos los niveles socioeconómicos para evitar la marginación de las personas de escasos recursos, bajo nivel educativo o sin acceso a tecnologías de cómputo mediante las cuales puedan acceder a estos modelos.

En 2023 “se lanzó desde el Senado de la República, la Alianza Nacional de Inteligencia Artificial (ANIA), un mecanismo multiactor que reconoce y fortalece el ecosistema de tecnologías emergentes e inteligencia artificial en México desde una perspectiva integral, plural y multidisciplinaria” (Lagunes A., 2023, p. 18) y esto enfatiza la necesidad de establecer un marco regulativo para el buen uso de las tecnologías de *GenAI* así como las penalidades que se le deberán imponer a quien haga un uso indebido de ellas.

CONCLUSIONES

La incesante búsqueda de la humanidad por mejorar su estilo de vida ha impulsado un avance constante en las tecnologías a lo largo de la historia. *GenAI* no es una excepción. Su vertiginoso crecimiento desde finales del 2022 hasta mediados del 2024 pone de manifiesto el enorme potencial de estas tecnologías, aún en sus primeras etapas de desarrollo. Mientras se exploran nuevas aplicaciones, se hace patente la necesidad de establecer marcos regulatorios que fomenten un uso responsable de la *GenAI*.

Al igual que cualquier tecnología, los avances en *GenAI*, en términos de perfeccionamiento, adaptabilidad y casos de uso, experimentarán una evolución constante. Por lo tanto, el marco regulatorio debe ser dinámico y adaptarse a estos avances, persiguiendo una mejora continua, idealmente anticipándose a las formas en las que se puedan emplear estas tecnologías para usos indebidos.

El autor expresa su preocupación por el posible rezago de Jalisco en el ámbito de la Inteligencia Artificial. Si las empresas mexicanas tardan en adoptar estas tecnologías, sus competidores podrían tener una ventaja considerable.

El mal uso de estas tecnologías también es motivo de inquietud. En México ya se registran numerosos casos de extorsión, falsificación de documentos y robo de identidad. Por ello, es fundamental establecer un marco legal que regule el uso responsable de la Inteligencia Artificial y sancione a quienes la empleen para actividades ilícitas como estafas, intimidaciones o robos.

A continuación, podrán ver unas cuantas citas que evocan a la reflexión en cuanto a los avances tecnológicos a los que se ha enfrentado la humanidad en los últimos 100 años. La cita de Sundar Pichai (CEO de *Alphabet Inc.* desde 2009) es la que motivó al autor de este trabajo a buscar el objetivo de pertenecer a *Google*, particularmente al departamento de *Vertex AI* para ser partícipe de la evolución de *GenAI*.

- “*La tecnología es mejor cuando une a las personas*” - Matt Mullenweg, CEO de Automattic
- “*En la tecnología, todo lo que se pueda hacer, se va a lograr*” – Andy Grove, Ex – CEO de Intel
- “*La Inteligencia Artificial es una de las cosas más profundas en las que estamos trabajando como humanidad. Es más profundo que el descubrimiento del fuego o la electricidad*” – Sundar Pichai, CEO de *Alphabet* y *Google*

BIBLIOGRAFÍA

¿Qué son las alucinaciones de la IA? (2024). Obtenido de IBM: <https://www.ibm.com/topics/ai-hallucinations>

Agency, C. a. (s.f.). *Los bots en redes sociales*. Obtenido de CISA: https://www.cisa.gov/sites/default/files/publications/social-media-bots-infographic-set-spanish_508.pdf

AI, G. V. (2024). *Descripción general de la IA generativa en Vertex AI*. Obtenido de Google Cloud: <https://cloud.google.com/vertex-ai/generative-ai/docs/learn/overview?hl=es-419>

AI, G. V. (2024). *Descripción general de los fundamentos*. Obtenido de Google Cloud: <https://cloud.google.com/vertex-ai/generative-ai/docs/grounding/overview?hl=es-419>

Aleem, Z. (21 de Diciembre de 2023). *AI-generated weapons of mass misinformation have arrived*. Obtenido de MSNBC: <https://www.msnbc.com/opinion/msnbc-opinion/ai-misinformation-fake-news-rcna130523>

Angela Watercutter, W. B. (9 de Noviembre de 2023). *La huelga de actores de Hollywood termina con un acuerdo que impactará a la IA y al streaming por décadas*. Obtenido de Wired: <https://es.wired.com/articulos/huelga-de-actores-de-hollywood-termina-con-acuerdo-que-impactara-a-ia-y-al-streaming-por-decadas#:~:text=9%20de%20noviembre%20de%202023.%20La%20huelga,a%20un%20acuerdo%20provisional%20con%20los%20estudios>.

