

TRABAJO FIN DE MÁSTER

CURSO ACADÉMICO 2021- 2022



PROBLEMAS EN TORNO A LA ENSEÑANZA DE LAS HUMANIDADES EN LA ERA DE LA INTELIGENCIA ARTIFICIAL

Máster Universitario en Formación del profesorado de
Educación Secundaria Obligatoria y Bachillerato

Aurora Irene García Carreras

Director: Luis Llera Cantero

Especialidad en Filosofía

Madrid, 16 de junio de 20

AGRADECIMIENTOS

Para la investigadora, la elaboración de este texto ha supuesto un reto considerable. La relevancia y amplitud de los temas escogidos, a saber, el ámbito de la inteligencia artificial, la reflexión educativa, así como el problema de las humanidades ya exigían como previo requisito la imposición de una notable disciplina. Sin una investigación extensa y variada este proyecto ni siquiera habría podido materializarse. Sin embargo, si el reto propuesto ha sido concluido de una manera satisfactoria, eso es algo que ha de ser juzgado por el lector.

Quería agradecer a mi familia y amigos cercanos porque sin la existencia de estos lazos personales ni siquiera podría pensar. A los compañeros del máster por contribuir con su implicación a generar un clima de compañerismo y de trabajo. A Bea y Andrés, en varias ocasiones compañeros infatigables de biblioteca.

En especial quería agradecer al tutor del trabajo Luis Llera por haber manifestado desde las clases del máster la enorme relevancia educativa del ámbito de la inteligencia artificial. Además, sin sus guías y recomendaciones bibliográficas hubiera sido prácticamente imposible orientarse.

También quería agradecer las facilidades aportadas por la universidad y su fondo bibliográfico. Y en todo caso, a la profesora Tamara Díaz Fouz cuya sesión sobre inteligencia artificial y educación ayudó al enfoque del trabajo una vez la investigación se encontraba en máximo proceso acumulativo de fuentes.

En último lugar, quería agradecer a todos aquellos investigadores que incansablemente se encuentran trabajando sobre pedagogía, inteligencia artificial y reflexión humanística. Sin su trabajo esta investigación no habría podido tener lugar.

ÍNDICE DE CONTENIDOS

1. Resumen y palabras clave	3
2. Introducción	4
3. Justificación	8
4. Comentario a la legislación	12
4.1. Un comentario al proyecto de los <i>neuroderechos</i> desde un prisma educativo	14
4.1.1 Algoritmo humano demasiado humano	17
5. Estado de la cuestión.	19
5.1. Qué es la inteligencia artificial y qué problemas nos presenta.	19
5.2. Aplicación efectiva de la inteligencia artificial en el ámbito educativo.	27
6. El papel de las humanidades en la era de la inteligencia artificial	34
6.1. Sobre máquinas de enseñanza-aprendizaje	36
6.1.1. Sobre máquinas de enseñanza	36
6.1.2. Sobre máquinas de aprendizaje	39
6.2. Una pluralidad técnica ante una pluralidad de fines	42
6.3. Aspectos de resistencia computacional de la enseñanza de las humanidades	44
6.4. ¿Humanidades?	55
Conclusiones	59
Bibliografía	61
ANEXO 1. GLOSARIO DE TÉRMINOS TÉCNICOS	67

1. Resumen y palabras clave

En el presente escrito se plantea una discusión de tres planos en torno al problema de la inteligencia artificial en el ámbito educativo. Para ello, nos preguntamos por las discusiones aparejadas a la disciplina de la inteligencia artificial, así como por sus aplicaciones al ámbito educativo. La reflexión nos obliga a transitar por concepciones propias de estos campos que, en ciertas ocasiones, tienen proximidad con la reflexión transhumanista. Realizamos, de esta manera, un rastreo histórico de los enfoques pedagógicos que irían entrelazados con la inteligencia artificial. No obstante, este movimiento no nos conduce a acuerdo pues en el terreno estudiado coexistirían una pluralidad de concepciones de la enseñanza. Por último, realizamos un abordaje del específico tema de nuestra pregunta: el problema de las humanidades. Sin embargo, la investigación realizada nos devuelve un cuestionamiento de la pregunta guía de la investigación. ¿Es la era de la inteligencia artificial una oportunidad única para constituir una tercera cultura educativa?

Palabras clave: Inteligencia artificial, humanidades, educación, tercera cultura, dataísmo, tecnología.

Abstract:

This paper studies a three-pronged discussion of the problem of artificial intelligence in education. To do this, we wonder about the debates associated with the discipline of artificial intelligence, as well as its applications to the educational field. Reflection forces us to move through the characteristic concepts of these fields, which, on certain occasions, are close to transhumanism. Thus, we carry out a historical tracing of the pedagogical approaches that would be intertwined with artificial intelligence. However, this movement does not lead us to agreement because in the studied field there are a coexisting plurality of teaching conceptions. Finally, we address the main issue developed in this paper: the problem of the humanities. Nonetheless, the research makes us question the guiding formulation of the research. Is the artificial intelligence era a unique opportunity to establish a third educational culture?

Keywords: Artificial intelligence, humanities, education, third culture, dataism, technology.

2. Introducción

En este trabajo se va a plantear una discusión en torno al estado de la enseñanza de las humanidades en el panorama de la floreciente disciplina de la inteligencia artificial. Sin embargo, esta cuestión será abordada en último lugar puesto que constituye un requisito previo la anterior reflexión acerca de la relación existente entre esta tecnología y la pedagogía.

De esta manera, la cuestión protagonista se encuentra atravesada por otra de encuadre más amplio que es aquella incursión de la inteligencia artificial en el ámbito educativo. Pero también, esta vía de análisis tan solo constituye una pequeña senda dentro de ese inmenso paisaje que representaría la problemática que describe la inteligencia artificial.

Por ello, no podemos limitar nuestro “movimiento de cámara” a un mero plano detalle. Ha de tomarse en cuenta que la inteligencia artificial augura un cambio social y económico a nivel global que ha de ser transitado de manera ineludible. Nuestro trabajo, de este modo, orbitará en un movimiento de tres planos. Esto es debido a que el plano detalle estaría incompleto si desconocemos tanto el plano medio como el plano general.

De manera preliminar, hemos de señalar que la inteligencia artificial no ha de ser comprendida como una aspiración lejana puesto que ya constituye una herramienta que se utiliza de manera cotidiana. Los fieles móviles y ordenadores de nuestros estudiantes y los nuestros ya cuentan con esta tecnología. La inteligencia artificial nos recomienda series y productos y gestiona nuestras redes sociales. No presenta una cara desconocida. Como reseña el famoso teórico transhumanista Nick Bostrom, en su éxito de ventas *Superinteligencia*, el mayor sistema de inteligencia artificial construido hasta la fecha sería mismísimo motor de búsqueda de *Google*.

En un plano general, nos enfrentamos, de esta manera, a un contexto de una progresiva digitalización y virtualización de la vida. Las clases trascienden las barreras espaciales y los plazos temporales que habrían constreñido una enseñanza exclusivamente analógica. Nuestra propia identidad se digitaliza y parece convertirse en información. El contenido de nuestra existencia, en suma, parece ser materia bruta que estaría disponible para su posterior tratamiento como datos.

Entonces, la *inteligencia artificial*, conocida también para algunos como ciencia *de datos*, se presenta como herramienta que configuraría el propio entendimiento de

nuestra vida. El trato con la misma, convertida en datos, estaría disponible a ella. Lo existente sería propiamente lo medible -o “datable”- y lo medible hablaría por sí solo. Algunos autores han denominado a esta coyuntura, en cierta medida ideológica, *dataísmo*.

No obstante, el panorama no es enteramente desolador. Aunque las discusiones en algunas ocasiones tornen un cariz futurista y se pregunten si acaso seremos reemplazados progresivamente por máquinas y perderemos nuestro trabajo e incluso la vida, o si llegará acaso aquel *giro traicionero* mediante el cual la *inteligencia artificial* se vuelva contra nosotros con el propósito de lograr sus propios objetivos a nuestra costa, lo cierto es que gran parte de las implicaciones de esta disciplina ya se están dejando ver. Es constatable que existe un cambio en la manera de trabajar y de comunicarse, así como, incluso, de dirigirse al conocimiento.

Advertimos así que, en un plano medio, en el ámbito educativo, el puro contenido parece perder su valía. Como se verá en el desarrollo posterior, aquella incidencia en el estudio memorístico podrá perder relevancia frente la riqueza de la información disponible en la red. Si bien, esta cuestión nos conduce a replantearnos si la memoria puede entenderse acaso como mero procesamiento de la información en una analogía perfecta con la máquina.

