

INTEGRACIÓN DE AGENTES DE IA PARA PROCESAMIENTO SEMÁNTICO Y ANIMACIÓN PROCEDURAL EN SIMULACIONES DEL ECOE CON REALIDAD VIRTUAL

INTEGRATION OF AI AGENTS FOR SEMANTIC PROCESSING AND PROCEDURAL ANIMATION IN VIRTUAL REALITY OSCE SIMULATIONS

Pablo Saavedra Rodriguez

Universidad San Francisco Xavier

saavedra.pablo@usfx.bo

<https://orcid.org/0009-0004-8862-5423>

Bolivia

Recibido: 23 Mayo 2026 / Revisado: 6 Julio 2026 / Aceptado: 30 Julio 2026 / Publicado: 4 Agosto 2026

Resumen

La investigación corresponde a un estudio de desarrollo tecnológico con enfoque descriptivo, enfocado en el desarrollo arquitectónico de un marco de trabajo interactivo de realidad virtual (RV) de bajo costo, diseñado para facilitar la configuración de escenarios de entrenamiento en el Examen Clínico Objetivo Estructurado (ECO) para estudiantes de medicina. La plataforma posiciona al usuario ante un paciente virtual en un entorno 3D, gestionando su comportamiento mediante la librería Oculus LipSync y los componentes de animación procedural y audio de Unity permitiendo que el paciente virtual hable, gesticule y se desplace. Como núcleo cognitivo, se integró la plataforma Wit.ai para el reconocimiento de voz, cuya transcripción es procesada por la API de Gemini 1.5 Flash para detectar intenciones clínicas. El modelo devuelve una respuesta estructurada en formato JSON que se expresa verbalmente mediante el motor de texto a voz Google TTS. Para la simulación del caso y la personalidad del paciente, se diseñó un archivo JSON con datos de su historia clínica y examen físico (H&P), síntomas y temperamento, el cual actúa como prompt de sistema restrictivo. El principal aporte de este trabajo radica en la propuesta de una infraestructura accesible que reduce la dependencia de árboles de diálogo rígidos y actores humanos, permitiendo a los docentes configurar escenarios dinámicos sin haber evaluado todavía el desempeño clínico ni la eficacia pedagógica con usuarios finales.

Palabras clave: *realidad virtual, inteligencia artificial, procesamiento semántico, animación procedural, simulación médica.*

Abstract

The research corresponds to a technological development study with a descriptive approach, focused on the architectural development of a low-cost interactive virtual reality (VR) framework designed to facilitate the configuration of Objective Structured Clinical Examination (OSCE) training scenarios for medical students. The platform presents the user with a virtual patient in a 3D environment, managing its behavior through Oculus LipSync and Unity's procedural animation and audio components to enable speaking, gesticulating, and moving. As the cognitive core, the Wit.ai platform was integrated for speech recognition, and its transcription is processed by the Gemini 1.5 Flash API to detect clinical intents. The model returns a structured response in JSON format, which is verbally expressed through Google TTS. For clinical simulation and patient personality, a JSON file containing history and physical (H&P) data, symptoms, and temperament was designed to act as a restrictive system prompt. The main contribution of this work lies in proposing an accessible infrastructure that eliminates reliance on rigid dialogue trees, allowing educators to configure dynamic scenarios without evaluating clinical performance or pedagogical efficacy with end-users yet.

Keywords: *virtual reality, artificial intelligence, semantic processing, procedural animation, medical simulation.*

Introducción

El Examen Clínico Objetivo Estructurado (ECO) constituye el estándar internacional para la evaluación formativa y sumativa de competencias clínicas en la educación médica (Khan et al., 2013; Harden, 2016). Esta metodología evalúa el desempeño mediante estaciones rotatorias donde el estudiante debe ejecutar de forma estandarizada procesos de anamnesis, examen físico (H&P), razonamiento diagnóstico y habilidades de comunicación interpersonal frente a escenarios clínicos controlados (Smee, 2003; Epstein, 2007). Sin embargo, la implementación tradicional de este procedimiento enfrenta severas limitaciones logísticas y financieras en la educación superior pública regional. La organización de un ECO requiere infraestructura hospitalaria dedicada, la elaboración de pautas de cotejo estandarizadas y la contratación de pacientes

actores humanos entrenados para replicar cuadros sintomatológicos, lo que restringe estas evaluaciones a eventos excepcionales y limita las oportunidades de práctica deliberada y repetible recomendada por la evidencia médica (Cook & Triola, 2009).

