

**LA LLEGADA DE LA INTELIGENCIA ARTIFICIAL Y
EL PROBLEMA DE LA EVALUACIÓN EN LA DOCENCIA UNIVERSITARIA.
EL SISTEMA EDUCATIVO EN CRISIS**

Carlo Stella Serrano

*Máster en Desarrollo Económico y Políticas Públicas.
Universidad Autónoma de Madrid*

RESUMEN

Antes de la llegada de la Inteligencia Artificial (IA) el sistema de evaluación universitario demostraba ciertas flaquezas que lo hacían cuestionable desde muchos ángulos. Este artículo hace un repaso de las características y consecuencias de la IA con particular énfasis en la evaluación de la docencia en el mundo académico. Se sugiere que no tiene sentido luchar contra el fenómeno de la IA ya que formará parte dentro de nuestras vidas tal y como lo hace internet o el teléfono móvil. La solución pasa por incorporar dichas ventajas al sistema de aprendizaje, modernizándose y renovándose éste a su vez.

Posteriormente se hace un repaso crítico del sistema de evaluación actual, diferenciando entre el método de evaluación tradicional y el método de evaluación continua y la dificultad que encuentran ambos métodos en valorar los aspectos más teórico-académicos con aquellos más prácticos o subjetivos. En dicha línea se proponen una serie de cambios dentro del modelo de evaluación y de enseñanza con el objetivo de lograr un modelo más útil en la vida personal y profesional del estudiante universitario, pero, sobre todo, más motivador. Se sugiere una evaluación mixta, adaptable para cada asignatura o materia y que, de importancia al aspecto práctico y al desarrollo de nuevas competencias como la reflexión, el análisis o el pensamiento crítico que, por otra parte, van a ser las más demandas en la era de la IA. Por último, se concluye haciendo una reflexión sobre el sistema universitario que se quiere para el futuro.

1. INTRODUCCIÓN

La tecnología ha revolucionado todos los aspectos de nuestra vida desde el ámbito más cotidiano -como en las tareas del hogar- hasta el ámbito profesional -como en la gestión de la agenda y la realización de tareas-. En la educación, la tecnología ha sido y es una herramienta muy útil que facilita el aprendizaje. Es más, con el internet y los buscadores tenemos a nuestro alcance todo el conocimiento del mundo. Se da la curiosa paradoja de que vivimos en un mundo de sobre información y contenidos en abierto: cuanta más información hay más difícil es discernir entre la buena y la mala (por ejemplo, en el mundo de la comunicación instantánea o redes sociales y la proliferación de *fake news*, noticias inexactas, manipuladas o denigratorias).

Mientras la tecnología se utiliza como herramienta indispensable en el sistema educativo (desde colgar las presentaciones y el uso de herramientas tradicionales como hipertexto, blog, vídeos elaborados por el profesorado en las plataformas virtuales y la toma de notas y apuntes directamente en el ordenador hasta la realización de tareas, ejercicios, informes o presentaciones por parte del estudiantado), el sistema de evaluación se ha mostrado siempre bastante más “analógico”, es decir, reticente a la incorporación de tecnología como herramienta o medio para su realización o con un uso parcial o minoritario. El

problema radica en que el mundo fuera del aula está avanzando a un ritmo más rápido que el mundo dentro del aula generando así un montón de incoherencias debido al despliegue tecnológico y posiblemente dejando “en ridículo” el sistema de evaluación ante la vertiginosa presencia de la tecnología disruptiva y sus posibles usos (y mal usos).

Recientemente una herramienta disruptiva nueva ha visto la luz para su utilización en masa: el ChatGPT que, lanzado a finales del 2022, es una aplicación a primera mano de lo que es la Inteligencia Artificial (IA en adelante). El ChatGPT, según su propia respuesta es “un modelo de lenguaje capaz de responder a preguntas y mantener conversaciones con personas en una forma natural”. ChatGPT es sólo uno de los múltiples chatbots de IA que existen. Éste está desarrollado por OpenAI y tiene múltiples versiones, cada una más poderosa que la anterior.

Por otra parte, Microsoft ha desarrollado en colaboración con OpenAI su Bing Chat y Google está desarrollando su chatbot equivalente el Bard. El factor revolucionario de estos buscadores es que son un híbrido entre un buscador tradicional y una conversación humana. Precisamente de eso se trata la IA, de lograr que la tecnología resulte lo más cercana a las interacciones humanas.

De hecho, tan trascendente es el tema que la RAE ha escogido a la Inteligencia Artificial como la palabra del año 2022. En concreto la define como “una disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico”. En concreto el chatbot GPT es tan poderoso porque es un sistema en continuo aprendizaje. Sus siglas GPT (*Generative Pre-trained Transformer*) indican lo novedoso que es el sistema puesto que es “transformador” ya que utiliza redes neuronales que procesan el texto de entrada y generan uno de salida; “pre-entrenado” porque se entrena en el lenguaje humano para ajustarlo a tareas específicas; y “generativo” porque genera un texto nuevo cada vez en función de los inputs que reciben (Pedreño, 2023).

El poder de esta herramienta radica en su capacidad de aprendizaje profundo (*deep learning*) que genera un output creíble. La novedad no es la IA generativa, sino las técnicas de aprendizaje automático con capacidades crecientes que, mediante la estructura de algoritmos en capas, crean una red neuronal artificial que aprende y toma decisiones como un humano (Muñoz, 2023.). Los usos de la IA son infinitos, pero muchos de ellos generan gran repulsa e indignación, como el uso para la vigilancia social mediante análisis facial, el control y manipulación de redes sociales, la creación de contenido falso, el uso indebido de datos personales o los robots autónomos como los drones usados en la guerra (Martínez, 2023).

A raíz del proceso de aprendizaje de la IA surge toda la problemática de la protección de datos de los usuarios. Hasta ahora, la UE tiene publicado un Libro Blanco sobre la IA (2020) que se centra en minimizar los riesgos para los derechos fundamentales y advierte de impacto que pueden tener las aplicaciones de IA de alto riesgo, pero sigue trabajando intensamente para “proteger” de los posibles efectos indeseados de la IA a todos los ciudadanos europeos y no europeos

Los acontecimientos van tan rápido que, al finalizar este artículo, me doy cuenta de que ya se ha desactualizado. El 23 de marzo de 2023 salió la versión de Chat GPT-4 que es capaz de analizar entradas en imágenes (tanto fotos, como problemas matemáticos manuscritos), por no decir que entiende bromas y memes. Por si esto pareciera poco, la capacidad creativa del GPT-4 es aún mayor y la exactitud de sus respuestas también.

La llegada de esta herramienta al mundo académico supone al mismo tiempo un potencial y un riesgo puesto que va mucho más allá de un simple buscador ya que es capaz de efectuar tareas humanas como el resumen de un texto y la generación de un ensayo (por no citar otras funciones). El potencial de generación de contenidos con dicha herramienta es infinito.

Ello hace necesario replantearse todo el sistema educativo con particular énfasis en el sistema de evaluación (muy basado en memorizar información y contenidos y realizar ensayos y trabajos fuera del aula) que debe cambiar para adaptarse a las nuevas tecnologías disponibles. Demonizar la IA y querer luchar contra ella es cómo querer luchar contra el fenómeno de las redes sociales, siendo por tanto una batalla perdida, al menos fuera del aula. Por tanto, una primera premisa para mejorar la calidad docente y la evaluación incorporando la IA pasará por evitar la masificación en el aula (grupos de 60-60 estudiantes parecen inviables) y por introducir criterios de integridad académica y honestidad en el proceso de enseñanza-aprendizaje (tanto por parte del profesorado como del estudiantado).