Asimismo, las propuestas de modificación del currículo de organismos como la UNESCO parecen abogar por una concepción de la enseñanza más centrada en la dotación de herramientas o competencias a los alumnos a través de las cuales éstos sepan aprovechar los medios disponibles y no quedar subsumidos por los mismos. Como se sugerirá, las *cámaras de eco* de los algoritmos parecen tener su correlato en las concepciones radicales o particularistas de la existencia mientras que las *burbujas de filtro* en cierta manera podrían estar reforzando una visión sesgada o maniquea de la realidad.

Lo “humano”, de esta manera, no es raramente presentado como alteridad salvadora ante el sinsentido alcanzado por el progreso técnico. Ya en el mercado laboral presente, encontramos un elemento de distancia frente a la cada vez más universal formación técnica y especialista.

Esta se encontraría, por ejemplo, en el progresivo superior reconocimiento de las habilidades no correspondientes con un perfil técnico sino más propiamente humanas o

sociales. Nos referimos a aquellas conocidas como las *soft skills* entre las cuales se encuentran la capacidad de trabajar en equipo o de gestionar el tiempo.

De igual manera, en el panorama educativo hacemos frente ante una creciente propuesta de modificación del currículo hacia habilidades igualmente comunicativas, pero además hacia un enfoque interdisciplinar. Se advierte un panorama de pérdida de trabajos, o *jobsolescence*, a causa de esta eclosión tecnológica y, en consecuencia, se plantean adaptaciones académicas y laborales que permitan a los individuos del mañana tener un perfil adaptativo. Esta posibilidad de adaptación parece plantearse, en cierta medida, en un panorama de posible sustitución de los seres humanos por máquinas.

De acuerdo con lo expuesto, estaríamos atendiendo a cómo la enseñanza y sus exigencias parecen cambiar en sus contenidos en la medida en la que el modelo laboral y productivo se ve alterado. Más aún en este contexto de auge de la ciencia de datos, no son extrañas las declaraciones de organizaciones y expertos del entorno de Silicon Valley que proponen una valoración del papel de las humanidades en los tiempos presentes. Organismos como la UNESCO ya han resaltado la extrema importancia de que los alumnos reciban una formación “humanista”.

De esta manera, podríamos apreciar una paradoja. Se plantea que el mercado de trabajo está cambiando y que cada vez más actividades podrán ser automatizadas. Por otro lado, aquellas actividades o disciplinas que se resistan a la automatización cobrarán una mayor valía siendo estas competencias eminentemente atribuidas al ámbito de las humanidades, así como de la programación y de la ingeniería. La paradoja, en definitiva, es la siguiente: a causa de un razonamiento utilitarista, las humanidades podrían ser revalorizadas.

Recordamos aquí las enseñanzas de Unamuno en la novela *Amor y pedagogía* (1902), ¿qué es esto de buscar en la utilidad de la ciencia una respuesta a las preguntas más humanas? Más aún, ¿qué es esto de dotar de valía a las preguntas más humanas por su utilidad?

De este modo, la posesión de atributos específicos no posibles como propiedades en las inteligencias artificiales podrían dar una ventaja competitiva a los individuos. De igual manera, competencias como la interdisciplinariedad también estarían en auge. Pareciera entonces que las fluctuaciones sociales y del mercado habrían provocado de nuevo, un retorno al ideal del renacimiento.

Reiteramos, paradójicamente un enfoque holístico del ser humano ofrecería un perfil de mayor adaptabilidad ante la sociedad futura y, por tanto, en una fina ironía de la historia sería aquel modelo defendido por los razonamientos utilitaristas. Prueba de este razonamiento son libros del corte de *Menos tech y más Platón* con su justificación del valor empresarial de las humanidades.

Sin embargo, como precisión metodológica, hemos de señalar que no compartiremos un enfoque que divida de manera radical aquellas disciplinas entendidas como “humanísticas” frente a aquellas que serían comprendidas como “ciencias puras”. Por el contrario, para nosotros lo propiamente humano es aquello que actualmente no puede ser desempeñado por las máquinas, como pudiera ser la formulación de la teoría de la relatividad.

Trataremos así, de eludir aquella división falaz que ya el físico y escritor Charles Percy Snow advertía. En este trabajo no se abogará por ninguna de las dos culturas sino, en todo caso, por una visión holística del ser humano y un enfoque transversal e interdisciplinar del conocimiento.

De esta manera, compartiremos objetivos con este razonamiento utilitarista, pero no razones. Quizás algunos, en una ironía propia del tema tratado, no sean por tanto capaces de distinguir nuestros argumentos a favor de las humanidades de la conducta simulada de defensa del valor empresarial de las humanidades. Pero, eso, es otra historia.

En consecuencia, dentro del ámbito de lo posible en este escrito, tan solo se atenderá al estudio de las posibilidades y concreciones efectivas de la inteligencia artificial en el ámbito educativo, y todo ello constreñido bajo nuestro prisma de análisis que es el de vislumbrar el papel de las humanidades en esta tesitura de los tiempos.

3. Justificación

Dada la amplitud y actualidad del tema confesamos que nos ha costado llegar a la acotación de éste. Son demasiados los frentes abiertos por la pujanza de esta novedosa técnica. En este sentido, nuestro tema de investigación se erige como una modesta propuesta de abordaje dentro de todo este caos.

Este limitado enfoque, no obstante, a nuestro juicio, pugna por abordar un aspecto fundamental para nuestra formación como ciudadanos y, en definitiva, como seres humanos que viven en una sociedad: el tratamiento de la enseñanza de aquellas asignaturas definidas como “humanidades” en esta era de la inteligencia artificial.

Con ello, no podemos evitar hacernos las siguientes preguntas. ¿Qué será de las futuras generaciones y de sus capacidades críticas ante el arrollador transcurso de la inteligencia artificial? ¿Seremos capaces de resistir a una ola de entusiasmo y de *novolatría* por la inclusión de nuevas técnicas? ¿Desplazará la tecnología el pensamiento humanista tal y como lo conocemos?

No esperamos, ni mucho menos, cerrar la discusión. Al contrario, en todo caso, pretendemos acrecentarla. Si bien, a nuestro juicio, esta empresa no por ello será superflua. Quizás el adecuado planteamiento de estos problemas contribuya de manera modesta a su posterior clarificación.

¿La I.A. llevará a una robotización de la sociedad, como denuncia la *declaración de Barcelona* o, por el contrario, nos liberará de la automatización del hombre? ¿La realización por parte de *robots* de tareas mecánicas nos hará mecánicos o más libres? Ninguna de estas cuestiones transcurrirá de manera automática puesto que el curso de la historia en cierta medida se produce a causa de decisiones políticas como sociedad. ¿Cómo navegaremos esta corriente que nos viene?

En lo pertinente a nuestro campo de estudio, podríamos decir que a nivel educativo tenemos cierto margen de acción. En el mero plano del aula, incorporando estas tecnologías... ¿Qué papel queremos que tengan los alumnos?

Consideramos que nuestro trabajo resulta pertinente debido al reto que plantean las características particulares de disciplinas como Historia, Filosofía, Lengua y las lenguas extranjeras. Si habláramos en una jerga propia de la ciencia de datos, podríamos decir que las disciplinas humanísticas presentan la misma problemática que la de los *datos*

no estructurados y es que, no presentan de manera apriorística una forma o estructura que se preste a una categorización en filas y columnas. La historia no son fechas, la Filosofía no es una lista de categorías y la conversación de las lenguas extranjeras no se presta de manera fácil a este registro automático.

Como se verá, todas estas disciplinas, en su estudio, presentan rasgos que superan con creces las capacidades de la inteligencia artificial. El estudio de estos campos implica la capacidad de transferir conocimientos previos a contextos nuevos. Por ejemplo, en Filosofía podemos relacionar la crítica contra el principio de causalidad de Hume con la diferencia entre correlación y causalidad que se establece en estadística. Asimismo, en Historia tratamos de buscar la comprensión de la sucesión temporal, el porqué de los eventos. De igual manera, el estudio del propio lenguaje también implica una especial atención al factor contextual, a los modos de vida al cual sirve, si queremos ser wittgenstenianos.

Ahora sí, una vez sembrado este panorama de justificación, vamos a ofrecer una lista de motivos para los cuales estimamos que nuestra investigación resulta pertinente.