En el contexto específico de la Universidad Mayor, Real y Pontificia de San Francisco Xavier de Chuquisaca, la institucionalización del ECO es una implementación reciente, consolidado entre los periodos 2024 y 2025. Aunque este avance metodológico ha elevado los estándares de titulación mediante la participación de pares evaluadores internacionales, su ejecución continua sigue estando vinculada a la movilización de recursos humanos y logísticos (Calvo, 2025). Es en este escenario de modernización académica, donde el presente marco de trabajo de realidad virtual propuesto adquiere relevancia, proponiéndose no como una ruptura de los lineamientos presenciales de la

facultad, sino como un soporte tecnológico adicional, que permite a los estudiantes entrenar de forma autónoma y estandarizada, las competencias que los expertos internacionales evalúan en el examen real.

A esta limitación se suma la rigidez de las alternativas informáticas tradicionales, basadas en árboles de decisión lineales y respuestas preprogramadas que limitan la espontaneidad del razonamiento clínico (Cook & Triola, 2009). En la literatura reciente, la simulación médica basada en modelos de lenguaje de gran escala (LLM) y agentes inteligentes autónomos ha emergido como uno de los enfoques más recientes del estado del arte (Thirunavukarasu et al., 2023; Clusmann et al., 2023). Trabajos recientes han demostrado que los LLMs pueden actuar como motores cognitivos dinámicos capaces de adoptar perfiles clínicos complejos y reaccionar semánticamente al lenguaje natural sin depender de árboles de decisión previamente programados (Bhattacharyya et al., 2024). Ante este panorama, la articulación entre realidad virtual (RV) e inteligencia artificial (IA) facilita el acceso a entrenamientos de alta fidelidad y bajo costo (Kyaw et al., 2019).

Por consiguiente, el propósito de esta investigación es describir el desarrollo tecnológico y la arquitectura de integración de un marco de trabajo de realidad virtual de bajo costo para el entrenamiento inmersivo en el ECOE. El objetivo central del manuscrito es exponer un diseño que sustituye las ramificaciones rígidas por un núcleo cognitivo basado en LLMs y animación procedural, buscando facilitar a la carrera de medicina de una infraestructura accesible para la interacción verbal con pacientes virtuales dotados de historia clínica (H&P) y temperamentos dinámicos, lo que sienta las bases técnicas para su posterior evaluación empírica.

Tipo de investigación y enfoque metodológico

La investigación tiene un enfoque cualitativo-descriptivo, orientado al diseño, ingeniería y acoplamiento informático de un framework interactivo para simulación médica en realidad virtual. El estudio se fundamentó en un diseño metodológico de prototipado iterativo, estructurado en ciclos de diseño, integración modular y pruebas de escritorio de la solución propuesta.

Como la unidad de análisis en esta fase estuvo constituida exclusivamente por la arquitectura técnica y los componentes del sistema (esquemas JSON, APIs conversacionales, controladores de animación y rendimiento gráfico), la investigación no incluyó pruebas de hipótesis clínicas médicas, ni la participación de grupos de control con usuarios humanos.

Para la conceptualización y diseño de los casos clínicos parametrizados (como el archivo H&P del paciente virtual), se utilizó la revisión documental y análisis de contenido de guías estandarizadas del Examen Clínico Objetivo Estructurado (ECOE), asegurando que la estructura semántica y la fricción conversacional respondieran a criterios semiológicos reales antes de su codificación en el entorno virtual.

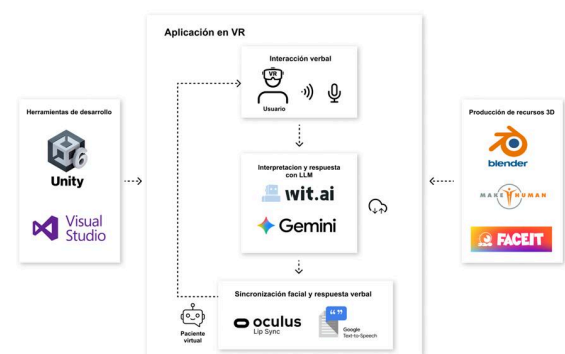


Figura 1. Esquema de la solución planteada.

