

## **Explorando oportunidades y desafíos de la Inteligencia Artificial Generativa en imágenes médicas y analítica de salud con ChatGPT**

### *Posibles oportunidades y ventajas de utilizar modelos generativos de IA*

La generación de datos sintéticos mediante el uso de redes generativas adversarias (GANs) tiene un potencial significativo en la investigación médica, particularmente en el estudio de enfermedades poco frecuentes y en el fortalecimiento de los conjuntos de datos de entrenamiento. Cuando los casos reales de una enfermedad son escasos, la generación de datos sintéticos permite crear un conjunto de datos más grande y diverso. Esto puede ser fundamental para realizar análisis estadísticos significativos y desarrollar modelos de predicción confiables.

Los datos sintéticos brindan la oportunidad de realizar experimentos y pruebas en un entorno virtual antes de aplicar estrategias y tratamientos en pacientes reales. Esto puede acelerar el proceso de desarrollo de terapias y enfoques de atención.

La generación de datos sintéticos garantiza que la privacidad de los pacientes sea protegida. Los datos reales de pacientes no necesitan ser compartidos o expuestos, lo que cumple con los estándares de protección de datos en la investigación médica.

Contar con conjuntos de datos de entrenamiento grandes y variados es esencial para desarrollar modelos de aprendizaje automático precisos. Los datos sintéticos pueden enriquecer estos conjuntos de las siguientes maneras. Las GANs pueden generar una amplia gama de datos que representen diferentes condiciones clínicas, edades, géneros, y otras características relevantes. Esto mejora la capacidad del modelo para generalizar y adaptarse a diversos escenarios clínicos.

### *Actividades específicas a ser agilizadas o automatizadas mediante estos modelos*

La generación de imágenes sintéticas puede acelerar la interpretación radiológica y la detección de patologías. Modelos generativos pueden crear resúmenes precisos y concisos de investigaciones médicas, reduciendo la carga cognitiva de los profesionales de la salud.

### *El nuevo rol de los seres humanos al contar con herramientas de IA como soporte*

A medida que la IA generativa se integre en la atención médica, el papel de los profesionales de la salud evolucionará. Los médicos supervisarán y validarán las decisiones de la IA, aportando su experiencia y juicio clínico. Los médicos podrán pasar

más tiempo interactuando con los pacientes, brindando un enfoque más humano y compasivo en la atención. Los profesionales se involucrarán en el entrenamiento de modelos y contribuirán a su mejora continua.

### *Riesgos sociales, económicos o morales debido al uso de modelos generativos de IA*

El uso de datos sintéticos puede mantener sesgos existentes en los datos, lo que podría resultar en inequidad en el resultado y en la utilidad del modelo. La toma de decisiones automatizada plantea la cuestión de quién es responsable en caso de errores o decisiones éticamente cuestionables. Por lo cual hay una necesidad de una supervisión ética y una regulación para garantizar su uso justo y responsable, e iniciativas como *AI for Good*, que centren en el uso IA para abordar desafíos globales, en pro del bienestar y el progreso social.

### Referencias:

Elicit. (2023, 16 septiembre). Generative ia, radiology, real word evidence, health analytics. <https://elicit.com/>

OpenAI. (2023, 16 septiembre). These papers collectively highlight the potential of generative AI in the field of radiology. ChatGPT [Modelo de lenguaje GPT-3.5].

Arora, A., & Arora, A. (2022). Generative adversarial networks and synthetic patient data: current challenges and future perspectives. *Future Healthcare Journal*, 9(2), 190.

Stiglic, G., Kopitar, L., Gosak, L., Kocbek, P., He, Z., Chakraborty, P., ... & Bian, J. (2023). Improving Primary Healthcare Workflow Using Extreme Summarization of Scientific Literature Based on Generative AI. *arXiv preprint arXiv:2307.15715*.

Xie, Q., Schenck, E. J., Yang, H. S., Chen, Y., Peng, Y., & Wang, F. (2023). Faithful AI in Healthcare and Medicine. *medRxiv*, 2023-04.

Patil, P. V., Dadpe, S. S., & Sultanpure, V. M. (2021). Health analysis in artificial intelligence. *International Journal of Scientific Research in Science, Engineering, and Technology*, 8(2), 7.

OpenAI. (2023, 16 septiembre). human with circuits colors cyan red black Nicolas Winding Refn and Shepard Fairey and realistic style. DALL-E.

OpenAI. (2023, 16 septiembre). x-ray machine with a skeleton in a futuristic metro station. DALL-E.