1) Observando los resultados del informe sobre el proyecto de la asignatura *Pensamiento computacional e inteligencia artificial* del INTEF vemos que este tipo de disciplinas podrían presentar cierto grado de resistencia a la hora de sincronizarse con el cada vez más en auge pensamiento computacional. De manera sumaria, podríamos señalar que este tipo de pensamiento se presta más a una circularidad y presentidad mientras que las disciplinas humanísticas suelen tener un proceder circunscrito a la temporalidad. ¿Son las ciencias más proclives para adaptarse a la metodología computacional? Quizás esto tenga que ver con la diferencia que establecía el neokantiano Rickert: el método histórico es individualizador mientras que las ciencias naturales generalizan.

2) En un panorama en el cual la tecnología parece haber tomado la delantera y los esfuerzos de la sociedad se centran en el progresivo mejoramiento técnico, las humanidades pueden ayudar a formar ciudadanos que emprendan la búsqueda de fines. Las herramientas son medios, pero su proyección no tiene por qué estar establecida de antemano. En este punto podríamos preguntarnos si acaso la inteligencia artificial es verdaderamente inteligente puesto que, como se pregunta Bostrom, quizás la inteligencia pueda estar estrictamente vinculada con la capacidad de conducirse según fines. Es

posible que, como seres inteligentes, precisamente nuestra potestad sería el establecimiento de finalidades a una simulada inteligencia que además es artificial.

3) En un plano de atención a lo puramente efectivo, hemos de pensar en por qué las humanidades actualmente podrían ser adaptativas. Frente al humano, las máquinas tienen la ventaja de no tener que descansar, pueden analizar cantidades ingentes de información a la vez, así como poseen un margen de error ínfimo. Debemos pensar en las implicaciones del hecho de que las humanidades puedan servir como ventaja competitiva frente a las máquinas, así como debemos pensar acerca de si esto efectivamente es así. Podríamos tratar de sospechar y plantearnos si acaso no se estaría utilizando a las humanidades para tratar de dar un rostro amable a la futura era de la inteligencia artificial. Aun así, no podemos dejar de pensar... ¿qué nos ofrecerían las humanidades? ¿son estas características imprescindibles en las asinaturas de humanidades?

4) La inteligencia artificial nos hace reflexionar sobre el fenómeno del aprendizaje. Más aún, nos ofrece un escenario privilegiado en el cual pensar sobre las vinculaciones que pueden existir entre la pedagogía y la tecnología. Al respecto, son interesantes los trabajos del psicólogo cognitivista Manuel Vega en los cuales se plantea una distinción entre el modo de pensamiento computacional y el cotidianamente humano. También merece una especial atención el estudio comparado del cerebro de los infantes y los modelos de inteligencia artificial realizados por el ingeniero Minsky y el psicólogo Papert, los cuales llevaron a una determinada concepción educativa, así como al desarrollo de un lenguaje de programación para niños que se conoce como *Logo*.

5) Este trabajo también se justificaría como esfuerzo de clarificación ante el complejo estado de la literatura actual sobre el tema. Las fuentes consultadas muchas veces eran demasiado futuristas o divulgativas y tendían a augurar un paraíso tecnológico o bien el fin de los tiempos. Del mismo modo, algunas eran excesivamente técnicas y poco accesibles. Por ello, aunque nuestro trabajo no pueda resolver gran cosa estimamos que quizás pueda servir de aliciente para que personas más capacitadas se pregunten e incluso solucionen las problemáticas evocadas.

6) Es necesario plantearse si precisamente este desarrollo tecnológico puede representar un camino para reconciliar la cultura humanística y la cultura científica. En la reflexión sobre aquello que como profesores podemos enseñar que exceda a la ciencia de

datos, quizás encontremos que muchas competencias también son propias de disciplinas científicas. ¿Son la creatividad o la facultad crítica exclusivas de las humanidades?

4. Comentario a la legislación

En esta sección se va a realizar un comentario crítico a la legislación en torno a la inteligencia artificial en el ámbito educativo. No obstante, debido a la novedad del ámbito, este análisis, en ocasiones, va a tener que tomar como objetivos centrales más las ausencias que las presencias de regulaciones. De manera preliminar, ya podemos señalar que apenas encontramos legislación específica y esto es disonante con la relevancia material, así como los desafíos éticos que plantea el tema de la inteligencia artificial en el ámbito educativo.

En un análisis de la actual ley de educación podemos ver que la expresión “inteligencia artificial” aparece un total de cero veces. Por el contrario, la palabra “datos” estaría presente en torno a dieciocho ocasiones mientras que el término “humanidades” tan solo cuatro.

Este hecho no deja de sorprendernos, debido a que desde 2018, con comunicados como “Inteligencia artificial para Europa” por parte de la Comisión europea, se ha incidido de manera notable en el potencial transformador de la herramienta de inteligencia artificial. Si bien, no es hasta finales de este año cuando se planea el lanzamiento de una legislación pionera que regule de manera conjunta las posibilidades de aplicación de la herramienta.

En la LOMLOE, no existe ningún tipo de anticipo con respecto a los desafíos que presentaría la inteligencia artificial. Siendo generosos, podríamos señalar que en esta ley se remarca la importancia de la protección de los datos personales, así como las garantías de los derechos digitales en la disposición final cuarta del texto. En estas últimas, de manera acertada, se menciona la importancia de adquirir competencias digitales, así como del desarrollo de una especial atención hacia situaciones de violencia en la red.

Sin embargo, atendiendo al panorama en disputa, esta novedosa ley educativa no deja de resultarnos arcaica. Los algoritmos de inteligencia artificial tienen una serie de potencialidades ante las cuales nuestros alumnos fácilmente quedarían indefensos.

Las violaciones de privacidad de la ley penan un tipo de hecho que se limita a las acciones realizadas por un agente particular como pudiera ser un alumno o una entidad como el centro educativo. Pero ¿qué tipo de agente particular es un algoritmo? Ante una filtración indebida de los datos de nuestros alumnos, que, por otro lado, habrán de ser recolectados de manera masiva para que el dispositivo de inteligencia artificial funcione,

¿quién cobrará responsabilidad penal? Asimismo, ¿qué ocurrirá si un algoritmo utilizado resulta discriminatorio hacia los alumnos de una etnia determinada? ¿A quién atribuimos responsabilidad si los datos de nuestros alumnos son masivamente recolectados por parte de una empresa para mejorar sus servicios?

Al respecto, encontramos un, ya mencionado, proyecto de legislación específica en la iniciativa de la Comisión europea. Este proyecto de marco legal en torno a la inteligencia artificial para los estados miembros tiene como objetivo el respeto a los derechos fundamentales de los ciudadanos, así como un aumento en la confianza en estas tecnologías (Comisión Europea, 2021).

Con este tipo de legislaciones se observa la estrecha vinculación que parece existir entre el sistema político del territorio y el marco de aplicación posible de esta novedosa herramienta. Tanto es así que, atendemos a que la posibilidad de expansión de la inteligencia artificial en Europa, a través de esta legislación, estaría correlacionada con el hecho de que sus aplicaciones sean cabales en el entorno de una democracia liberal.

Por ejemplo, vemos en esta iniciativa de la Comisión Europea, una propuesta de sanciones millonarias ante el caso de que ciertas empresas en el uso de estas tecnologías infrinjan los derechos fundamentales. De igual manera, atendemos en esta propuesta a una prohibición generalizada del uso de los dispositivos de identificación biométrica remota, tan frecuentes en China¹, así como una prohibición de la aplicación de la inteligencia artificial para propósitos de elaborar sistemas para puntuación social por parte del gobierno.

¹ Los dispositivos de identificación biométrica son un tipo de *softwares* asociados a cámaras que se especializan en el reconocimiento de seres humanos específicos ya sea por sus rasgos faciales particulares, su manera de andar, por sus huellas dactilares o por su voz. Está planeado que este tipo de dispositivos sean prohibidos, de manera generalizada en el espacio público, en el entorno de la Unión Europea debido a que violarían el derecho a la privacidad al posibilitar al gobierno un seguimiento casi absoluto de los movimientos de cada gobernado. No obstante, se plantea que su uso sea aceptado para casos policiales específicos como para el supuesto de que hubiera que encontrar al sospechoso de un asesinato. Sin embargo, esta tecnología también entraña riesgos pues, aunque su efectividad de reconocimiento pueda ser de un 99 por ciento, aun así, existiría el riesgo de un uno por ciento de culpar a una persona inocente de un determinado crimen como consecuencia del fallo del dispositivo de reconocimiento.