A continuación, se harán una serie de reflexiones sobre el fenómeno de la IA en la enseñanza y sobre el sistema de evaluación, que ya parecía mostrar sus flaquezas mucho antes de la llegada de la IA al circuito de uso comercial. La calidad de la enseñanza futura dependerá de su capacidad de adaptación esta nueva revolución tecnológica, pero sobre todo de un nuevo sistema de evaluación que sepa incorporar aspectos que valoren a la persona en su conjunto y vaya más allá de un rígido sistema escalar autoimpuesto.

2. LA INTELIGENCIA ARTIFICIAL: CARACTERÍSTICAS Y CONSECUENCIAS

No cabe ninguna duda de que la IA y sus versiones comerciales disponibles para todos los usuarios en versiones de chats inteligentes suponen un auténtico elemento disruptivo en nuestras vidas. La misma IA es el motor clave en la cuarta revolución industrial debido a la su capacidad de aprendizaje (WEF, 2023). Lo disruptivo en este tipo de IA es que de repente se integran funciones propias del ser humano en la tecnología o en su interacción. Hasta entonces el internet de las cosas (IoT) “comercial” abarcaba un radio limitado como el cumplimiento de órdenes concretas o básicas en el hogar (por ejemplo, en todas las funciones de los altavoces inteligentes o asistentes virtuales cómo controlar el encendido de las luces de casa con unas palabras, programar la calefacción, solicitar una nueva *playlist* o preguntar qué tiempo hace fuera). Sin embargo, el nacimiento de dichos asistentes en formato de chat ha generado toda una revolución puesto que van más allá de las funciones previamente atribuidas a un buscador de información (tales como Google o Bing) o a un asistente virtual (tales Siri o Alexa), sino que son capaces de dar respuestas de forma humana y aprender.

Tan poderosa es la IA que no solo otorga una respuesta (output) a una cuestión (input), sino que es capaz de hacer mucho más: puede programar, analizar textos, recopilar información dispersa, solucionar problemas matemáticos, generar conversaciones/escenarios, hacer resúmenes, encontrar similitudes/diferencias y redactar ensayos. Por si fuera poco, el modelo de Bing Chat es capaz incluso de citar las fuentes de donde obtiene la información o de citar paso a paso la resolución de un problema matemático. El Chat GPT-4 destaca entre los modelos anteriores por el potencial aún mayor que tiene: acepta múltiples formatos de entradas/inputs (fotos, textos, imágenes); compone canciones; redacta discursos; escribe guiones; aprende el estilo de escritura... etc. ¿Dónde está el límite a la capacidad de estos dispositivos? Y por si pareciera poco, en el ámbito de la literatura, el Chat GPT se está usando para publicar nada menos que libros y artículos de ciencia ficción y las editoras han tenido que decir “basta”. Eso significa que si se utiliza para publicar o crear contenido ¿hay que otorgarle los correspondientes derechos de autor? ¿Hasta dónde nos sorprenderá los usos que tiene la IA? Como vemos este es un tema muy complejo y por mucho que se intente prohibir el uso de la IA, se va a seguir usando porque hasta ahora no hay detector válido que diferencie el texto humano del no-humano.

En el ámbito educativo, la llegada de la IA ha generado un sinnúmero de reuniones académicas y momentos de crisis entre los profesores y varias instituciones como los colegios de Nueva York o la prestigiosa SiencesPo han decidido prohibirla en parte con algunas matizaciones. En relación con el tema de los datos personales, la agencia Italiana de Protección de Datos, conocida como *il Garante* ha sido la primera en tomar cartas en el asunto. El 31 de marzo 2023 decidió prohibir su uso por la recogida ilícita de datos y ausencia de sistema de verificación de edad. La Unión Europea por su parte tiene activada toda su maquinaria para sacar un Reglamento de Inteligencia Artificial. Aunque desde la UE ya tenían una propuesta redactada en 2021 (después del surgimiento del Libro Blanco), ésta no se

imaginaba el potencial de los nuevos *chatbots*. A raíz de los acontecimientos en Italia, en apenas 11 días el Consejo adaptó el borrador e impulsó los trámites para la aprobación de la ley.

Además, con motivo de la prohibición en Italia, el Chat GPT-4 incorporó una nueva configuración para mejorar la privacidad de los usuarios. Esta nueva versión permite evitar el almacenamiento del historial y que la información que el usuario comparte (o introduzca en el input) no se utilice para el entrenamiento o aprendizaje de la máquina. Aunque, ¿qué garantías tenemos sobre esto? Al final es una herramienta gratuita, y las cosas gratuitas se pagan de varias formas, entre ellas con nuestros datos. Tras esta modificación, el ChatGPT volvía a estar disponible en Italia, aunque a la velocidad que parecen estar yendo los avances, todo apunta a que siga habiendo muchas novedades en el futuro.

Toda esta característica de la IA pone en la encrucijada al sistema educativo. La integridad académica, definida como la adopción y adecuación de los valores fundamentales que deben regir toda actividad académica, siendo éstos: honestidad, confianza, responsabilidad, respeto y justicia o equidad (Comas, 2011) es una de las vías por las que la Universidad debe apostar. Las prácticas deshonestas hasta ahora conocidas se quedan “pequeñas” en comparación con el potencial que tiene la IA. El problema está en que la IA es capaz de aprobar y sacar buenos resultados en exámenes oficiales, como viene demostrando ya en varias ocasiones. Ya en su versión anterior, el GPT-3 aprobó el examen final del MBA de la Wharton School de la Universidad de Pennsylvania, el examen de acceso a la abogacía en EE.UU. conocido como el Bar exam o incluso el examen para obtener la licencia médica en EE.UU. (Múñoz, 2023). En España, la revista Redacción Médica sometió al Chat GPT al MIR que superó con un 25% de respuestas incorrectas.

Los niveles de innovación y disrupción en el plano de posibilidades no parecen terminar. Otro elemento disruptivo no citado en la lista anterior, es la posibilidad que tiene la IA de generar imágenes, video y audio basados en personajes reales. Este tipo de producción gráfica se la conoce como *synthetic media* y a su contenido mediático *deepfake* (ultra falso). Hay personas especializadas en la creación de dicho contenido. Por tanto, cada vez es más difícil diferenciar lo falso de lo real y cada vez hay mayor facilidad para generar bulos. Nos dirigimos hacia un mundo donde no distinguiremos la ficción de la realidad y esto abona el terreno para la desconfianza general hacia todo tipo de imágenes, incluso las reales (Gonzalo, 2023). Parece ser que una nueva necesidad va ser formar a los humanos en diferenciar lo real de lo falso, que en el caso del *synthetic media* puede intentarse observando lo siguiente: las luces y sombras, los reflejos en los ojos, los detalles en las articulaciones del cuerpo, la calidad, las proporciones del cuerpo, la perfección en los cuerpos, la regularidad en los rostros o las exageradas expresiones faciales), sin pasar a enumerar las implicaciones del metaverso en la educación.

El revuelo que está causando la IA es tan grande que el Instituto *Future of life* patrocinado por Elon Musk ha hecho un llamamiento conjunto para realizar una moratoria sobre la IA por los principales riesgos que conlleva. En la presentación de la versión del GPT-4, el propio CEO de OpenAI manifestó públicamente su preocupación por los usos delictivos (desinformación o ciberataques) de su creación, añadiendo, “creo que la gente debería estar contenta de que estamos asustados sobre esto”.

