

**DANIEL ULISES CAMPOS DELGADO**  
*duc@dciencias.uaslp.mx*  
FACULTAD DE INGENIERÍA, UASLP



# La transcripción médica en la era de la inteligencia artificial

Todos nos vemos abrumados cuando el trabajo se apila, no logramos sacarlo adelante, y por lo tanto, se acumula progresivamente. En este sentido, visualizar un médico durante una larga jornada de trabajo, con una rutina diaria de al menos 20 pacientes, y al final de cada jornada teniendo que documentar cada expediente con el mayor cuidado posible, pues los detalles se pierden con el tiempo. En definitiva, esta carga administrativa puede consumir horas extras del día laboral y afectar la calidad de vida del médico. ¿Existe alguna herramienta que pueda facilitar esta rutina de documentación? La huella de la inteligencia artificial (IA) en la medicina es innegable, desde las tecnologías para apoyar el diagnóstico médico, el descubrimiento y la optimización de nuevos fármacos, el diseño de tratamientos personalizados y la optimización de los flujos de trabajo en la clínica. En particular, esta columna quiero dedicarla a una herramienta para eficientizar el trabajo de los especialistas en medicina, la documentación o transcripción médica. Este es un proceso que consume tiempo del médico, y que, en la actualidad, ya puede ser automatizado por herramientas de IA, con una muy buena precisión. La base de estas herramientas son las estrategias de conversión “voz-a-texto” (“voice-to-text”) basadas en IA; inclusive también existe el proceso inverso, es decir “texto-a-voz” (“text-to-voice”), que en otra columna describiremos.

Estas herramientas pueden implementarse en-línea, por ejemplo para obtener notas a partir del audio generado en una videoconferencia, o fuera-de-línea a partir de un archivo previamente grabado en formatos WAV, MP3 o MP4. Quizás ya más de uno haya empleado estas tecnologías, por ejemplo durante una reunión en Zoom, Google Meet o Microsoft Teams. En una conversión “voz-a-texto”, primero se reconocen los sonidos en el audio, luego se entiende su significado (“modelos de lenguaje” de IA), y finalmente se convierte en texto claro. Las ventajas de estas herramientas son diversas, entre ellas, la precisión de las transcripciones (>90%), el soporte y traducción entre varios idiomas, la versatilidad de su integración con otras tecnologías, y su automatización.

Ahora, para que esta tecnología pueda ser exitosa en el ámbito médico, debe ajustarse al lenguaje del especialista, ya que existen términos y conceptos muy particulares que deben distinguirse y describirse puntualmente, por ejemplo

para un ginecólogo, un cardiólogo, un pediatra o un oftalmólogo. También pueden distinguirse diversos escenarios de aplicación: durante la interacción con el paciente, o durante el trabajo de análisis y planeación. Por ejemplo, un pediatra podría dictar sus observaciones mientras revisa a su pequeño paciente, y la herramienta generaría automáticamente la evaluación clínica; o el cardiólogo podría dictar el tratamiento a seguir mientras analiza los resultados de sangre del paciente y contrasta con su ultrasonido. La ventaja de estas tecnologías de IA es que solamente cambiando el modelo de lenguaje se puede ajustar al escenario práctico, es decir tienen una alta versatilidad. Sin embargo, aún deben considerarse como un elemento de apoyo al médico, es decir las transcripciones generadas deben ser revisadas y certificadas por cuestiones éticas y de precisión; ya que es bien sabido de las “alucinaciones” de los modelos de IA. La tecnología ayuda, pero la última palabra siempre será del especialista. En el mercado existen varias plataformas que ofrecen este servicio, enfocadas al ámbito médico, por grandes compañías tecnológicas como “Google Speech-to-Text Clinical Conversation”, “OpenAI Speech-to-Text”, “Amazon Transcribe Medical”, y “Microsoft Azure Speech-to-Text engine”. Las herramientas de Amazon, Google y Microsoft pueden emplearse en la modalidad fuera-de-línea y no tienen costo por un periodo de prueba o por un número limitado de “créditos”. Por otro lado, la herramienta de OpenAI “Whisper” es la única que es totalmente abierta, multi-idioma y con diferentes modelos de lenguaje.

Cabe mencionar que estas herramientas ya son una realidad en diversos hospitales en el mundo, aunque su costo puede restringir su acceso por el momento. De hecho, Microsoft anunció en 2025 su nueva plataforma “Microsoft Dragon Copilot” como un ecosistema completo para la creación de la documentación de los pacientes, el dictado automático por voz, la generación de notas y el seguimiento de los expedientes electrónicos. Ahora, dicha tecnología está aún lejos de nuestro día a día en México, sin embargo, conforme se implemente a nivel mundial sus costos disminuirán y podrá ser una herramienta poderosa para reducir la carga de trabajo en los hospitales y las clínicas. Por lo que si se adopta adecuadamente podría devolver a los médicos algo invaluable: tiempo para escuchar e interactuar con sus pacientes. **UP**