## **Exploring opportunities and challenges of Generative Artificial Intelligence in medical imaging and health analytics with ChatGPT**

### *Possible Opportunities and Advantages of Using Generative AI Models*

The generation of synthetic data using Generative Adversarial Networks (GANs) holds significant potential in medical research, especially in studying rare diseases and strengthening training datasets. When real cases of a disease are scarce, generating synthetic data allows for the creation of larger and more diverse datasets. This can be crucial for conducting meaningful statistical analyses and developing reliable prediction models.

Synthetic data provides the opportunity to conduct experiments and tests in a virtual environment before applying strategies and treatments to real patients. This can expedite the development of therapies and healthcare approaches.

The generation of synthetic data ensures the privacy of patients is protected. Real patient data does not need to be shared or exposed, aligning with data protection standards in medical research.

Having large and diverse training datasets is essential for developing accurate machine learning models. Synthetic data can enrich these datasets in the following ways. GANs can generate a wide range of data representing different clinical conditions, ages, genders, and other relevant characteristics. This enhances the model's ability to generalize and adapt to various clinical scenarios.

### *Specific Activities to be Streamlined or Automated by These Models*

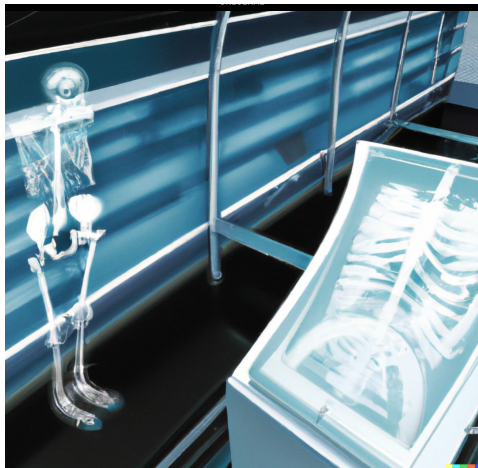
The generation of synthetic images can expedite radiological interpretation and pathology detection. Generative models can create accurate and concise summaries of medical research, reducing the cognitive burden on healthcare professionals.

### *The New Role of Humans with AI Tools as Support*

As generative AI becomes integrated into healthcare, the role of healthcare professionals will evolve. Doctors will oversee and validate AI's decisions, providing their expertise and clinical judgment. Doctors will have more time for patient interaction, offering a more human and compassionate approach to care. Professionals will be involved in model training and contribute to its continuous improvement.

## *Social, Economic, or Moral Risks Due to the Use of Generative AI Models*

The use of synthetic data can perpetuate biases present in the original data, potentially resulting in inequities in model outcomes and utility. Automated decision-making raises questions about who is responsible in case of errors or ethically questionable decisions. Therefore, there is a need for ethical oversight and regulation to ensure fair and responsible use. Initiatives like AI for Good, which focus on using AI to address global challenges for the betterment of society and social progress, are essential in this context.



## **Dilemas Éticos en el Uso de ChatGPT y Modelos Generativos de IA**

Si un profesional de la salud utiliza ChatGPT para obtener información sobre un diagnóstico y tratamiento de una enfermedad específica, y el modelo presenta recomendaciones sesgadas que reflejan estereotipos culturales o de género, esto podría llevar a una atención inapropiada o insatisfactoria.

### *Actores involucrados*

Profesionales de la salud, su interés principal es proporcionar la mejor atención posible a los pacientes y tomar decisiones clínicas basadas en información precisa y confiable.

Desarrolladores de IA, tienen un interés en mejorar la eficacia y utilidad de sus modelos.

Pacientes, les interesa confiar en la atención que reciben y en la relación con sus médicos.

Comites de ética, tienen interés en garantizar que la atención médica se brinde de manera justa y equitativa, sin discriminación basada en género, raza u otros factores.

### *Principios éticos en conflicto*

Precisión vs sesgo, existe un conflicto entre el principio de proporcionar información precisa y el riesgo de que los modelos de lenguaje generen respuestas sesgadas debido a sesgos en los datos de entrenamiento.

Autonomía vs confianza, la autonomía del profesional de la salud para utilizar herramientas de IA choca con la necesidad de que los pacientes confíen en la atención que reciben y en la tecnología utilizada en su atención.

Justicia vs discriminación, el principio de justicia en la atención médica se opone a la discriminación basada en género, raza u otros factores.

### *Consecuencias a corto y largo plazo*

En el corto plazo, las respuestas sesgadas generadas por modelos como ChatGPT podrían resultar en diagnósticos y tratamientos inapropiados para ciertos pacientes. La falta de confianza en la tecnología de IA podría llevar a la resistencia por parte de algunos profesionales para adoptar estas herramientas en su práctica diaria. Si los

sesgos persisten, podrían surgir desigualdades crónicas en la atención médica, lo que afectaría negativamente la salud y el bienestar de grupos marginados o minoritarios.