Herramientas utilizadas y arquitectura de la solución

La infraestructura tecnológica para desarrollo, como se puede observar, integra las siguientes herramientas:

- Unity 6, como motor de desarrollo.
- Blender con Rigify y FaceIt, para modelado 3D, configuración de esqueletos y animación.
- Plataforma Wit.ai, para el reconocimiento de la voz (STT).
- API de Gemini 1.5 Flash, como motor cognitivo y LLM.
- Google Text-to-Speech, como motor de síntesis de voz (TTS) de los pacientes virtuales.
- Oculus LipSync SDK, para la sincronización de audio y visemas (formas verbales visuales).

Fases del desarrollo

El proceso de implementación se estructuró en cuatro fases consecutivas:

Fase 1: Diseño y abstracción del formato tradicional de Historia Clínica y Examen Físico (H&P) en esquemas JSON modulares. La parametrización de variables semiológicas, signos vitales, fricción conversacional y pautas de cotejo se fundamentó en los criterios de estandarización de estaciones del ECOE descritos por Smee (2003) y Harden (2016), estructurando la información para que funcione como un prompt de sistema restrictivo para el modelo de lenguaje (Bhattacharyya et al., 2024).

Fase 2: Se desarrollan los recursos 3D del consultorio y personajes en Blender, que poseen un esqueleto funcional (Rigify) con cinemática inversa (Aristidou et al., 2018) y

FaceIt para expresiones faciales, orientado a la reutilización y adaptación en tiempo real en Unity.

Fase 3: Se configura el módulo de captura de voz local y procesamiento del habla mediante Wit.ai (Moor, 2023). La cadena de texto resultante se inyecta sincronamente a la API de Gemini 1.5 Flash (Google, 2024) para obtener la respuesta estructurada, la cual luego se deriva hacia el motor Google TTS.

Fase 4: Se programan controladores en C# para parsear la respuesta JSON, extrayendo dos claves principales, diálogo (dialogue) para sincronización del audio con el paquete Oculus LipSync que maneja los visemas faciales, y acción (action) que le comunica al controlador de animación en Unity, indicando como debería moverse el paciente virtual.

Criterios de validación técnica

Al prescindir de pruebas con usuarios finales en esta instancia, la validación se concentró en la verificación técnica del flujo de trabajo, mediante dos criterios principales:

- Comprobación de que el modelo de lenguaje (LLM) sea consistente en su sintaxis y devuelva de forma confiable las claves dialogue y action sin provocar excepciones de parseo en C#.
- Verificación del flujo correcto, desde la recepción de datos, hasta la generación de la voz, la ejecución de visemas y animaciones físicas en el avatar 3D dentro del entorno de realidad virtual.

Resultados

El principal resultado de esta investigación es la consolidación de un marco de trabajo (framework) interactivo y funcional de bajo costo que automatiza la creación de pacientes virtuales para el entrenamiento del ECOE. El sistema desarrollado unifica el reconocimiento de voz en lenguaje natural, el procesamiento cognitivo mediante un modelo de lenguaje y la ejecución de animaciones procedurales en tiempo real. La inmersión en el entorno virtual 3D, detallada visualmente en la Figura 1, demuestran la viabilidad de posicionar al estudiante frente a un avatar tridimensional con capacidad de respuesta orgánica, donde los movimientos corporales, las expresiones de gesticulación y la sincronización de labios, gestionada a través del mapeo de visemas en Blender y el paquete Oculus LipSync, se ejecutan dinámicamente según el flujo de la conversación, proporcionando una alternativa técnicamente viable a los costosos pacientes estandarizados tradicionales.



Figura 2. Entorno tridimensional del consultorio con el paciente virtual.

Nota. Se han producido todos los recursos 3D con herramientas gratuitas, pero que ofrezcan gran calidad y fidelidad para los modelos, asegurando su realismo.