Asimismo, ya se puede dar cuenta de usos polémicos de este tipo de dispositivos en determinados sistemas políticos. Por ejemplo, en el siguiente documental se expone el seguimiento que estaría realizando el gobierno chino contra los integrantes de la etnia musulmana uigur a través del empleo de estos dispositivos. <https://www.youtube.com/watch?app=desktop&v=v7AYyUqrMuQ>

De este modo, entendemos que quizás, el modo de desarrollo de la I.A. en nuestro entorno tomará un camino alternativo al del gigante asiático. En este sentido, quizás nos enfrentemos a un escenario bien distinto a aquel según el cual en las aulas se ubicarían dispositivos biométricos para la detección de las emociones o de los lapsos de atención de nuestros estudiantes².

Aun así, resulta del todo punto desconcertante las posibilidades de control ofrecidas por este tipo de herramientas. Asimismo, cabe preguntarse por los riesgos específicos que serían implícitos a la aplicación de estas herramientas en nuestro contexto determinado.

Suele hablarse de que existe un peligro real de que, a través de una aplicación generalizada de este uso de herramientas, la educación se convierta en un negocio extractor de datos. Este hecho es una posibilidad real puesto que cabe preguntarse por las características de los servicios ofrecidos por estas empresas tecnológicas. ¿Serán estos servicios gratuitos? En este caso este hecho levantaría nuestra sospecha. ¿Serán de pago? ¿Seguirán leyes lo suficientemente estrictas de cara a lo que es la protección de datos?

Estas cuestiones nos conducen a un escepticismo ineludible. La expansión ética de la inteligencia artificial conlleva la realización de un esfuerzo paralelo para el desarrollo de tecnologías parejas de protección de datos. ¿Esperaremos a poder seguir los estándares éticos o tomaremos ciertos sacrificios para poder ser competitivos?

4.1. Un comentario al proyecto de los *neuroderechos* desde un prisma educativo

Regresando al ámbito europeo, no podemos dejar de hacer referencia al proyecto de legislación que está teniendo lugar tras la publicación del “informe sobre Inteligencia artificial en la era digital”.

² Gracias a la *inteligencia artificial*, en países como China, se está pudiendo cumplir el sueño de la monitorización continua del estudiante. Para más información:

Parra, S. (2020, 24 febrero). El colegio en China gracias a la inteligencia artificial. *MuyInteresante.es*. <https://www.muyinteresante.es/tecnologia/articulo/asi-es-el-colegio-en-china-gracias-a-la-inteligencia-artificial-191582559478>

Linares, F. M. (2019, 2 febrero). Así usa China la inteligencia artificial para controlar a sus ciudadanos. *La Vanguardia*. <https://www.lavanguardia.com/tecnologia/20190202/46161322564/china-inteligencia-artificial-ciudadanos.html>

En este se atendía de manera implícita a la dificultad de alcanzar el nivel de I.A. de Estados Unidos o de China al tiempo que se perseguía la configuración de la constitución de una legislación equivalente al Reglamento General de Protección de datos en materia de inteligencia artificial. Al respecto, algunos *expertos* en la materia recriminarían la insuficiencia de la legislación. Su principal vindicación giraría en torno a un desamparo de aquello que sería entendido como “neuroderechos”.

Los *neuroderechos* constituirían una propuesta de ampliación de los derechos fundamentales con base en los recientes estudios del cerebro. A juicio de esta iniciativa, habría que configurar una legislación que velara por la no manipulación cerebral a través del uso de las nuevas tecnologías. La propuesta ha sido iniciada por una plataforma denominada *Neurorights Initiative*.

Nos encontramos así, que, la todavía no presente legislación en torno a inteligencia artificial no estaría siquiera actualizada con respecto a los conocimientos sobre nuestro cerebro que proporciona la neurociencia. El aumento de conocimiento sobre la actividad cerebral que ha proporcionado la neurociencia también ha ocasionado el desarrollo de disciplinas como la neurotecnología y el neuromarketing. Se argumenta que, de esta manera, todavía no existe una legislación suficiente que vele por la seguridad cerebral de los ciudadanos frente al potencial de estas técnicas.

La plataforma que inició esta propuesta, *Neurorights Initiative* habla de cinco neuroderechos, los cuales encontramos preciso comentar y conectar con el tema de nuestra investigación. Aunque su formulación terminológica, desde un prisma filosófico, nos resulte ciertamente metafísica encontramos que el fondo del mensaje expresado por los mismos es del todo punto reseñable. Los neuroderechos comentados serían los siguientes:

1) Derecho a la identidad personal.

Este derecho se fundamenta en la posibilidad de que la neurotecnología cambie nuestra propia percepción del yo.

En el ámbito educativo, podemos preguntarnos si acaso los modelos predictivos de las plataformas inteligentes pudieran tener algún grado de influencia en la construcción del yo adolescente. Al tratar de generar un modelo conductual de la persona, los matices

de esta y el juego de expectativas que influyen en la autopercepción podrían quedar diluidos.

Los modelos predictivos, asimismo, pueden contribuir al refuerzo de ciertas conductas ocasionando “profecías autocumplidas” en el rol desempeñado por el alumno. En las sugerencias de aprendizaje de los algoritmos educativos podría ponerse en riesgo la individualidad de cada estudiante (Holmes et al., 2019). Esto, a su vez, podría en último término llevar a una reducción de la agencia del estudiante por haber sido su desarrollo de aprendizaje condicionado a un principio de estandarización tecnológica.

2) Derecho al libre albedrío.

Este neuroderecho hablaría de la necesidad de protegerse frente a que las neuroinfluencias pudieran tener algún calado en nuestras libres decisiones.

Al respecto, estaría bien preguntarse por el papel que desempeña el carácter adictivo de las herramientas de inteligencia artificial que incorporan algún grado de gamificación. Asimismo, también es pertinente preguntarse por el grado de control que pueden ejercer estas herramientas a través de su recopilación masiva de datos.

Quizás debería tomarse en cuenta aquella sugerencia de la UNESCO (2019) según la cual debería implantarse un principio de mínimos para la recolección de datos. Esto es, los dispositivos habrían de recolectar los datos requeridos para su justo funcionamiento en concordancia con el beneficio público. Todo exceso al respecto, en consecuencia, sería una ilegítima intromisión en la libertad del individuo.

3) Privacidad mental.

Este neuroderecho plantea la necesidad de legislar que la persona pueda decidir cómo se usan los datos de su actividad cerebral. Es decir, para hacer uso de estos se requiere el consentimiento de la persona.

En este aspecto podríamos encontrarnos un problema en los centros. Si nuestro método de enseñanza se basa en este tipo de dispositivos de recolección de datos y los padres de un alumno no dan su consentimiento para esa colecta. ¿Habría de proporcionarse una formación no tecnológica alternativa al alumno en cuestión?

4. Acceso equitativo.

Se sostiene que la mejora en las capacidades cerebrales que traerá la neurotecnología deberían de estar al alcance de todos.

Este postulado tiene un cierto cariz transhumanista. Sin embargo, en lo pertinente al mismo no podemos dejar de ser escépticos.

5. Protección contra los sesgos.

Este derecho se fundamentaría en la necesidad de realizar una protección contra las discriminaciones. Al respecto no podemos dejar de hablar de que esto parece complicado puesto que la caja negra de los algoritmos nos amenaza con una opacidad del proceso. ¿Cómo evitar un resultado de un proceso que desconoces?

A nuestro juicio, deberíamos de empezar a asumir que la inteligencia artificial también está situada. Constituye un riesgo real de la enseñanza que los algoritmos se sitúen como modelo evaluador de objetividad pues, como reiteradamente queda demostrado, los algoritmos también tienen prejuicios.

4.1.1 Algoritmo humano demasiado humano

En la *Declaración de Barcelona*, reunión consistente en plantear la problemática de la posible robotización de la sociedad, se mostraban dos efectos de los algoritmos. Las *cámaras de eco* y las *burbujas de filtro*.

Las *cámaras de eco* serían aquel efecto de repetición de la información que se produciría como consecuencia de la búsqueda de personalización de la experiencia tecnológica. Este efecto supone un riesgo para la cognición puesto que refuerza los sesgos confirmatorios e incluso disminuiría la información recibida desde el entorno.

Las *burbujas de filtro*, por otro lado, serían aquella limitación de la información recibida a causa de los algoritmos que seleccionan el contenido mostrado. Este simple funcionamiento de la tecnología que posibilita una supuesta mejora a causa de la personalización de cara al usuario también entraña el riesgo de crear una burbuja, como gráficamente lo muestra la expresión, o sesgo ideológico o cognoscitivo sobre el usuario. De esta manera se filtrarían aquellas informaciones que no coincidirían con el punto de vista del usuario a su vez generando un refuerzo de sus propias convicciones.