Desde su utilización masiva en el ámbito académico hasta el potencial de desinformación a gran escala que tiene pasando por los potenciales usos ilícitos o delictivos, no ha habido tiempo para regular dicho ámbito. Los avances en este campo están ocurriendo a un ritmo vertiginoso lo que me hace dudar sobre si su salida a mercado no era algo cuidadosamente planificado con anterioridad. De hecho, el panorama es bastante confuso. Resulta extraña la postura de Elon Musk que encontrándose inmerso en el desarrollo del Chat con OpenAI y Twitter (Chat GPT se construyó utilizando datos de las interacciones de usuarios en Twitter), para posteriormente cortar por completo la relación, reclamar una moratoria de 6 meses y advertir de los riesgos que tiene la IA para la humanidad. Todo ocurría cuando a finales de abril del 2023 anunciaba la creación de su propia empresa de IA, llamada X.AI, y una noticia del NY Times mencionaba su “larga y complicada historia en relación a la IA, gobernada por sus contradictorias

visiones sobre la tecnología”. Por lo tanto, hasta en el ámbito de los máximos exponentes e impulsores de la IA el panorama es confuso.

En el ámbito educativo, una de las características principales de la IA es la *humanización en la respuesta* (en este caso, escrita). Lo realmente impresionante y revolucionario de la IA es la humanización de la respuesta tanto en la naturalidad del lenguaje como en la capacidad creativa. Ésta última hace que cada respuesta dada sea distinta de la anterior (capacidad generativa), lo que difícilmente se traduce en una respuesta monótona o “robotizada”. Otros factores impresionantes son la inmediatez del proceso, la precisión de la respuesta (sin entrar en la veracidad o certeza de la misma) y los múltiples usos que puede tener previamente mencionados.

Esto no significa que la labor del humano vaya a desaparecer puesto que deberá aprender a preguntar a la IA con el método deductivo e inductivo. Por tanto, otra característica es la *necesidad de supervisión humana en el uso de la inteligencia artificial*. Los outputs deben ser analizados e interpretados por un humano puesto que es posible que las respuestas sean incoherentes. La revisión y corrección resulta necesaria en todo caso. Precisamente sobre esto matiza Llorens (2023), que resalta la importancia de “tener en cuenta que la IA no puede reemplazar completamente al profesorado humano, ya que la enseñanza y el aprendizaje son actividades complejas que requieren la interacción humana y el pensamiento crítico”.

Ciñéndose al ámbito académico, se pueden citar las siguientes consecuencias:

- *El riesgo de plagio o fraude*. Éste es el principal problema en los trabajos académicos puesto que la capacidad creativa de la IA permite obtener respuestas distintas equiparándolas a contenido redactado por un humano. De esa forma los detectores de plagio (por ejemplo, el Turnitin usado en las universidades) ya no tienen ningún sentido. Esto hace que haya que repensar por completo el método de evaluación, de lo que se hablará en el próximo apartado.
- *De la búsqueda activa de información a la actitud pasiva*. La principal diferencia con respecto a un buscador tradicional es que el uso de la IA permite usar un output desde una posición totalmente pasiva ante el aprendizaje. Bajo esta dinámica, el output dado por la IA es directamente la respuesta a algo concreto lo que anula la necesidad de buscar, sintetizar, analizar o contrastar información. Por ejemplo, una cosa es tener que redactar un ensayo buscando activamente las distintas fuentes de información (requiere un esfuerzo en analizar, sintetizar, vincular conceptos...etc.), otra cosa es redactar el ensayo desde las respuestas inmediatas y exactas sobre el tema que se ha preguntado (no existe creatividad humana, solo se potencia la capacidad de “adornar” lo que viene a ser un “corta y pega” de información).
- *La desaparición del buscador tradicional*. Siguiendo el punto anterior, el buscador tradicional que implica una búsqueda activa de información (contraste de páginas webs, fuentes...etc.) para formar una opinión personal puede perder peso frente el uso de los Chats de IA que van al grano de la pregunta (sin generar valor añadido por la inteligencia natural o humana). De hecho, el buscador tradicional –anticipándose a la llegada de la IA- ya había modificado la forma en la que presentaba los resultados de la búsqueda: Google mostraba ya una respuesta rápida sin necesidad de entrar en el link de la página web, bajo el apartado de “Otras preguntas de usuarios”. Esta característica también tiene que ver con el desplazamiento de la búsqueda activa hacia la búsqueda pasiva.
- *Desde el inicio del pensamiento al fin del mismo*. La forma en la que se utiliza la IA corresponde a una decisión estrictamente personal que determinará si se utiliza para marcar el inicio del pensamiento o el fin del mismo. Si su utilización corresponde a un uso que permite fomentar la reflexión crítica, profundización en un tema o discusión sobre las respuestas entonces se estará usando la IA de forma constructiva. En el ejemplo del ensayo, entra en juego el factor “ética e integridad” de cada persona que será determinante en el tipo de intenciones que se tiene en el uso de la IA.
- *El potencial de aprendizaje*. Uno debe desconfiar de los servicios que se ofrece de forma gratuita porque generalmente se pagan con el valor de nuestros datos personales (por ejemplo,

para utilizarlo es necesario ceder el número de teléfono, correo electrónico...etc.) (incompatible con la protección de datos en la Unión Europea). Pero más allá de eso la IA tiene el potencial de aprender a través de las preguntas que se le hacen (capacidad de aprendizaje o entrenamiento), lo que con el paso del tiempo podrá mejorar aún más las respuestas.

- *El potencial uso en el ámbito investigador.* Si ya en el ámbito de la docencia el ChatGPT puede parecer revolucionario, en el ámbito investigador lo es aún más. Es necesario incorporarlo en la investigación porque cómo indica Pedreño (2023), el futuro avanza hacia un contenido abierto y colaborativo. El potencial de la IA debe explotarse para lograr avances científicos relevantes (en algunos campos como el de la Medicina existen múltiples ejemplos). Además, otra curiosidad es que, aunque se desapruebe su “coautoría” en los trabajos (cosa que ahora parecería algo extraña), esto no quiere decir que no se haya empleado, por lo tanto, aparezca o no, la IA va a permanecer una herramienta siempre disponible y accesible.
- *La generación de desigualdades y brechas.* La permisividad de su uso o no podrá generar grandes desigualdades entre personas, centros o países en cuanto a los resultados de las publicaciones en el ámbito profesional (por ejemplo, con la extraña presión por publicar “papers” la IA puede agilizarlos enormemente). Dado que prohibir semejante herramienta no parece factible, se hace necesario regular su uso y otorgarle la respectiva autoría cuando se utilice (lo que está directamente relacionado con la ética y la integridad académica).

En relación con esto último, el profesor F. Llorens ha sido el primero en publicar una entrada en Universidad con el ChatGPT en coautoría. Según este autor, la IA desempeñará un papel clave en la escritura de textos académicos, no tanto porque sustituirá al estudiante sino porque lo ayudará a estructurar, organizar o corregir sus ideas. Creo que, en el caso de los ensayos académicos, todo depende del tipo de uso que el escritor dé a la IA porque de ser un “soporte” para la investigación y la creatividad humana a utilizar el “copia y pega” directamente obtenido del ChatGPT hay bastante diferencia.