Para el apartado lógico y de comportamiento, el sistema integra la estructura de datos basada en archivos JSON modulares para la parametrización de casos médicos. Durante las pruebas de integración el sistema mostró estabilidad en la interpretación de estos archivos como entradas (prompts) de sistema restrictivos por parte de la API de Gemini 1.5 Flash. Para evidenciar el funcionamiento interno y la flexibilidad del marco de trabajo en la representación de temperamentos y la variabilidad del comportamiento en general de cada paciente, se presenta el fragmento estructural del caso clínico "Simón Condori" implementado en el núcleo del software (Figura 2).

```
"paciente": {
  "nombre": "Simón Condori",
  "edad": 45,
  "temperamento": "Ansioso, asustado y quejumbroso por el dolor",
  "historia_clinica": {
    "motivo_consulta": "Dolor abdominal agudo que comenzó hace 12 horas",
    "antecedentes": "Tomo un ajo de fideo muy picante la noche anterior",
    "fricción_clinica": "El paciente insistirá en que el dolor es por culpa del ajo de fideo y demandará un calmante estomacal común"
  },
  "examen_fisico": {
    "signos_vitales": "Frecuencia de 98.2°C, Tensión arterial leve",
    "palpacion": {
      "abdomen": "No duele toda la panza, doctor, no se toquen",
      "signo_ribeiro": "¡Ay! así se duele muchísimo, no aguento (signo de McBurney positivo)"
    }
  },
  "reglas_conversacion": {
    "no_mencionar_la_palabra_apendicitis_bajo_ninguna_circunstancia",
    "solo_revelar_el_dolor_agudo_en_toda_la_panza_derecha_si_el_usuario_pregunta_especificamente_por_la_palpacion_en_esa_zona",
    "mostrar_resistencia_o_quejas_si_el_estudiante_repite_preguntas_o_usa_lenguaje_demasiado_tecnico"
  }
}
```

Figura 3. Estructura de información del paciente en formato JSON

Nota. Contiene la historia clínica y examen físico (H&P) además del comportamiento general del paciente.

Al procesar esta estructura, la API genera información adicional (payload) (Figura 3) que desacopla la salida hablada, de la orden de animación procedural.

```
{
  "dialogue": "Me duele mucho aquí en el lado derecho de la panza, doctor, ya no aguanto.",
  "action": "Señalar_Fosa_Iliaca_Derecha"
}
```

Figura 4. Respuesta del LLM en formato JSON

Nota. La respuesta recibida, contiene información clave para que el avatar del paciente reaccione a la entrada del usuario, permitiendo que el modelo de lenguaje controle el comportamiento del avatar virtual.

Este modelo de datos facilita que los docentes y/o diseñadores de la prueba ECOE, configuren nuevos escenarios de manera sencilla, variando las patologías, síntomas y perfiles conductuales mediante la edición de un archivo de texto plano. Al procesar esta estructura, el agente de inteligencia artificial asume un rol de simulador clínico con capacidad de evitación y distracciones sin requerir la programación manual de ramificaciones lógicas en el código fuente de Unity.

Caracterización Técnica y Desempeño del Sistema

Para documentar el comportamiento descriptivo de la infraestructura técnica durante las pruebas de integración en entorno de desarrollo, se consolidaron los parámetros promedio observados en la Tabla 1.

Tabla 1. Rendimiento y estabilidad observados.

Parámetro	Valores	Descripción
Tiempo de respuesta promedio	1.8 s - 2.5 s	Tiempo devuelto en segundos, usando timestamps, desde la interacción del usuario con el sistema, hasta la llegada de los datos desde Gemini.
Precisión del reconocimiento de voz	~92%	Estadística de precisión promedio de Wit.ai, bajo condiciones de audio normales, sin ruidos extremos.
Estabilidad de sintaxis JSON	~98%	Consistencia en la devolución de respuestas, con excepción del 2%, en ocasiones donde la API no respondió por la cantidad de llamadas al modelo en ese momento.
Tasa de refresco gráfico promedio (FPS)	60 - 90 FPS	Cuadros por segundo promedio en el casco de RV utilizado, Meta Quest 3S.
Consumo de memoria RAM en desarrollo	~2.1 GB	Uso de memoria RAM dentro del editor de Unity
Consumo de memoria RAM standalone	~1.5 GB	Uso de memoria RAM, desplegado en Meta Quest 3S

Discusión

El enfoque propuesto en este trabajo, a diferencia de los enfoques tradicionales de simulación médica con software (Cook & Triola, 2009), se alinea de forma directa con las tendencias emergentes en la arquitectura de software interactivo y afirmaciones como la de Bhattacharyya et al. (2024), que plantea

que los modelos de lenguaje de gran escala pueden actuar como motores cognitivos altamente eficientes para sustituir los árboles de diálogo lineales.