Ambos efectos, constituyen problemas graves que afectan a nuestros alumnos y a la sociedad como conjunto. De hecho, un efecto curioso de este tipo de resultados es el hecho de que encapsula al individuo de tal manera que, si a este se le muestra una

información que contradice sus creencias, éste se refuerza más en los prejuicios que perpetuaría la burbuja de filtro. Esto podría ser ciertamente relacionado con la presente proliferación de tendencias conspiracionistas y radicales a nivel social.

¿Qué consecuencia tendrá el uso de la inteligencia artificial en el ámbito educativo? ¿Se crearán burbujas de filtro en lo enseñado fomentando una visión compacta del conocimiento obviando la pluralidad de matices del mismo y su propio cuestionamiento? ¿Qué consecuencias puede tener a nivel social educar con burbujas de filtro?

5. Estado de la cuestión.

A continuación, vamos a realizar una breve demarcación del estado de la literatura adecuada al planteamiento de la cuestión abordada. Para ello, realizaremos un recorrido por los planos más amplios del problema, a saber, el panorama general de qué es aquello de la inteligencia artificial y, posteriormente, el plano medio el cual se pregunta por la concreta aplicación de la inteligencia artificial en el ámbito educativo.

Con estos dos movimientos, pretendemos despojarnos de ciertos prejuicios en torno a este auge tecnológico. Para ello, atenderemos a que la inteligencia artificial es algo concreto cuya aplicación en la actualidad no coincidiría con la representación que tomaría cuerpo en los imaginarios ficticios. De igual manera, cartografiaremos una serie de posibles usos de la herramienta que podrían tener lugar en el aula. Todo ello nos servirá para, en la discusión posterior, preguntarnos por aquella vinculación que existiría entre pedagogía y tecnología en el uso de la *inteligencia artificial* en el aula.

5.1. Qué es la inteligencia artificial y qué problemas nos presenta.

El término *inteligencia artificial* o I.A. alude tanto a una herramienta como a la sucesión de disciplinas que posibilitan la aplicación de esa herramienta. De esta manera, *inteligencia artificial* no se corresponde con un término unívoco. Cuando hablamos de la I.A. podemos referirnos indistintamente a un rico panorama de aplicaciones. Podemos aludir tanto a mecanismos de filtrado de información como a *softwares* de reconocimiento de objetos e incluso a dispositivos de personalización de la experiencia de compra.

No obstante, un aspecto común a esta serie de dispositivos estaría constituido por el hecho de que, en todos estos usos, la herramienta *artificial* o computacional estaría llevando a cabo una serie de decisiones que no habrían sido estrictamente programadas por los informáticos.

Asimismo, la I.A. no se diferencia en exclusiva por campo de aplicación, sino que también presenta una diversidad de funcionamientos. Podemos tener, por ejemplo, sistemas de reconocimientos de patrones de datos como el *machine learning*, cuyo modo de proceder o de aprender, a su vez, puede tomar diferentes versiones. El algoritmo de *machine learning* puede funcionar por aprendizaje supervisado, el cual generaría modelos basándose en previas categorizaciones de los datos dadas, por aprendizaje no supervisado, en el cual el algoritmo sería capaz de extraer categorizaciones de los datos sin previas indicaciones e incluso aprendizaje por refuerzo, en el cual, el modelo predictivo generado

a partir de los datos para tomar decisiones se iría configurando gracias al *feedback* que iría recibiendo el algoritmo (Holmes et al., 2019). De igual manera, cuando hablamos de la I.A. también podemos referirnos a disciplinas como el *Deep learning* cuyo modo de funcionamiento se basaría en una modelización del funcionamiento del cerebro humano. El *Deep learning* opera a través de capas de redes de neuronas artificiales con unos resultados sorprendentes³.

No obstante, la cuestión planteada no puede ser, ni mucho menos, resuelta a través de una mera enumeración de las aplicaciones actuales de aquello que en el presente conocemos como *inteligencia artificial*. Ello es debido a que, la propia denominación del término “inteligencia artificial” alude a un problema filosófico complejo.

Asimismo, este problema filosófico, a nuestro juicio, no podría resolverse en una mera reformulación de las palabras empleadas. Filósofos como Markus Gabriel parecerían aludir a una suerte de pseudoproblemática planteando que, en realidad, el debate sobre la constitución artificial de la inteligencia sería falsario por ser ya la inteligencia humana constituida artificialmente a través de la cultura⁴. Con esta propuesta no podemos estar de acuerdo.

Sostenemos que la denominación “inteligencia artificial” contiene de suyo una problemática histórica que nos hace reflexionar sobre las posibilidades de la técnica, así como sobre la naturaleza de nuestra propia inteligencia. En el debate sobre la posibilidad de la constitución de una inteligencia artificial siempre ha estado de fondo el problema de la modelización de la inteligencia humana.

Ya en uno de los padres de la disciplina, Allen Newell, encontramos una concepción sobre el humano y sobre su aprendizaje muy específica que, a su vez, se basa en un modelo computacional. Para este autor, existiría una analogía casi idéntica entre el funcionamiento de la mente humana y el del ordenador. Siendo así que, al igual que el computador:

“La mente humana también es un sistema de procesamiento de la información, y una descripción de la misma puede darse en una aproximación de las mismas características. [...] El

³ Esta serie de precisiones técnicas en torno a la inteligencia artificial quedan explicadas de una manera más detallada en el glosario anexo.

⁴ Estamos aludiendo a una tesis sostenida por éste en la obra *El sentido del pensamiento*. En: Gabriel, M. (2022). *El sentido del pensamiento*. Ediciones de Pasado y Presente.

modelo de procesamiento humano puede dividirse en tres subsistemas que interaccionan entre sí (1) el sistema perceptual, (2) el sistema motor y (3) el sistema cognitivo, cada uno con sus propias memorias y procesadores⁵. (Card et al., 1983, p. 23)

Esta cuestión, por tanto, no puede eludirse en un mero desplazamiento de la pregunta por cuanto que, entre otras cosas, estaría en juego la imposición de un modelo computacional e incluso conductista de la inteligencia que, para nuestro caso, puede influir en el modo de concebir la enseñanza. Pensar en qué queremos decir sobre las cosas cuando nos planteamos por la constitución de una inteligencia artificial es, en consecuencia, del todo punto imprescindible.

Asimismo, como veníamos diciendo, la denominación “inteligencia artificial” también alude a la posibilidad de simular la inteligencia propia del humano a través de medios no orgánicos. Es en este punto en el cual se observaría un enraizamiento de esta cuestión con la problemática del *transhumanismo*.

Esta filosofía o ideología, aprovecharía la temática de la producción artificial del humano para llegar al postulado de un progresivo mejoramiento técnico del mismo. El humano habría entonces de ser superado en una evolución controlada. Esta escalada llegaría incluso al punto de haber trascendido nuestra condición biológica por completo y, en consecuencia, alcanzado la *singularidad tecnológica*. (Kurzweil, 2021).

No obstante, a nuestro juicio, estos planteamientos no son uniformes con respecto a lo que debe mejorarse en esta evolución técnica controlada. Incluso autores como Nick Bostrom (2016), los cuales parecen centrarse en el aspecto inteligente de la mejora, no serían del todo claros con respecto a lo que es la inteligencia. Más aún, en el campo de reflexión acerca de la concreta *inteligencia artificial* se aprecia que, puesto que no hay un acuerdo en torno a la consistencia de nuestra inteligencia, tampoco lo habría en la definición conceptual del término.

Autores como Marvin Minsky, parecen optar por una definición de la I.A. basada en una noción conductista de la inteligencia. Para este autor, la *inteligencia artificial* sería aquella disciplina que permitiría desarrollar una serie de acciones de las cuales si fueran realizadas por humanos inferiríamos inteligencia (Enciclopedia Herder, 2009).

⁵ La traducción es nuestra.

Otros organismos parecerían incidir en el aspecto inteligente de la atribución de fines. De este modo, *inteligencia artificial* también podría aludir a sistemas que mostrarían un comportamiento inteligente siendo capaces de realizar un análisis del entorno, así como de realizar acciones con cierta independencia con el propósito de alcanzar objetivos específicos (Comisión europea, 2018).

Del mismo modo, otros parecen hacer hincapié en el carácter de automatización de cálculos complejos como vía de simulación de actividad inteligente⁶. Richard Bellman en *An Introduction To Artificial Intelligence* señalaría así que la inteligencia artificial es “la automatización de actividades que vinculamos con procesos de pensamiento humano, actividades tales como toma de decisiones, resolución de problemas, aprendizaje...”. (Moreno Padilla, 2019, p.262).