Brin, L. P. (3 de Diciembre de 2004). <https://blog.google/alphabet/letter-from-larry-and-sergey/>. Obtenido de The Keyword: <https://blog.google/alphabet/letter-from-larry-and-sergey/>

Cloud, G. (18 de Mayo de 2021). *Google Cloud Launches Vertex AI, Making Machine Learning More Accessible and Useful For Developers and Businesses*. Obtenido de PR Newswire: <https://www.prnewswire.com/news-releases/google-cloud-launches-vertex-ai-making->

machine-learning-more-accessible-and-useful-for-developers-and-businesses-301294157.html

Cloud, G. (2024). *Acelera el desarrollo de IA con Google Cloud TPU*. Obtenido de Google Cloud: <https://cloud.google.com/tpu?hl=es-419>

Colillas, D. (5 de Marzo de 2024). *La IA es el tema de conversación de la ciudad, pero las empresas aún no están preparadas para ello, según muestra una encuesta*. Obtenido de CNBC: <https://www.cnbc.com/2024/03/06/generative-ai-holds-massive-potential-but-businesses-arent-ready-yet.html>

Dastin, J. (2023, Febrero 6). *Google unveils ChatGPT rival Bard, AI search plans in battle with Microsoft*. Retrieved from Reuters: <https://www.reuters.com/technology/google-opens-bard-chatbot-test-users-plans-more-ai-search-2023-02-06/>

David Wright, D. S. (4 de Diciembre de 2023). *Gartner Magic Quadrant for Strategic Cloud Platform Services*. Obtenido de Gartner: <https://www.gartner.com/reviews/market/strategic-cloud-platform-services>

Deming, P. (s.f.). *Ciclo de Deming*. Obtenido de Wikipedia: https://es.wikipedia.org/wiki/Ciclo_de_Deming

Dignan, L. (30 de Enero de 2024). *Google Cloud stronger than expected in Q4, \$36 billion annual revenue run rate*. Obtenido de Constellation Research: <https://www.constellationr.com/blog-news/insights/google-cloud-stronger-expected-q4-36-billion-annual-revenue-run-rate>

Doubling down on Waymo One. (26 de Julio de 2023). Obtenido de Waymo: <https://waymo.com/blog/2023/07/doubling-down-on-waymo-one/>

Edwards, B. (30 de Noviembre de 2023). *Año uno de ChatGPT: así ha cambiado el mundo desde la llegada del chatbot de OpenAI*. Obtenido de Wired: <https://es.wired.com/articulos/ano-uno-de-chatgpt-asi-ha-cambiado-el-mundo-desde-la-llegada-del-chatbot-de->

openai#:~:text=Hoy%20hace%20un%20a%C3%B1o%2C%20el,una%20computadora%20que%20hable%20contigo.

El enfoque de Google para luchar contra la desinformación online. (s.f.). Obtenido de Google Centro de Seguridad: https://safety.google/intl/es_es/stories/fighting-misinformation-online/

Forbes. (2023, Enero 27). *Microsoft Confirms Its \$10 Billion Investment Into ChatGPT, Changing How Microsoft Competes With Google, Apple And Other Tech Giants*. Retrieved from Forbes: <https://www.forbes.com/sites/qai/2023/01/27/microsoft-confirms-its-10-billion-investment-into-chatgpt-changing-how-microsoft-competes-with-google-apple-and-other-tech-giants/?sh=13fe0df13624>

Google. (22 de Abril de 2024). *Guía detallada sobre cómo funciona la Búsqueda de Google*. Obtenido de Google Search Central: <https://developers.google.com/search/docs/fundamentals/how-search-works?hl=es-419#introducing-the-three-stages-of-google-search>

Google. (16 de Abril de 2024). *Versiones del modelo y ciclo de vida*. Obtenido de Google Cloud:
<https://cloud.google.com/vertex-ai/generative-ai/docs/learn/model-versioning?hl=es-419>

Google. (s.f.). *Chrome*. Obtenido de Google: <https://www.google.com/chrome/>

Google. (s.f.). *Community Guidelines*. Obtenido de Google: <https://about.google/community-guidelines/#:~:text=When%20communicating%2C%20follow%20the%20three,need%20to%20own%20them%20collectively.>

Google. (s.f.). *Gemini*. Obtenido de Gemini: <https://gemini.google.com/>

IA Responsable. (2024). Obtenido de Google Cloud: <https://cloud.google.com/responsible-ai?hl=es>