El potencial de los chats de IA es enorme y mucho mayor de lo que podemos imaginar. Como indican los expertos, el ChatGPT sólo es la “punta del iceberg” de lo que está por llegar (Pedreño, 2023). Algunos muy catastrofistas auguran que es el “fin de la humanidad” tal y como la entendemos ahora. En este caso parece conveniente adoptar una postura racional e intermedia. Aunque haya tecnófilos y tecnófobos, lo que es peligroso es que se conviertan en tecnófilos ingenuos y los tecnófobos recalcitrantes (Llorens, 2022). En el ámbito académico la solución no pasa por prohibirlo sino por incorporarlo de la forma más constructiva posible. Su uso para la realización de trabajos concretos, plantear preguntas, aplicar conocimiento o profundizar sobre un tema puede transformar y mejorar la calidad de la educación. Todo dependerá del uso que se le dé para sacarle el máximo provecho y siempre seguir aprendiendo (antes hablábamos de la importancia del intercambio de conocimientos profesor-estudiante, ahora podemos hablar del intercambio de conocimientos profesor-máquina-estudiante). Más que en el lado del estudiante, la dificultad pasa a estar en el lado del docente: éste deberá enfocar la pregunta de tal forma que permita el desarrollo de cualidades humanas (por ejemplo, pensamiento crítico, síntesis, análisis, evaluación...etc.) que la IA no puede hacer (aún).

Por otra parte, es posible que por naturaleza ciertas tareas o artículos de opinión/reflexión jamás puedan sustituirse por la IA. Por ejemplo, pienso en este mismo artículo que al haberlo finalizado incorporaba más de 22 horas de edición total (visto en el contador de información del archivo de Word) y más de 5 horas de supervisión de mi tutora académica en el programa de posgrado. Mi método de redacción ha sido el de lectura, recopilación y reflexión personal. ¿Podría haberlo hecho de igual forma la IA? O mejor aún, ¿hasta qué punto el uso del ChatGPT podría haber reducido el tiempo de edición y mejorado el contenido? Son preguntas que dan miedo porque en ciertas tareas (sea la enumeración de características, ventajas o desventajas), la IA seguramente sí que podría haber ayudado bastante más de lo que uno quisiera creer.

Por último, conviene también citar las limitaciones de la IA como el sesgo de datos; la inexactitud de las afirmaciones (argumentos extremadamente convincentes pero incorrectos); la fuga de información

confidencial y la controversia sobre derechos de autor (Pedreño, 2023). Ciertamente es que ninguna de estas cosas es nueva en el mundo digital. Por ejemplo, el sesgo de datos es algo que existe en la misma ciencia, precisamente el trabajo científico inicial consiste en encontrar los mejores datos para estudiar un problema concreto y poder cumplir con el enfoque ARE (Argumentos-Razonamientos-Evidencias). Quizás debería asustar algo más la fuga de información confidencial (que actualmente es una de las grandes preocupaciones de las organizaciones públicas y privadas del mundo) o los derechos de autoría en la multitud de trabajos que está realizando la IA. Lo primero es un riesgo importante en cuanto se permite a la aplicación acceder a los ficheros personales del ordenador y lo segundo plantea el problema grave de no conocer fuente (cosa que soluciona el Bing Chat). Esta mejora es una muestra más del potencial que tienen estos sistemas de IA y de lo rápido que pueden evolucionar, más aún cuando se piensa en el potencial de aprendizaje (“entrenamiento”) que tiene dicha aplicación. Es normal y comprensible que estos avances nos asusten.

3. EL SISTEMA DE EVALUACIÓN DOCENTE EN LA ENCRUCIJADA

La adaptación del sistema universitario español al Espacio Europeo de Educación Superior (EEES) ha traído consigo el modelo de evaluación continua en el sistema de aprendizaje (Vallespín, 2023). El objetivo era el de evaluar las competencias y habilidades que van más allá de la mera incorporación de conocimiento memorístico y, por tanto, incorpora procedimientos de retroalimentación continua a estudiantes. Esto se diferenciaba del modelo de evaluación tradicional que se basaba primordialmente en exámenes parciales y un examen final. Con la evaluación continua, la evaluación final de la signatura se dividió en dos partes: una que evaluaba las competencias y habilidades a través de trabajos prácticos en el aula (autoevaluaciones, ensayos, informes, presentaciones, prácticas, actividades o ejercicios) y otra que se correspondía al examen final de la asignatura.

Ya desde hace tiempo, ha habido muchas críticas al sistema de evaluación continua pero hoy aún más, con la aparición de los chats de IA, este sistema parece perder aún más legitimidad. Esto se da por el hecho de que los trabajos de la parte “continua” (ya sean ensayos, informes, notas o recopilaciones, entre otros) pueden efectuarse de forma íntegra con los chats de IA. De hecho, varios estudiantes tanto de bachillerato como de universidad ya lo han hecho y esto está provocando un auténtico escándalo o revolución en los colegios y universidades. Por lo tanto, ¿el profesor evalúa al estudiante o a la máquina? El sistema de evaluación docente se encuentra en la encrucijada de evaluar aptitudes que van más allá del contenido teórico que tan fácilmente puede reproducir la máquina de IA como antes lo hacían los estudiantes.

En la universidad, es cada vez más habitual que los docentes comenten formal e informalmente sobre el descenso gradual en cuanto a preparación y aptitud de los estudiantes que entran en primer curso. Esto está principalmente caracterizado por las mayores diferencias interpersonales entre las capacidades de los estudiantes dentro de una clase. No obstante, una posible explicación a esta anomalía es el efecto de la pandemia del Covid-19 para los que se encontraban estudiando bachillerato en los años 2019/20 y 2020/21, por no citar los que entraron a la universidad durante uno de esos cursos. La docencia virtual, la semi-presencialidad, los efectos sobre la salud mental, la falta de un lugar para estudiar y las varias preocupaciones vividas durante dicho periodo podían haber complicado el proceso de aprendizaje. En concreto, las universidades públicas (y todos sus servicios) cumplen una función de equidad social a través de la igualdad de oportunidades de acceso a la educación universitaria puesto que permiten a todos los estudiantes potenciar en mayor o menor medida sus habilidades dentro de las mismas condiciones de partida. Es indudable que la pandemia perjudicó el estado físico y mental de los jóvenes, por la falta de actividad, como por la cantidad de tiempo pasado delante de dispositivos electrónicos. En cuanto al lado académico, el cierre de los centros afectó negativamente el rendimiento de cada individuo dependiendo de la situación socio-económica o familiar en la que pudo encontrarse el estudiante. Esto ocurre porque los jóvenes en familias de renta baja suelen vivir en condiciones que hacen difícil el aprendizaje en el hogar (Cifuentes-Faura, 2020).

Sin embargo, pasada la pandemia, el ámbito académico vuelve a verse afectado por un nuevo fenómeno, y éste es el de la llegada de los chats de IA. El problema de la evaluación vuelve a parecer como tema principal, incluso más que la dinámica dentro del aula. A continuación, se hacen unas reflexiones sobre el sistema de evaluación docente en la actualidad desde mi perspectiva actual como estudiante de un Máster Oficial en la Universidad Autónoma de Madrid:

- *El problema de evaluación continua por exceso.* Para Vallespín (2023) el problema de la evaluación continua *por exceso* implica que, con el calendario académico comprimido en cuatrimestres, la evaluación continua genera una sobre carga de trabajo adicional entre pruebas que se solapan en el tiempo y complican la vida de los estudiantes. Esto genera un doble problema tanto para los profesores (por excesiva burocratización) como para los estudiantes (que les complica la compatibilización universidad-trabajo). Es más, tanto énfasis en la evaluación continua podría dejar de lado el tiempo de estudio necesario para adquirir conocimientos esenciales de la asignatura.
- *El problema de la evaluación continua por defecto.* De la misma manera Vallespín (2003) indica que la evaluación continua *por defecto* implica que en la parte de “continua” se incorporen exámenes parciales en su cómputo. De esa forma el sistema de evaluación continua pierde toda su esencia cuando las competencias y habilidades que se pretenden evaluar se equiparan a contenidos evaluables de forma mucho más rápida (por ejemplo, exámenes parciales tipo test o autoevaluaciones semanales tipo test). Sin embargo, los parciales liberatorios tipo test podrían ser la solución para la evaluación sin la IA.
- *El problema del trabajo en grupo.* La dinámica de la evaluación continua suele ser la misma en todas las asignaturas, componiéndose de un trabajo y presentación en grupo con el objetivo de potenciar las habilidades de comunicación oral e interpersonal. Estas habilidades son muy importantes, sin embargo, realizar la misma dinámica durante los cuatro años de carrera puede resultar tedioso. Más que nada porque si los trabajos que se encargan son extensos, la aparición de los chats de IA y su uso desvirtúan por completo el objetivo de la evaluación. Además, cabe mencionar que casi nunca se logra una repartición igualitaria de las partes entre los integrantes del grupo. La parte de revisión y edición final del trabajo conlleva un mayor esfuerzo adicional y de coordinación, si bien puede introducir competencias de liderazgo en algunos miembros del grupo.
- *La homogeneización de asignaturas.* Dicho sistema de evaluación continua ha generado tal homogeneización en las asignaturas que las transforma en un modelo rígido en el que todas las asignaturas contienen siempre un trabajo escrito (generalmente informe sobre un tema y presentación) y un examen final. El problema no está en los primeros años, sino en los últimos, en los que la exigencia puede ser mayor y los trabajos de evaluación continua requieren grandes cantidades de inversión en tiempo para hacerse adecuadamente.
- *El riesgo moral en las tareas.* En los trabajos mandados a casa existe siempre un desajuste entre intereses del profesor y del estudiante. Sólo el estudiante sabe cómo ha efectuado la tarea. Si ya antes suponía una dificultad en la evaluación (tanto por el reparto desigualitario de la carga de trabajo, el “copiar” o el hacerse ayudar por otra persona), ahora con la llegada de los chats de IA, éstos tienen el potencial de hacer directamente el trabajo (aunque esto dependerá del tipo de trabajo que se encargue naturalmente).
- *Las diferentes situaciones de partida.* Las tareas de evaluación continua podrían generar desigualdades a la hora de medir el rendimiento porque impactan de forma distinta sobre los estudiantes según la facilidad que tengan para realizarlas (por ejemplo, perjudicando a aquellos que tienen que trabajar, a aquellos que no disponen de un lugar tranquilo en su hogar y muchos otros condicionantes y factores socio-económicos.) Por ello, las actividades prácticas que no se realizan directamente en el aula incorporan una dosis de subjetividad en la medición del rendimiento.

Estas reflexiones hacen reconsiderar el modelo actual de evaluación, teniendo en cuenta que muchas veces la nota no es un fiel reflejo de las capacidades del estudiante. Si lo comparamos con otros

sistemas, el británico otorga mucha importancia a lo que va más allá del mero conocimiento memorístico. Por ejemplo, la estructura de la nota en una típica pregunta de ensayo se divide entre los puntos de conocimiento (*knowledge, K*), aplicación (*application, AP*), análisis (*analysis, AN*) y evaluación (*evaluation, EV*). Por tanto, una pregunta de diez puntos puede tener la siguiente estructura de corrección: 2K, 2AP, 3AN, 3EV. Lo que nosotros entendemos como evaluación (en forma de test, examen o valoración) nada tiene que ver con la evaluación de este contexto. En este caso la evaluación se refiere a una habilidad específica como la de dar una opinión ofreciendo una visión crítica y analítica a la pregunta propuesta, formando parte de la puntuación final de la pregunta.

Otro tema interesante es la fijación del aprobado o el suspenso donde en la valoración puede haber una gran dosis de subjetividad. Por ejemplo, ¿qué diferencia en conocimiento el que tiene un 4.3 (suspenso), del que tiene un 5.1 (aprobado)? ¿Cómo se fija ese umbral mínimo de conocimiento, habilidades o capacidades? De la misma forma se puede hacer el análisis por lo alto: ¿qué mucho mejor es el que tiene un 9.7 (con matrícula) de un 8.9 (notable)? ¿Qué esfuerzo o capacidad adicional permite llegar a esa nota máxima? Quizás todo esto tenga que ver con el rígido sistema de evaluación que hay, donde poner una nota sea la prioridad máxima del profesor.

De esa forma, la crítica del riesgo moral en las tareas mandadas a casa se entiende por el hecho de que existe una necesidad por parte del docente de tener algo escrito que “evaluar”, sin entender realmente que muchas veces estas tareas sobrecargan a los estudiantes (en tiempo en casa y tiempo de clase en presentaciones) sin generar ningún aumento significativo en conocimientos, habilidades o competencias. Con la llegada de la IA, varios docentes han constatado la rapidez con la que los estudiantes realizaban las prácticas de tipo ensayo o trabajo escrito. Si la IA se utiliza para realizar la mayoría de los trabajos y esto se convierte en una norma informal, existe un doble problema: por parte del estudiante la falta de interés (para ir rápido y quitarse de encima los trabajos); pero también por parte del docente donde no se analiza la utilidad e implicaciones del trabajo que se solicita. De hecho, algo bueno de la IA puede ser que marque el fin de los interminables trabajos y tareas en casa fomentando el trabajo en equipo con el profesor en el aula (Pedreño, 2023).

Pensando en las presentaciones en grupo, ¿qué porcentaje de estudiantes está atento u está a otras cosas? O en aquellos primeros años, ¿cuántas veces no se han pactado esas preguntas con los compañeros? Son varios los estudiantes que afirman que la monotonía de los trabajos mandados a lo largo de los 4 años de grado es evidente y poco constructiva. Muchas veces, todo ocurre por la necesidad por parte del profesor de tener algo que evaluar para justificar o tener evidencia sobre una nota. Incluso varios profesores se muestran críticos con esta dinámica que resulta estresante y que de “continua” solo tenga el nombre (Vallespín, 2023). Quizás la vuelta del examen parcial liberatorio podría reducir esa carga evaluadora y permitir conjugar dentro del aula distintas actividades que resulten más fructíferas para el estudiantado mientras que se estudia en casa el contenido más teórico. Por último, pero no menos importante, la forma en la que se evalúa la realización de los TFGs y los TFMs tiene que ser reinventada si no se quiere caer en la trampa de la IA. Habrá que dar más peso en la evaluación a las primeras fases del seguimiento del TFG y del TFM y a las tutorías y a la última fase de presentación del trabajo ante un tribunal donde el estudiante tendrá que demostrar su maestría en el tema, su razonamiento y su pensamiento crítico. También en el caso de los TFMs tendrán que impregnarse de la filosofía *Open Science* (donde los trabajos puedan replicarse con los mismos datos, la recogida de evidencias, etc.)

En el ámbito de las becas universitarias, también existen incongruencias, aunque no vengán provocadas por la IA. Un buen ejemplo es el caso de la valoración de las solicitudes de becas FPU (ayudas pre doctorales otorgadas por el Ministerio de Universidades). La peculiaridad es que en el 2022 ha habido una pequeña reforma en cuanto a su tramitación, pasando de dos a una única fase en la que se valora la nota media del expediente académico (70%) y la trayectoria académico-profesional del candidato (30%). En esta segunda parte se valoran los premios académicos, la formación complementaria, los idiomas, la participación en proyectos de investigación o méritos de docencia u actividad investigadora (valorando participaciones en congresos, cursos, eventos de difusión y

publicaciones académicas). Este último aspecto es el que más polémica ha generado ya que se trata de actividades muy difíciles de llevar a cabo en los momentos previos a la concesión de estas becas.