Dicho esto, ya hemos expuesto varios motivos por los cuales deberíamos de apreciar que la expresión *inteligencia artificial* no es un vocablo unívoco. No obstante, todavía queda por expresar una última apreciación polisémica del término que aqueja a la propia denominación de la disciplina y que la pone en cuestión tal y como fue concebida como proyecto inicial.

Nos estamos refiriendo a la fundamental distinción, acuñada por el filósofo John Searle en 1980, que existiría entre la *inteligencia artificial* en un sentido *fuerte* y la *inteligencia artificial* en un sentido *débil*. Esta última actualmente no ha sido alcanzada y a juicio de algunos, como el propio Searle, no podrá serlo nunca⁷. A pesar de todo, la *inteligencia artificial fuerte* o *inteligencia general* de las máquinas es la que más frecuentemente aparece representada en las obras de ficción que versan sobre la inteligencia artificial. Por ejemplo, en la novela *El quinteto de Cambridge* de John L. Casti, la discusión de los personajes orbitaría en torno a la posible constitución de una inteligencia general en las máquinas. De igual manera, en la novela *Klara y el sol* del

⁶ En este fragmento hemos puesto como ejemplos varios modos distintos de conceputar lo que sería la inteligencia artificial. No obstante, el número de definiciones de esta es extremadamente elevado. Al respecto, consideramos de gran valía la distinción entre inteligencia artificial como simulación de la inteligencia humana y la inteligencia artificial concebida desde la constitución de una racionalidad abstracta que queda expresada en la entrada sobre inteligencia artificial de la Enciclopedia Stanford. Para ampliar en esta cuestión consultar: Bringsjord, Selmer, & Naveen Sundar Govindarajulu. (2020). Artificial Intelligence. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2020/entries/artificial-intelligence/>

⁷ La argumentación de por qué no podrá ser alcanzada la inteligencia artificial en sentido fuerte se encuentra en el famoso artículo de Searle referenciado a continuación. Searle, J. R. (1980). “Minds, brains, and programs. *Behavioral and Brain Science*”. vol. 3, n.º 3, 417-457.

nobel de literatura Kazuo Ishiguro, el personaje protagonista, Klara, aludiría a un robot que ya ha alcanzado la inteligencia general y que, en consecuencia, presenta un comportamiento adaptativo ante diversas situaciones.

La *inteligencia general* o *inteligencia artificial en sentido fuerte* consistiría en el logro de hacer que las máquinas adoptaran un tipo de inteligencia capaz de adaptarse a múltiples circunstancias. Esta inteligencia sería evolutivamente similar a la de un ser humano.

El humano tiene la posibilidad de actuar según las condiciones variables del contexto y de transportar sus aprendizajes a otros campos que no son los específicos en los cuales adquirió estos. Esta facultad en el campo de la enseñanza y de la psicología ha recibido el nombre de *transferencia*. En psicología, el *transfer* conlleva que la persona utilice lo transferido para contextos distintos a los cuales fue concebido inicialmente (López, 2019). Por el contrario, en pedagogía la sistematización de la *transferencia* es más exhaustiva. Hay múltiples criterios de diferenciación de la transferencia aplicada al aprendizaje de entre las cuales podríamos reseñar la diferencia que existiría entre la transferencia positiva, neutra y negativa.

La transferencia positiva sería aquella que se produciría cuando el entrenamiento en una tarea mejora el desempeño en otra (López, 2019). Por ejemplo, podemos pensar que la lectura de novelas puede mejorar la habilidad en realizar comentarios de texto. La transferencia negativa sería aquella en la cual el desempeño de una actividad empeora la competencia en otra. Por ejemplo, podemos pensar que jugar cuatro horas diarias al *fortnite* puede empeorar la capacidad de atención en clase de un alumno. Por último, la transferencia neutra sería aquella en la cual se demarcaría que el desarrollo de una habilidad es indistinto o no afecta al desarrollo de otra.

Pues bien, el problema de la *inteligencia general* de las máquinas queda expresado de manera nítida con el ejemplo siguiente.

Cuando se tenía como sueño lograr la constitución de un robot que jugara al ajedrez y que pudiera vencer al ser humano, se pensaba que se habría dado con aquella piedra de toque que configuraría a la inteligencia humana como tal. Jugar al ajedrez era una habilidad que se concebiría como representante de la excelencia del intelecto humano. En consecuencia, se infería que, si se lograra configurar una máquina que venciera al mejor ajedrecista del mundo, esto supondría el hito de haber configurado a través de

nuestra propia inteligencia una inteligencia superior a nosotros. No obstante, cuando la famosa computadora diseñada por IBM, *Deep Blue*, ganó en el año 95 al conocido ajedrecista Kasparov, la victoria resultó agrídulce.

Se había logrado que la máquina ganara a un humano en la competencia específica de jugar al ajedrez, pero, sin embargo, *Deep blue* no sabía hacer nada más. *Deep blue* no era capaz de tener una conversación, de servir un vaso de limonada o de hacer una simple cuenta. Esta máquina verdaderamente no había logrado igualar a Kasparov en el sentido global de su inteligencia, sino que tan solo sabía maquinar combinatorias muy buenas para esa finalidad específica de jugar al ajedrez.

Es por ello, por lo cual, en el posterior análisis de la aplicación efectiva de la *inteligencia artificial* en el aula, tan solo atenderemos a aplicaciones correspondientes a aquello que sería cualificado como *inteligencia artificial* en sentido estrecho o *débil*. Son lo actualmente disponible, aunque, no obstante, la cualidad específica de la inteligencia general que nos caracteriza como humanos es lo que parece estar impregnando las propuestas de reconfiguración del currículum. Esto será visto con mayor detalle en los apartados posteriores.

No obstante, algo que si podemos preguntarnos en este momento de la exposición es si acaso la denominación *inteligencia artificial* es la más indicada para la efectiva aplicación de técnicas como el *machine learning* en el ámbito de la escuela. Según algunos autores una denominación más adecuada al efectivo transcurso de la disciplina sería “ciencia de datos”.

La *inteligencia artificial* en ese caso quedaría ligada no a una suerte de realización de un sueño metafísico de autoproducción de la inteligencia humana sino, por el contrario, a una concreta e incluso “modesta” pero ambiciosa disciplina del tratamiento de datos. Eso sí, sería una técnica que desde sí misma estaría cambiando nuestro modo de aproximarnos a la realidad, las necesidades de las empresas e incluso produciendo lo que algunos especialistas en inteligencia artificial como Andrew Ng denominan “cuarta revolución industrial” (Rouhiainen, 2018).

Asimismo, como es evidente, la posibilidad de análisis de los datos estaría aparejada al desarrollo de la *ciencia de datos*. En consecuencia, con una frase atribuida al millonario Clive Humby, podríamos decir que “los datos son el nuevo petróleo”. Si bien, ha de reseñarse que esta cuestión también podría generar desigualdades sociales por

cuanto que los recursos quedarían recentrados en aquellas grandes compañías con la posibilidad de recabar un enorme número de datos y, en consecuencia, generar los mejores modelos predictivos de la disciplina (Alok Jha, 2022-presente, 24m40s).

Como veníamos diciendo, el auge material de la *ciencia de datos* habría traído aparejado la eclosión de una serie de conceptualizaciones e ideologías en torno a los datos. Ello se manifestaría, en cierto modo, en el uso continuado de expresiones como “sociedad de la información” o “sociedad del conocimiento” por parte de organismos como la UNESCO y la Comisión Europea.

Aunque el término “sociedad de la información” fuera acuñado en el año 1973 por el sociólogo estadounidense Daniell Bell en *El advenimiento de la sociedad postindustrial*, el calado del término es del todo punto operativo en las concepciones de los días presentes. Este término, aludiría a un tipo de sociedad en la cual el conocimiento y la información serían el principal eje quedando las ideologías desenmascaradas como excedente a la misma⁸. (Burch, 2005).

No obstante, podríamos pensar que este planteamiento de la información como instancia no ideológica es, desde sí misma, bastante ideológica. De hecho, el pensar que la información es algo por sí mismo objetivo, independientemente de los plexos de sentido, es un esquema de pensamiento que en la actualidad se encontraría conservado estrictamente en el culto indiscriminado hacia la objetividad de los datos.

Este culto ha sido conceptualizado como *dataísmo* y estudiado por el filósofo Byung-Chul Han en obras como *Psicopolítica* y *La salvación de lo bello*. El *dataísmo* sería una ideología que sostendría que la realidad podría ser interpretada a partir de la mera lectura de los datos. Tanto es así que ingenieros como Chris Anderson plantearían una quiebra del método científico a causa de la disciplina de datos⁹.