### *Posibles soluciones*

Entrenamiento con datos diversos y representativos. Garantizar que los modelos de lenguaje se entrenen con datos que sean diversos y representativos de la población, lo que puede ayudar a reducir sesgos culturales y de género en las respuestas. Puede ser desafiante recopilar datos verdaderamente representativos y mantener la privacidad de los pacientes al hacerlo. Revisión humana y corrección activa, implementar un proceso de revisión humana de las respuestas generadas por la IA antes de que lleguen a los profesionales de la salud. Puede aumentar la carga de trabajo de revisores humanos y ralentizar el proceso de obtención de respuestas. Transparencia y explicabilidad, proporcionar información detallada sobre cómo se generaron las respuestas de la IA, incluidos los datos de entrenamiento y los algoritmos utilizados. La explicación completa de los modelos de lenguaje puede ser compleja y requerir un esfuerzo adicional para comprenderla.

### *Recomendaciones*

Entrenamiento con datos diversos y representativos, esto ayudará a abordar los sesgos desde la base y a proporcionar respuestas más equitativas. Revisión humana y corrección activa: implementar un proceso de revisión humana para identificar y corregir sesgos y errores en las respuestas generadas. Transparencia y explicabilidad, proporcionar información transparente sobre cómo funcionan los modelos de lenguaje y cómo se generan las respuestas. Esto permitirá a los profesionales de la salud evaluar críticamente la información proporcionada por la IA.

### *Reflexión personal*

Los profesionales de la salud tienen responsabilidad al utilizar modelos de IA generativos para la toma de decisiones. Es esencial que se capaciten en como usar esta tecnología, conocer sus beneficios y limitaciones. También involucrarse más para trabajar en conjunto con desarrolladores y reguladores, para garantizar que se cumplan los principios éticos, de responsabilidad y equidad.

## **Ethical Dilemmas in the Use of ChatGPT and Generative AI Models**

If a healthcare professional uses ChatGPT to obtain information on the diagnosis and treatment of a specific disease, and the model provides biased recommendations reflecting cultural or gender stereotypes, this could lead to inappropriate or unsatisfactory care.

### *Stakeholders*

Healthcare professionals, whose primary interest is providing the best possible care to patients and making clinically sound decisions based on accurate and reliable information.

AI developers, who have an interest in improving the effectiveness and utility of their models.

Patients, who are interested in trusting the care they receive and their relationship with their doctors.

Ethics committees, who are interested in ensuring that healthcare is provided fairly and equitably without discrimination based on gender, race, or other factors.

### *Conflicting Ethical Principles*

Precision vs. bias, there is a conflict between the principle of providing accurate information and the risk of language models generating biased responses due to biases in training data.

Autonomy vs. trust, the autonomy of healthcare professionals to use AI tools clashes with the need for patients to trust the care they receive and the technology used in their care.

Justice vs. discrimination, the principle of justice in healthcare opposes discrimination based on gender, race, or other factors.

### *Short and Long-Term Consequences*

In the short term, biased responses generated by models like ChatGPT could result in inappropriate diagnoses and treatments for certain patients. Lack of trust in AI technology could lead to resistance from some professionals in adopting these tools in

their daily practice. If biases persist, chronic healthcare inequalities may emerge, negatively affecting the health and well-being of marginalized or minority groups.

### *Possible Solutions*

**Training with Diverse and Representative Data.** Ensure language models are trained on data that is diverse and representative of the population, which can help reduce cultural and gender biases in responses. The challenge lies in collecting truly representative data while maintaining patient privacy. **Human Review and Active Correction:** Implement a process of human review of responses generated by AI before they reach healthcare professionals. This can identify and correct biases and errors but may increase the workload of human reviewers and slow down response times. **Transparency and Explainability:** Provide detailed information on how AI responses are generated, including training data and algorithms used. Full explanation of language models can be complex and require additional effort to understand.

### *Recommendations*

A combination of solutions is recommended. **Training with Diverse Data:** This will address biases at the foundational level and provide more equitable responses. **Human Review and Correction:** Implement a process of human review to identify and correct biases and errors. **Transparency and Explainability:** Provide transparent information on AI operations to allow healthcare professionals to critically evaluate the information provided.

### *Personal Reflection*

Healthcare professionals have a responsibility when using generative AI models for decision-making. It is essential for them to be trained in how to use this technology, understand its benefits and limitations, and become more involved in working collaboratively with developers and regulators to ensure ethical, responsible, and equitable use of AI.

### *References:*

OpenAI. (2023, 16 septiembre). These papers collectively highlight the potential of generative AI in the field of radiology. ChatGPT [Modelo de lenguaje GPT-3.5].