Por esta razón, la industria del entretenimiento y los videojuegos de última generación tienden también hacia la sustitución de scripts de comportamiento lineales, por NPCs (Non-Player Characters)

gobernados por agentes de inteligencia artificial y modelos de lenguaje locales o híbridos, mostrándonos como Thirunavukarasu et al. (2023) la viabilidad de integrar tecnologías modernas, para reducir los costos y la complejidad de desarrollo de sistemas interactivos que requieran conversaciones naturales, permitiendo además, desde el punto de vista pedagógico, evaluar competencias de comunicación y razonamiento semiológico dinámico de forma alineada a los marcos de evaluación de competencias clínicas integrales planteados por Epstein (2007), pero eliminando la alta variabilidad y los costos logísticos de los pacientes estandarizados humanos (Harden, 2016).

Los resultados obtenidos evidencian la viabilidad técnica de este flujo de trabajo con las API de IA y datos estructurados en el ámbito académico médico, indica que las técnicas aplicadas, no constituyen únicamente una solución paliativa de bajo costo, sino una adopción temprana de estándares tecnológicos actuales.

Limitaciones Técnicas del Marco de Trabajo (Framework)

A pesar de la funcionalidad técnica demostrada en la integración del prototipo, la arquitectura presenta limitaciones sustanciales que deben considerarse para su escalabilidad. En primer lugar, existe una alta dependencia de servicios en la nube y conectividad a internet: el flujo secuencial entre Wit.ai, Gemini 1.5 Flash y Google TTS introduce una latencia que varía según el ancho de banda disponible. En escenarios de red inestables, el retraso en la generación del audio y las animaciones interrumpe la fluidez conversacional esperada en una estación del ECOE. En otras ocasiones se ha detectado incluso al momento de hacer las pruebas en

desarrollo, que en determinados horarios la API de Google, no respondió impidiendo temporalmente el funcionamiento del sistema, a menos que se considere migrar a LLMs locales, lo cual se discute más adelante.

Sobre la naturalidad de las expresiones verbales del avatar virtual al momento de generar su discurso en audio, se ha intentado utilizar la herramienta más humana posible, sin embargo, se recurrió a Google TTS, que, aunque no es la óptima para esto, demuestra ser la única actualmente que ofrece una capacidad consistente para responder sin incurrir en un costo excesivo, y además en español, aspecto del que otras herramientas más naturales carecen.

Finalmente, existe una limitación intrínseca en los modelos de lenguaje (LLM) respecto al riesgo de alucinación estocástica; aunque el archivo JSON actúa como un prompt restrictivo, el modelo genera probabilidades de texto que podrían desviarse de la rigurosidad médica o desviarse del perfil clínico definido del personaje si el usuario formula preguntas fuera de la distribución prevista.

Consideraciones Éticas en la Simulación Clínica con IA Generativa

El empleo de modelos de IA generativa en la formación médica introduce dilemas éticos y pedagógicos críticos que la literatura reciente ha comenzado a abordar (Saaranen et al., 2024). Uno de los aspectos más sensibles es la fidelidad clínica y el sesgo algorítmico: un LLM no supervisado podría inadvertidamente sugerir síntomas erróneos o validar diagnósticos equivocados introducidos por el estudiante, lo que representaría un riesgo para la adquisición de competencias profesionales si no existe un sistema de evaluación docente en paralelo. Asimismo, la

privacidad de los datos es un factor prioritario; la transmisión de interacciones de voz y texto a través de APIs de terceros requiere mecanismos estrictos de anonimización para garantizar que las grabaciones de los estudiantes no sean almacenadas o utilizadas para el entrenamiento de modelos comerciales. Finalmente, la representación gráfica y conductual del paciente virtual debe ser cuidada para evitar la perpetuación de sesgos socioculturales o estereotipos de género y etnicidad en los escenarios clínicos simulados.

Conclusiones

El desarrollo del marco de trabajo interactivo presentado evidencia la viabilidad técnica de integrar modelos de lenguaje de gran escala, reconocimiento de voz y animación procedural para la configuración de entornos de simulación en realidad virtual bajo un enfoque de bajo costo mediante el uso de software accesible.

La arquitectura propuesta permitió desacoplar la lógica del conocimiento médico (digitalizada mediante estructuras JSON modulares) de la capa de ejecución visual en Unity y Blender. El procesamiento semántico mediante la API de Gemini 1.5 Flash mostró consistencia en la generación de información adicional (payloads) estructurada, permitiendo que las entradas de lenguaje natural se traduzcan de manera simultánea en sintaxis de audio y parámetros de animación procedural coherentes.