⁸ Esta concepción también calaría en la idea de una sociedad en la cual la producción intelectual no ha de estar estrictamente vinculada a la manual para su posterior explotación. Como indica el sociólogo Manuel Castells: “Por primera vez en la historia, la mente humana es una fuerza productiva directa, no sólo un elemento decisivo del sistema de producción”. (Castells citado en Burch, 2005, p. 4).

⁹ Para este autor, la ciencia de datos habría llegado a tal desarrollo que podría “desbancar” al resto de ciencias. De esta manera, para comprender el mundo solo haría falta un ordenador, ingentes cantidades de datos y herramientas estadísticas. No importaría la correlación. No harían falta modelos coherentes, teorías unificadas ni explicaciones mecanicistas. En: Anderson, C. (2008, 23 junio). “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete”. *Wired*. <https://www.wired.com/2008/06/pb-theory/>

El dato quedaría concebido como pura información y transparencia. Asimismo, la verdad de las cosas quedaría progresivamente desvelada por la reiterada acumulación de datos. Este proceder *dataísta*, a juicio de Han, estaría llevando a la sociedad a una suerte de segunda ilustración. Esta segunda ilustración también buscaría la liberación de mitos en el conocimiento con el matiz de que este proceso depurativo tomaría cuerpo a través de la ideal transparencia derivada de la acumulación de datos. (Han, 2014). Convirtiendo las cosas en datos e información sería como las cosas se harían transparentes. Más aún, el puro dato se presentaría en ocasiones como modelo de objetividad frente a la subjetividad del conocimiento.

En el ámbito que nos corresponde, esta reflexión cobra especial relevancia puesto que, como ya hemos advertido en el comentario a la legislación, corremos el riesgo de pensar que los dispositivos de inteligencia artificial no tienen prejuicios, sino que están liberados de toda la subjetividad humana. Tanto es así que, habríamos de vencer una suerte de tendencia a delegar nuestra labor como profesores hacia la supuesta objetividad de los algoritmos. Pero, todo ello, será planteado de una manera menos sumaria en el desarrollo posterior.

En consecuencia, en el desarrollo previo, hemos atendido a que en la actualidad “tan solo” se habría logrado desarrollar una *inteligencia artificial débil* no siendo esta correspondiente con el imaginario ficticio en torno a lo que la inteligencia artificial representaría. De igual modo, apreciamos que la denominación “inteligencia artificial” traía consigo la demarcación de un problema filosófico muy complejo que es la pregunta por la posibilidad de reproducción de la inteligencia humana, así como por el discernimiento de las características de esta. Esta reflexión sobre las posibilidades de la técnica se vinculaba con la cuestión transhumanista. Además, veíamos así que la propia conceptualización del funcionamiento de la “mente” humana podría influir en nuestro desempeño como profesores. Asimismo, atendíamos a la fundamental cuestión de la inteligencia general humana y a cómo ésta se relacionaría con facultades imprescindibles en el fenómeno de aprendizaje como es la capacidad de *transferencia*.

Una vez, realizada esta delimitación nos vemos en condiciones de realizar una exposición crítica de las efectivas aplicaciones de la *inteligencia artificial* en el ámbito de la enseñanza.

5.2. Aplicación efectiva de la inteligencia artificial en el ámbito educativo.

A continuación, vamos a realizar una breve delimitación de las principales aplicaciones de la *inteligencia artificial* que estarían siendo llevadas a cabo en el campo de la enseñanza.

De manera preliminar, podemos señalar que las diversas herramientas no se fundamentan en una sola concepción de la enseñanza. Por el contrario, el desarrollo técnico concreto tiene a su base la atribución de una finalidad determinada, aspecto crítico que podríamos vincular con lo propiamente humanístico. Vemos así, aplicaciones basadas en el diálogo, así como otras que incorporarían algún grado de gamificación hasta llegar incluso a algunas que se fundamentan en una concepción constructivista de la enseñanza.

De este modo, realizaremos un comentario de seis aplicaciones posibles seguido de un último apartado destinado a la reflexión sobre aquello que posiblemente podrá tener lugar. Para demarcar la clasificación general de las aplicaciones educativas, hemos usado como texto guía el esquema planteado en el libro *Artificial Intelligence in Education* de Holmes, Bialik y Fadel.

No obstante, el mismo no se ha conservado intacto pues el mero parafraseo de esta referencia no es suficiente en el estado actual de la técnica. A pesar de ser muy detallado, el texto ha sido publicado hace apenas tres años, tiempo más que considerable en el desarrollo de esta disciplina. Por ello, complementaremos nuestra explicación con reestructuraciones, consideraciones críticas, así como con referencias adicionales.

La clasificación propuesta es la que sigue.

a) Sistemas de tutoría inteligentes (*Intelligent Tutoring Systems*)

Estos sistemas se caracterizan por proporcionar un aprendizaje guiado a los estudiantes en un entorno interactivo y personalizado (Rodríguez Chávez, 2021).

El primer sistema de tutoría inteligente fue diseñado en 1970 por el informático uruguayo Jaime Carbonell como contenido de su tesis doctoral. Este sistema, llamado SCHOLAR, era concebido como un tipo de instrucción asistida por ordenador (*Computer Aided Instruction*). SCHOLAR era un programa capaz de evaluar el conocimiento de un estudiante en un contexto determinado, como podría ser el campo de la geografía europea, llevando a cabo un diálogo en el cual ambos participantes podían tomar la iniciativa (Carbonell, 1970). De este modo, el sistema era capaz de dar instrucciones automáticas de

manera personalizada siendo posible establecer conversaciones rudimentarias con el mismo. Al respecto, podríamos sostener que, en la actualidad tenemos a nuestra disposición versiones modernas de este “antiguo” prototipo.

Este tipo de sistemas se caracterizan por proporcionar tutoriales paso a paso adaptados al alumno concreto. La adaptación consiste en que el sistema registra los errores que comete el alumno y en base a ello le configura un camino de aprendizaje personalizado. Asimismo, el sistema va adaptándose a los progresos del alumno pues según el alumno va mejorando o empeorando el sistema ajusta su nivel de dificultad. El asistente proporciona guía hasta que considera que el alumno ha dominado la materia.

No obstante, debe reseñarse que la aplicación de estos sistemas de tutoría inteligente está tomando un claro enfoque hacia las disciplinas de la Matemática y la Computación (Rodríguez Chávez, 2021). Ejemplos de estos sistemas de tutoría inteligente son *Andes*, ITS destinado a la instrucción de Físicas y *Mathia*, el cual da instrucción personalizada para el aprendizaje de las Matemáticas.

b) Sistemas de tutoría basados en el diálogo.

Estos sistemas en lugar de proporcionar caminos de aprendizaje personalizado lo que hacen es animar a los alumnos a emprender una conversación con ellos sobre una materia específica. En el transcurso del diálogo con el tutor es cuando se va produciendo el aprendizaje al tiempo que se generan modelos de personalización de este.

Ejemplos de este tipo de sistemas sería por ejemplo *AutoTutor* el cual lo que hace es tratar de enseñar a través del diálogo, en una inspiración socrática (Holmes et al., 2019). *AutoTutor* plantea una pregunta y el alumno trata de responderla introduciéndola por escrito. Si el sistema detecta que no hay comprensión del tema entonces asigna al alumno lecturas o le proporciona preguntas que le ayuden. El tutor va acompañando al alumno hasta que éste da con la respuesta idónea (Graesser et al., 2004).

c) Agentes de software conversacionales inteligentes (*chatbots*).

Los *chatbots* o agentes conversacionales, son mecanismos de software con la capacidad de generar interacciones con seres humanos a través del lenguaje natural. (Rodríguez et al., 2021).

Los *chatbots* habitualmente son planteados como meros asistentes de apoyo de las clases presenciales. Todavía no presentan un grado de desarrollo tal que pudieran sustituir

la labor docente y, posiblemente, considerando la imposibilidad de que desarrollen algún grado de inteligencia general, nunca lo harán.

Su finalidad suele consistir en atender a los estudiantes en tareas monótonas basándose en modelos predictivos de las dudas que suelen tener. Se plantea así, que los *chatbots* puedan liberar de tareas repetitivas a los profesores, generando respuestas automáticas ante dudas similares. De igual manera, suele plantearse que si estos tienen voz desarrollada debería de establecerse por ley que ellos mismos declararan que son robots antes de iniciar una conversación (Rouhiainen, 2018).