Por una parte, se genera una discriminación entre aquellos que, por motivos económicos no han podido realizar estancias extranjeras (por ejemplo, de Erasmus durante la carrera); por otra, la experiencia docente o la valoración de publicaciones en revistas científicas (valoradas según el cuartil JCR o SPI) es completamente absurdo para la obtención de dicha beca. Estas ayudas sirven para formar al futuro PDI, ¡valorar la docencia es como pedir a un pollito de gallina poner un huevo! (Del Corral, 2023). En cuanto a las publicaciones, ¿tiene sentido pedir las antes de la beca? En palabras del profesor Julio del Corral (2023), un investigador que tenga artículos o capítulos de libro antes de pedir un contrato FPU puede ser (a) porque no ha participado en ellos como autor principal; (b) porque ya tenga la tesis a medio hacer. En este caso vemos como la evaluación de las becas FPU está envuelto en un sistema ultra-competitivo lleno de sin-sentidos. Son varios los docentes universitarios que reclaman cambios en la forma de valoración, cómo cambios en la remuneración: no pueden convertirse en meras rentas de subsistencia. Dicho esto, se puede apreciar fácilmente lo extraño que puede resultar el sistema de evaluación, en este caso para una beca o ayuda del sistema universitario.

A todo este sistema debemos ahora añadirle lo que supone la llegada de la IA (y sus múltiples facetas) al ámbito educativo. Con razón se hace más aún necesario reconsiderar el modelo de evaluación porque parece anclado en una rígida dinámica que no corresponde con los tiempos que corren. A continuación, se tratará de exponer ciertas consideraciones para reinventar el modelo de evaluación dentro del nuevo marco de la presencia de IA dentro y fuera del aula.

4. HACIA UN NUEVO MODELO DE EVALUACIÓN DEL APRENDIZAJE: PROPUESTAS DE CAMBIO DOCENTE

Han pasado más de 20 años desde la adaptación de las universidades al EEES en el llamado plan o proceso Bolonia. Actualmente la evaluación tipo de una asignatura se estructura en la valoración de un examen final más la valoración de una parte de trabajo práctico (o evaluación continua). La llegada de la IA puede suponer un riesgo a la hora de evaluar esta última parte, sobre todo cuando se trata de redacciones, ensayos, informes o trabajos mandados a casa. Además, la dinámica existente en las formas de evaluación de las asignaturas puede resultar repetitiva y poco creativa, con el riesgo de no reflejar realmente las capacidades del estudiantado. Por ello se hace necesario reconsiderar la forma de evaluación mirando hacia un nuevo modelo que, entre otras, tenga en cuenta en mi opinión de estudiante las siguientes propuestas para conseguir aumentar la motivación fuera y dentro del aula:

- *Convertir el riesgo de la IA en oportunidad.* Es necesario adaptarse a la llegada de la IA al aula y explotarla para sacar su máximo provecho. En este contexto toma sentido una nueva dinámica de clase conocida como la clase invertida (*flipped-classroom*), dónde se aprende haciendo y no memorizando. Nuestros mayores nos hablan de la época en la que no existían el internet ni los correos electrónicos y oponerse a ello parecía ridículo. Hoy son dos herramientas imprescindibles en nuestro día a día. Es necesario “exprimir” la IA para completar el conocimiento a través de varias maneras como pidiendo que desarrolle ejemplos prácticos o que clarifique un concepto que no ha quedado claro. El profesor podrá analizar la exactitud y coherencia de la respuesta y como tarea el estudiante podrá evaluar la respuesta, completarla y sugerir mejoras. Un experimento de uso de aula invertida en estudiantes de ingeniería de la Universidad de Concepción (Chile) mostró una mejora en los resultados académicos y una mayor satisfacción por parte de los estudiantes (véase Maluenda et al, 2021).
- *Cambiar el barómetro de evaluación y pasar del conocimiento teórico al aplicado.* Como forma de combatir la IA se hace todavía más evidente valorar otro tipo de cualidades como la capacidad de análisis, reflexión, contraste de ideas, pensamiento crítico, profundidad del pensamiento o evaluación del contexto. La IA podrá contener mucha inteligencia “artificial” pero carecer de inteligencia “emocional”. La evaluación podría cambiar de un sistema menos

memorístico hacia uno más aplicado. Es necesario basar el sistema de evaluación de ensayos, redacciones o intervenciones siguiendo todos los calificativos anteriores (análisis, diagnóstico, evidencias, pensamiento crítico, evaluación...etc.).

- *Renovar la dinámica en el aula.* Con tantas facilidades existentes por adquirir conocimiento (como son los vídeos, libros, artículos online, IA...etc.), el profesor pasa a ser un actor de soporte y consulta, dando mayor énfasis a las tutorías. Siempre que sea posible el aula debe pasar a ser un espacio de debate y reflexión sobre la materia aprendida y las tutorías individuales y en grupo también. Se deberán fomentar las visitas prácticas hacia lugares relacionados con el tema de estudio (por ejemplo, trabajos de campo, excursiones, visitas a fábricas, empresas, organismos internacionales, organizaciones sin ánimo de lucro...etc.).
- *Potenciar las actividades orales dentro del aula.* Con la IA se hace necesario volver a potenciar las cualidades humanas, y una de ellas es la capacidad de interrelacionarse. Por ello veo conveniente fomentar el debate, la charla o el coloquio dentro del aula. Esto permitiría desarrollar en el estudiantado un cúmulo de habilidades que posteriormente le serán requeridas en el mundo profesional. El aula debe ser un espacio dónde se intercambien opiniones, donde se pregunte y debata, pero también se escuche. Muchos estudiantes no acuden a clase porque según dicen prefieren “leerse las diapositivas en sus casas”. Se hace necesario y urgente volver a motivar al estudiantado y no hay manera más beneficiosa que entregar los materiales anticipadamente e incentivar al estudiantado a través de las diversas dinámicas de participación. Como anécdota personal, en todo el tiempo que llevo estudiando, sólo una profesora incorporó el debate académico y nos enseñó esta técnica cómo actividad práctica en su asignatura. Aunque al principio no provocó mucha simpatía la actividad, después de realizarla todos estuvimos de acuerdo en que esa fue una de las mejores clases (y asignaturas) del Máster.
- *Volver al examen oral.* El examen oral vuelve a cobrar sentido como un posible sistema de evaluación, si bien presenta las dificultades propias derivadas del número de alumnos por grupo. En Italia, por ejemplo, es la forma habitual de examinarse. En este contexto, la evaluación de aspectos teóricos podría enfocarse a través de un tipo test (que además lo corrige una máquina) y del contenido aplicado a través de una redacción o un examen oral. La forma de expresarse, escribir, desarrollar y vincular conceptos son elementos personalísimos de cada persona en los que la IA no puede interferir. Es necesario potenciar la personalidad y originalidad en la enseñanza, los trabajos y la evaluación. Precisamente la repetición memorística y el “corta y pega” lo puede realizar la IA de forma mucho más eficiente y rápida y es aquí donde radica el principal problema.
- *Valorar las preguntas en sí mismas.* En un viaje de Máster en Desarrollo Económico y Políticas Públicas (UAM) los profesores nos pidieron pensar y anotar las preguntas que nos surgían durante las visitas a los lugares a un país en desarrollo. Este tipo de dinámica es muy útil a la hora de fomentar el pensamiento crítico y permite vincular conceptos teóricos aprendidos en el aula con conceptos prácticos del entorno socio-cultural o económico del lugar que se visita. Sobre la importancia de las preguntas y la IA hace hincapié Pedreño (2023) sosteniendo éstas pueden ser la evaluación en sí: permiten realizar una “prueba de carga” a una de las herramientas de conocimiento más potentes de las que dispone la humanidad. Además, plantea que no tiene sentido renunciar a través de preguntas inteligentes el poder de búsqueda y síntesis que tiene dicho instrumento. Por lo tanto, un elemento clave es hacer una pregunta inteligente, correcta, concisa y bien estructurada. Precisamente esto es lo que ha caracterizado a lo largo de la historia de la humanidad al “discípulo” o “aprendiz” brillante.
- *Volver a la lectura.* El uso de la “pregunta” cómo posible método de evaluación encuentra mayor significado cuando se complementa con lecturas. Los beneficios de este hábito son inmensurables porque además de instruir, mejora la capacidad comunicativa, mejora la expresión, desarrolla la reflexión, estimula el cerebro, Por lo tanto, volver a las tareas de “lectura” (acompañada de una redacción o discusión oral que requiera cierta aportación crítica) puede mejorar la didáctica de una asignatura a la vez que permitir al docente evaluar competencias del alumnado.