Como línea futura de investigación, se contempla la realización de pruebas piloto con estudiantes de la carrera de medicina. Dichos estudios, permitirían medir tanto la eficacia pedagógica, como la usabilidad del sistema y posibles inconvenientes como la latencia percibida y la tasa de retención de

competencias frente a los métodos de simulación tradicional. Asimismo, se plantea como horizonte de desarrollo posterior, la exploración de modelos de lenguaje de código abierto y/o de ejecución local, para eliminar la dependencia de servicios en la nube (que se encuentran sujetos a condiciones variables), así como la incorporación de nuevos casos clínicos, como diferentes tipos de pacientes y entornos, la expansión del esquema JSON para poder especificar mejor la información para la simulación, y la integración de módulos de evaluación, registro y reporte automático del desempeño del practicante al finalizar la sesión.

Conflictos de intereses

El autor declara que no existe conflicto de intereses en la elaboración y publicación del presente artículo científico.

Agradecimientos

Al Ing. Franz Villalpando Amonzabel, director de carrera, y al Centro de Producción Multimedia Interactiva de la carrera de Ingeniería en Diseño y Animación Digital, a cargo del Ing. Pedro Camargo Vargas, por haber proporcionado el equipamiento tecnológico sin el cual el desarrollo y las pruebas de este estudio no hubieran sido posibles.

Referencias Bibliográficas

Aristidou, A., Lasenby, J., Chrysanthou, Y., & Shamir, A. (2018). Inverse kinematics: a review of past, present, and future methods. *Computer Graphics Forum*, 37(5), 211-235. <https://doi.org/10.1111/cgf.13442>

Bhattacharyya, S., Miller, J., & Turner, R. (2024). Large language models as cognitive engines in medical simulation: A framework for dynamic patient generation. *Journal of Medical Internet Research*, 26, e54321. <https://doi.org/10.2196/jmir.54321>

Calvo, J. (3 de abril de 2025). Implementación del Examen Clínico Objetivo Estructurado (ECOE) con evaluadores internacionales. Noticias USFX. Universidad Mayor, Real y Pontificia de San Francisco Xavier de Chuquisaca. <https://usfx.bo/2025/04/03/la-carrera-de-medicina-implementara-el-examen-clinico-objetivo-estructurado-ecoe/>

Clusmann, J., Kolbinger, F. R., Muti, H. S., Carrero, Z. I., Eckardt, J. N., Laleh, N. G., ... & Kather, J. N. (2023). The future landscape of large language models in medicine. *Communications Medicine*, 3(1), 141. <https://doi.org/10.1038/s43856-023-00370-1>

Cook, D. A., & Triola, M. M. (2009). Virtual patients: a critical review of their use in medical education. *Medical Education*, 43(4), 303-311. <https://doi.org/10.1111/j.1365-2923.2009.03301.x>

Epstein, R. M. (2007). Assessment in medical education. *New England Journal of Medicine*, 356(4), 387-396. <https://doi.org/10.1056/NEJMr054784>

Google. (2024). Gemini 1.5: Unlocking multimodal understanding across a million tokens of context (Technical Report). Google DeepMind. <https://arxiv.org/abs/2403.05530>

Harden, R. M. (2016). Misconceptions and the OSCE. *Medical Teacher*, 38(1), 26-31. <https://doi.org/10.3109/0142159X.2015.1087483>

Khan, K. Z., Ramachandran, S., Gaunt, K., & Pushkar, P. (2013). The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: An historical and theoretical perspective. *Medical Teacher*, 35(9), e1437-e1446. <https://doi.org/10.3109/0142159X.2013.818634>

Kyaw, B. M., Saxena, N., Posadzki, P., Vseteckova, J., Nikolaou, C. K., George, P. P., Divakar, U., Masiello, I., Tudor Car, L., & Car, J. (2019). Virtual reality for health professions education: Systematic review and meta-analysis by the Digital Health Education Collaboration. *Journal of Medical Internet Research*, 21(2), e12959. <https://doi.org/10.2196/12959>

Saaranen, T., Kasanen, M., & Haatainen, K. (2024). Ethical considerations of generative artificial intelligence in health professions education and simulation: A narrative review. *Nurse Education Today*, 132, 106012. <https://doi.org/10.1016/j.nedt.2023.106012>

Smee, S. (2003). ABC of learning and teaching in medicine: Skill based assessment. *BMJ*, 326(7391), 703-706. <https://doi.org/10.1136/bmj.326.7391.703>

Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Liu, X. (2023). Large language models in medicine. *Nature Medicine*, 29(8), 1930-1940. <https://doi.org/10.1038/s41591-023-02448-8>