Inteligencia artificial responsable. (2024). Obtenido de Microsoft: <https://www.microsoft.com/es-es/ai/responsible-ai>

- Ives, D. (26 de Abril de 2024). *Microsoft y Google: "Momento 1995"*. Obtenido de Markets Insider: <https://markets.businessinsider.com/news/stocks/dan-ives-says-ai-is-the-fourth-industrial-revolution-after-robust-results-from-microsoft-and-google-1995-moment-1033295034>
- Izacard, G. (2023, Febrero 24). *Introducing LLaMA: A foundational, 65-billion-parameter large language model*. Retrieved from LLaMA: <https://ai.meta.com/blog/large-language-model-llama-meta-ai/>
- Jr, R. S. (21 de Junio de 2023). *Cómo Nvidia está utilizando la IA generativa para acelerar su crecimiento*. Obtenido de The Motley Fool: *Cómo Nvidia está utilizando la IA generativa para acelerar su crecimiento*
- Lagunes A., M. Y. (2023). *PROPUESTA DE AGENDA NACIONAL DE LA*. Obtenido de Alianza Nacional de Inteligencia Artificial (ANIA): <https://www.ania.org.mx/>
- Lu, M. (18 de Marzo de 2024). *Gráfico: Los trabajos más afectados por la IA*. Obtenido de Visual Capitalist: <https://www.visualcapitalist.com/charted-the-jobs-most-impacted-by-ai/>
- Mac, M. M. (27 de Diciembre de 2023). *The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work*. Obtenido de New York Times: <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>
- Mok, A. (1 de Abril de 2024). *Gigantes de la IA como OpenAI y Anthropic están luchando por conseguir suficientes datos para entrenar modelos*. Obtenido de Business Insider: <https://www.businessinsider.com/ai-giants-openai-anthropic-running-out-of-good-training-data-2024-4?amp>
- MongoDB. (2024). *¿Qué es la IA generativa?* Obtenido de MongoDB: <https://www.mongodb.com/es/basics/generative-ai>
- NASDAQ. (24 de Mayo de 2024). *Nvidia NASDAQ: Google*. Obtenido de Google: <https://www.google.com/search?q=nvidia+stock&client=firefox-b-1->

d&sca_esv=4a4d881d32876091&sca_upv=1&ei=nUdeZvGyC9q_kPIP6p6M-A8&ved=0ahUKEwix9fjBwcCGAxXaH0QIHWoPA_8Q4dUDCBA&uact=5&oq=nvidia+stock&gs_lp=Egxdnd3Mtd2l6LXNlcnAiDG52aWRpYSBzdG9jazIKEAAYsAMY1gQYR

NHTSA. (s.f.). *Vehículos Automatizados para Mayor Seguridad*. Obtenido de NHTSA: <https://www.nhtsa.gov/es/seguridad-de-vehiculos/vehiculos-automatizados-para-mayor-seguridad>

Pereira, D. (10 de Mayo de 2023). *Google Mission and Vision Statement*. Obtenido de Business Model Analyst: <https://businessmodelanalyst.com/google-mission-and-vision-statement/>

Revell, E. (17 de Marzo de 2024). *Microsoft dice a los reguladores europeos que Google tiene una ventaja en IA generativa*. Obtenido de Fox Negocios: <https://www.foxbusiness.com/technology/microsoft-tells-european-regulators-google-has-edge-generative-ai.amp>

Scheffler, E. (23 de Marzo de 2023). *Se viraliza una foto del Papa Francisco creada con una herramienta de inteligencia artificial*. Obtenido de Entrepreneur: <https://www.entrepreneur.com/es/noticias/se-viraliza-una-foto-del-papa-francisco-creada-con-una/448438>

Secretaría de Marina. (27 de Octubre de 2022). *Coordinadora General del Ciberespacio*. Obtenido de Gobierno de México: <https://www.gob.mx/semar/articulos/unidad-de-ciberseguridad-279197?idiom=es#:~:text=%2D%20335%2F2022%20del%2001%2F,Mayor%20General%20de%20la%20Armada>.

Shirer, M. (21 de Diciembre de 2023). *IDC Forecasts Spending on GenAI Solutions Will Double in 2024 and Grow to \$151.1 Billion in 2027*. Obtenido de International Data Corporation (IDC): <https://www.idc.com/getdoc.jsp?containerId=prUS51572023#:~:text=IDC%20Forecasts%20Spending%20on%20GenAI,to%20%24151.1%20Billion%20in%202027>

Varios. (8 de Mayo de 2024). *Diagrama de Flujo*. Obtenido de Wikipedia: https://es.wikipedia.org/wiki/Diagrama_de_flujo