- *Respetar la libertad de cátedra.* Cada docente debe transmitir el contenido de su asignatura como considere oportuno sin presiones algunas por parte de la estructura universitaria, cosa que no siempre ocurre. Desde una evaluación continua al sistema de evaluación tradicional existen muchos puntos intermedios, pasando por la realización de distintas actividades que permitan enseñar y aprender de diversas formas en una nueva dinámica de aula (por ejemplo, en las citadas visitas exteriores, actividades prácticas, excursiones...etc.). La libertad de cátedra debe permitir al docente adaptar para cada asignatura el método de evaluación que más conveniente estimara. Por ejemplo, habrá asignaturas que precisen de examen final, mientras que en otras la aptitud y los conocimientos del estudiante puede evaluarse de otras formas, sin necesidad de encorsetamientos. Precisamente los estudiantes recordamos a los profesores por su originalidad (en la forma de impartir la clase, enfocar la asignatura, transmitir conocimiento, mandar tareas o evaluar conocimiento). Esta “libertad” que debe potenciar la “originalidad” debería ser un elemento distintivo del sistema universitario.

Seguramente existan muchos más puntos y propuestas que permitan mejorar el sistema de evaluación docente en la enseñanza universitaria. Lo cierto es que ni el sistema de evaluación continua, ni el sistema de evaluación tradicional son perfectos y conviene aplicar un punto intermedio con flexibilidad según las necesidades de cada asignatura y preferencias del profesorado. La llegada de la IA refuerza la necesidad de cambiar, no tanto el sistema de evaluación, sino las modalidades de tareas y entregas solicitadas por el profesor. En este punto es clave el número de alumnos por aula (y por profesor. Por eso, es necesario reenfocar el contenido de la parte práctica hacia un conocimiento más aplicado, renovando tanto el sistema de evaluación como la propia dinámica del aula en función de diversos parámetros -ámbitos de conocimiento, grado de experimentalidad, número de estudiantes por grupo, etc.)

La solución no parece ser volver al modelo de evaluación tradicional, en todo caso mantener el actual, pero mejorándolo. El tensionamiento entre los modos de evaluación formativa vs. sumativa encuentra reflejo en el tensionamiento evaluación continua vs. tradicional. La experiencia hasta ahora indica que el rendimiento de los estudiantes en un entorno de evaluación formativa (en la que se valora el proceso de enseñanza-aprendizaje) es mayor que aquel de los estudiantes que optan por otras vías de evaluación (como por ejemplo la elección de la evaluación no-continua, yendo a examen final). En este tipo de metodología, el resultado de la evolución en el estudiante no es determinista y comparable, pero tiene un valor intrínseco y descriptivo acerca de los logros alcanzados a través de unos objetivos flexibles y adaptados a cada individuo (Melado-Moreno et al, 2021).

Sin embargo, todo dependerá del tipo de asignatura que se trate. En aquellas más prácticas esta noción se refuerza. Un experimento de la Universidad de Valladolid para el grado de Educación Infantil sugiere que la vía de aprendizaje formativa mostró un mejor rendimiento académico (Molina-Soria et al., 2020). Esto cobra aún más sentido dadas las recientes vivencias en de los estudiantes y profesores durante el confinamiento.

Otro tema muy importante es la evaluación de las asignaturas y docencia por parte del estudiantado. En este aspecto los estudiantes tenemos una función muy importante a la hora de lograr el cambio a través de nuestra opinión constructiva en las encuestas docentes que se realizan (en particular en las preguntas abiertas). De hecho, todos conocemos la insistencia con la que nuestros profesores nos recuerdan rellenar las encuestas. Pocos son los estudiantes que las completan debidamente. El paso al formato online de las encuestas (en este caso, de nuevo la tecnología) no parece haber ayudado en cuanto a interés en su cumplimentación.

Lo cierto es que mientras sigan existiendo tareas para casa de escaso valor añadido, la IA se convertirá seguramente en el principal suministrador de respuestas. El máximo provecho de la IA no se saca copiando y pegando su contenido, sino poniéndolo a prueba. Me remito a las palabras de la profesora Carmen Pérez-Esparrells, Directora de Calidad del blog Universidad, que contestando a una entrada comentaba como no hay vuelta atrás por mucho que se empeñen las Universidades o centros en

prohibir el chat. Las líneas de actuación parecen claras: sacarle el máximo provecho al chat, a la IA y al resto de las tecnologías disruptivas (metaverso en la docencia universitaria en la que por cuestiones de tiempo no vamos a entrar) cambiando la forma en la que se evalúa con inteligencia. Pero, sobre todo, respetar la libertad de cátedra y renovar la dinámica en el aula para lograr que la Universidad vuelva a motivar al “estudiante medio” o a la mayoría del estudiantado. Aunque suene pretencioso es uno de los mayores retos del futuro cercano del sistema educativo en su conjunto.

5. CONCLUSIÓN

Da miedo especular como estudiante de último curso universitario que soy actualmente ¿qué vendrá luego? Pensar que muchos de nuestros profesores habían nacido en una época sin internet ni teléfonos móviles, cuando muchos de los estudiantes de ahora recurren a sistemas de IA para realizar entregas son solo muestras de lo rápido que parece avanzar la tecnología en la sociedad del conocimiento. El mundo en el que vivimos ahora está asistido en todos los aspectos por la tecnología y nuestro futuro está orientado a vivir en una convivencia entre humanos y no humanos (Martínez, 2023).

No queda otra que aceptar dichos avances tecnológicos, aunque el futuro de vértigo, los retos éticos y regulatorios podrían solucionarse a través de los pactos de integridad académica entre estudiantes, profesores e investigadores. En el ámbito educativo, al igual que ocurre en el ámbito de la salud o en otros muchos ámbitos, el éxito estará determinado por la capacidad de integración de dichos progresos en IA en el sistema educativo para sacar el máximo provecho de los mismos y lograr el máximo aprendizaje y avance en la generación del conocimiento. Pedreño (2023) apunta a que se necesita una mente curiosa para sacar el máximo provecho a dicha tecnología logrando empoderar a los estudiantes de manera personalizada y adaptativa. Estas tecnologías serán imprescindibles en el futuro desarrollo profesional de los estudiantes y en su perfecta inserción en el mercado laboral del futuro.

Creo que la llegada de la IA al aula es toda una oportunidad para darle una vuelta de tuerca a la evaluación del sistema educativo en general y del universitario en particular. Un sistema en el que los estudiantes recurren a la vía rápida para quitarse los trabajos de encima, en el que se pasa lista para lograr asistencia o en el que se organizan interesantísimas conferencias presenciales a las que nadie asiste es un sistema colapsado. Creo que nuestro sistema educativo antes y después de la irrupción de la IA está en crisis (si esa es la palabra). Un sistema educativo que no motiva a la gran mayoría de estudiantes es un sistema fallido.

La modernización del sistema de aprendizaje debe liderarse desde la universidad, puesto que es aquí donde nacen los futuros docentes e investigadores. Las complicaciones legales existentes no deben suponer una limitación para no experimentar con nuevas formas de enseñar y aprender, más ahora cuando se presenta la oportunidad de la IA al alcance de cualquiera (a precio cero si no fuera por la captación de datos de los usuarios del chat GPT, de sus escuelas, universidades, empresas, etc.).

Son muchos los expertos que opinan que prohibir el chat GPT no es la solución. Hacerlo podría generar el efecto contrario. Hemos visto que en Italia se prohibió durante un mes y no se consiguió nada, ¿acaso no está prohibido también consumir drogas, descargarse series pirata o conducir sin cinturón? Quién inventa la ley, inventa la trampa. La tecnología avanza muy rápido y aunque se intente prohibir, los más ágiles encontrarán métodos para acceder a ello. Parece que la creatividad, el pensamiento crítico y la resolución de problemas parecen las únicas vías por las que se pueda apostar, además de los ya mencionados pactos de integridad académica. Una entrada de una profesora de la Universidad Complutense reforzaba la necesidad de complementar a la IA ya que sólo así “podremos desarrollar un aprendizaje continuo y con perspectiva analítica, inspirar, ilusionar y aportar la calidad y cercanía humana que es exigible en el *learning* o *mentoring* del verdadero maestro” (Muñoz, 2023).

La universidad debe ser un espacio que inculque a los estudiantes cualidades y valores que le serán válidos en todos los aspectos de la vida profesional y personal. En este contexto tiene sentido reformar la dinámica en el aula y el sistema de evaluación docente. La verdadera evaluación continua

debe ir más allá de la estructura típica de informe-presentación. Tiene que fomentar el pensamiento crítico y la reflexión. Incluso podría darse fuera del aula puesto que es ahí donde también se aprenden, aplican y evalúan contenidos prácticos en la vida real. Puede confirmarse con seguridad que la gran mayoría de egresados que ya están trabajando corroboran este último aspecto. Por ello creo que los viajes de estudios son la mejor forma de aplicar el conocimiento aprendido y además dan cabida a una forma de evaluación a través de las preguntas o redacciones reflexivas sobre el contexto. Lo práctico permite interiorizar y completar el aprendizaje teórico del aula en una forma muy efectiva. Evidentemente, cada asignatura tendrá una mayor o menor facilidad en acoger dicha dinámica y el número de alumnos en el aula condiciona las posibilidades de éxito en las nuevas formas de evaluación que están por venir.

La evaluación se encuentra actualmente en el difícil reto (o encrucijada) de compaginar la valoración de un contenido teórico, así como la valoración de otras destrezas del estudiante, que como se ha dicho son personales y humanas. Además del conocimiento, el mundo laboral reclama muchas otras cualidades que también deben ser aprendidas e interiorizadas en la universidad (humanismo, integridad, bondad, ética y creatividad).

La universidad no puede limitarse a ser una simple “fábrica de egresados” (en palabras de un profesor de sociología mío). La universidad debe pasar por la persona, no al revés. El sistema educativo y universitario debe inculcarnos principios y valores para hacernos mejores personas en todos los aspectos, desde el ámbito académico (personas más cultas, informadas y capaces) hasta el ámbito personal (llenas de principios y valores humanos). La universidad debe ser toda una experiencia que permita crecer personal y profesionalmente a través de las múltiples oportunidades y servicios que ofrece. Sino ¿para qué sirve su paso a través de ella? Seguramente las cosas en el mundo “adulto” (pienso en el mundo profesional y político) irían mucho mejor si existiera desde las instituciones educativas un esfuerzo por potenciar el debate y discusión, la reflexión, el pensamiento crítico, la aplicación de contenido, la capacidad creativa, la inteligencia emocional, la resolución de problemas o muchas otras cualidades complementarias al lado académico. ¡Qué difícil resulta medir estos parámetros! Antes o después de la IA. Aquí citaré una frase que nos repite el catedrático de la UAM José María Mella en sus clases: “no es tanto la aptitud, si no la actitud”.

Por último, me parece curioso finalizar analizando las raíces de las palabras academia y evaluar. La palabra academia viene del griego *akademeia* y es nada menos que el lugar dedicado a un héroe griego en el que Platón enseñaba sus doctrinas. Este lugar era un jardín. Por otra parte, la palabra evaluar proviene del latín *ex valere* que significa “ser fuerte hacia afuera”. Sin embargo, no define el qué ni cómo se mide dicho valor.

6. BIBLIOGRAFÍA

- CIFUENTES-FAURA, J. (2020). Consecuencias en los niños del cierre de escuelas por Covid-19: El papel del gobierno, profesores y padres. *Revista Internacional de Educación para la Justicia Social*, 9, 3e.
- COMAS, R., SUREDA, J., CASERO, A., & MOREY, M. (2011). La integridad académica entre el alumnado universitario español. *Estudios pedagógicos (Valdivia)*, 37(1), 207- 225.
- DEL CORRAL, Julio. (2023). Las nuevas FPU: alguna luz y muchas sombras. *El Blog de Studia XXI. Universidad.*
- GONZALO, Marilín. (2023). Como detectar imágenes, vídeos o audios generados con IA. *Newtral Periódico.*
- LLORENS, F. & ChatGPT. (2022). Cavilaciones invernales. *El Blog de Studia XXI. Universidad.*
- MALUENDA, J., VARAS, M., & CHACANO, D. (2021). Efectos del aula invertida y la evaluación auténtica en el aprendizaje de la matemática universitaria en estudiantes de primer año de ingeniería. *Educación*. Vol. 30. No. 58
- MARTÍNEZ, Dídac. (2023). El desconcierto creciente de las cosas inteligentes. *El Blog de Studia XXI. Universidad.*

- MELLADO-MORENO, P.C., SÁNCHEZ-ANTOLÍN, P., & BLANCO-GARCÍA, M. (2021). Tendencias de la evaluación formativa y sumativa del alumnado en Web of Sciences. *Alteridad*, 16(2), 170-183.
- MOLINA-SORIA, M., PASCUAL-ARIAS, C., & LÓPEZ PASTOR, V. (2020). El rendimiento académico y la evaluación formativa y compartida en formación del profesorado. *Alteridad*, 15(2), 204-215.
- MORENO, L., PEDREÑO, A. (2020). Europa frente a EE.UU. y China. Prevenir el declive en la era de la inteligencia artificial. Libro. Amazon.
- MUÑOZ, C. (2023). ¿Chat GPT en la Universidad? ¿Complementar el aprendizaje y cambiar el modelo educativo? *Diario La Ley, No. 10283, Sección Tribuna*.
- PEDREÑO, Andrés. (2023). ChatGPT y las universidades. El Blog de Studia XXI. Universidad.
- PEDREÑO, Andrés. (2023). ChatGPT y universidad: obteniendo las respuestas ¿conocemos las preguntas? El Blog de Studia XXI. Universidad.
- VALLESPÍN, David. (2023). La conveniente racionalización de la evaluación continua universitaria. El Blog de Studia XXI. Universidad.