Ejemplos de *chatbots* son Pepe, el *chatbot* de *Naturgy*, y Noa el asistente virtual de *CaixaBank*.

d) Evaluaciones automáticas de redacción.

Estos sistemas pueden adoptar dos enfoques. Hay algunos que ayudan al alumno en el mismo proceso de redacción, aportándoles sugerencias para escribir mejor. Por otro lado, hay otros que cuando se entrega la tarea escrita proporciona una puntuación automática del proceso de redacción (Holmes et al., 2019).

Esta última aplicación no deja de resultarnos sorprendente puesto que la inteligencia artificial tan solo es capaz de procesar información, pero no de comprenderla. ¿Qué patrones formales tiene establecidos para juzgar la calidad de un texto? Asimismo, ¿Qué pasa si los estudiantes tienen acceso a dispositivos de escritura automática y a la vez son evaluados con dispositivos de evaluación de ensayos automáticos?

Las perspectivas de crítica pueden multiplicarse. Podría argumentarse que los evaluadores automáticos concebirían la escritura como un mero acto de demostración formal y no como un acto comunicativo (Ruiz Herrero, 2016). De igual modo, desde un enfoque wittgensteniano podríamos argumentar que estos dispositivos de evaluación no tienen acceso al modo de vida en el cual fueron escritos por los estudiantes.

Ejemplos de este tipo de herramienta son *Write Tolearn*, *e-Rater* y *Open Essayist*.

e) Entornos exploratorios de aprendizaje (*Exploratory Learning Environment*)

Los entornos exploratorios de aprendizaje o ELEs comprenden un complejo interactivo que va desde simuladores, micromundos a laboratorios y juegos educativos (Mavrikis et al., 2016).

Este tipo de herramientas proporcionan paisajes con los cuales el alumno pueda interactuar y, de este modo, construir su propio conocimiento. Como se puede ver, los ELE's tienen un enfoque constructivista. No obstante, este carácter es motivo de crítica habitual puesto que se estima que, sin un cierto grado de instrucción, la enseñanza va a producirse de una manera no optimizada (Holmes et al., 2019). Es posible que, los estudiantes se enfrenten a una sobrecarga de estímulos ante la cual no saben cómo actuar.

Es por ello, por lo cual muchos de estos entornos de aprendizaje han incorporado una suerte de guía o tutor para proporcionar cierto grado de retroalimentación a los alumnos, así como sugerencias si estos se encuentran muy perdidos. Como se ve, es un enfoque completamente distinto al de los sistemas de tutorización a través del diálogo. Aquí no hay una serie de pasos prefijadas que el alumno deba dar, sino que él mismo se encuentra solo para habérselas con el paisaje diseñado y configurar su propio conocimiento.

Ejemplos de entornos exploratorios de aprendizaje son *Crystal Island* y *ECHOES* los cuales van a ser explicados a continuación de manera sumaria.

Crystal Island es un tipo de entorno gamificado. El alumno se convierte en un detective y tiene que ir resolviendo misterios que requieren el desarrollo de ciertos conocimientos. El juego puede seguir varias vías haciendo que el estudiante desarrolle competencias determinadas.

Del mismo modo, *ECHOES* es un ELE que también incorpora un enfoque de gamificación. No obstante, se especializa en ayudar a alumnos con trastornos del espectro autista. Un profesor humano externo elige una de las doce actividades programadas para que el niño interactúe con Andy, un asistente virtual. Con este tipo de actividades se pretende que el niño desarrolle habilidades comunicativas, así como la atención. El agente virtual Andy tiene una función que le permite detectar adonde se dirige la mirada del niño y así saber reconducir su atención. (Holmes et al., 2019)

De acuerdo con lo expuesto, apreciamos la enorme valía que pueden tener este tipo de aplicaciones. No obstante, no dejamos de poner de manifiesto la importancia de que estas herramientas supongan un complemento a la enseñanza y no un sustitutivo. Siempre quedaría el riesgo de que este tipo de herramientas virtuales, o más cercanas al campo de la simulación virtual, puedan convertirse en un entorno de escape en donde depositar a los alumnos con características del aprendizaje distintas a los demás.

f) Plataformas para el aprendizaje de idiomas.

Estas aplicaciones son quizás las más célebres. Ejemplos de estas plataformas online para el autoaprendizaje son *Duolingo* y *Babbel*.

Estas plataformas evalúan el nivel de inglés y en base a este van proporcionando conocimientos y vocabulario específico. Tienen software de reconocimiento de voz para apreciar la pronunciación, así como un procesador del lenguaje natural para valorar la calidad de la redacción y de la comprensión lectora. Son capaces de establecer niveles de progreso al tiempo que, *Duolingo*, por ejemplo, incorpora un cierto nivel de gamificación proporcionando puntos canjeables a los buenos estudiantes.

Numerosas personas están mejorando su nivel de idiomas actualmente desde sus dispositivos móviles a través del uso de estas herramientas. En este sentido, estas plataformas ofrecen la ventaja de “democratizar” el conocimiento haciéndolo más accesible a todos.

g) Plataformas online para el autoaprendizaje.

Estas plataformas también suelen ser designadas como plataformas *e-learning*. Se caracterizan por ofrecer la adquisición de una serie de conocimientos con independencia del horario lectivo a la vez que, por generar estadísticas, patrones conductuales e incluso por programar fechas límite y asignar tareas.

Este tipo de dispositivos son más típicos de la enseñanza no formal. Ejemplo de ello son los cursos en línea masivos y abiertos (MOOC) que ofrece la plataforma Coursera.org. Del mismo modo, en el ámbito educativo resaltaría la plataforma *Khan Academy* en la cual profesores de todo el mundo ofrecen gratuitamente vídeos instructivos sobre las más variadas disciplinas de primaria y secundaria.

Las ventajas que ofrecen este tipo de plataformas es que permiten llevar la educación y el conocimiento técnico de determinados asuntos a los lugares más remotos. (UNESCO, 2019). De igual manera permite la combinación de diferentes materiales de aprendizaje.

Una desventaja que apreciamos se derivaría de la posibilidad de que esta plataforma sustituyera el modo de educación presencial y en vivo, por cuanto que ello implicaría entre otras cosas la pérdida del diálogo humano y posiblemente de habilidades comunicativas e interpersonales en los alumnos. Asimismo, su aplicación siempre va a

estar sometida a la desigualdad inherente a la todavía vigente *brecha digital* (Almenara, 2006).

g) Qué usos y aplicaciones podrían tener lugar.

De acuerdo con los dispositivos expuestos podríamos preguntarnos por cómo estos, en sus diversos enfoques, podrían condicionar el estilo de enseñanza.

Para empezar, podríamos pensar en que los tutores inteligentes, así como los *chatbots* asistenciales tendrían el riesgo de dotar de un carácter más dependiente a los alumnos. El hecho de que un *chatbot*, por ejemplo, pudiera servir de apoyo a un estudiante a la hora de hacer los deberes representaría múltiples ventajas. No obstante, habría que cuidarse de que el alumno desarrollara una baja tolerancia a la frustración y que, ante cualquier duda, optase por pedir ayuda al dispositivo. Asimismo, el hecho de que los sistemas de tutoría inteligente se centren en ámbitos científicos también podría representar un riesgo. Las ciencias no pueden entenderse en exclusiva como métodos estandarizados para llegar a la respuesta correcta. Hay procedimientos científicos que requieren del uso de la creatividad. Por ejemplo, podemos pensar en los diversos modos que hay de realizar una demostración matemática. De igual modo, podemos realizar una demostración matemática por reducción al absurdo, pero ésta tampoco tiene una única manera de realizarse.

Por otro lado, debemos de dejar patente que no hemos mencionado los dispositivos de identificación biométrica en el aula puesto que se estima que no serían potencialmente aplicables al entorno político en el cual nos encontramos. Asimismo, podríamos señalar que quizás tener la posibilidad de medir las respuestas emocionales de nuestros alumnos y cronometrar sus lapsos atencionales, así como los ejercicios en los cuales emplearían más tiempo podría ayudarnos a optimizar el proceso de enseñanza, en unos parámetros fabriles, pero ello, a nuestro juicio, supondría un coste demasiado alto.

En último lugar, podríamos señalar diversas vías técnicas que podrían desarrollarse en los próximos años.

Es posible que se desarrollen aplicaciones que fomenten y supervisen el aprendizaje colaborativo. Se habla del desarrollo de herramientas de creación automática de grupos de trabajo (Holmes et al., 2019). Los alumnos quedarían agrupados según estándares preestablecidos en busca de perfiles compatibles. Éstos podrían asociarse, por

ejemplo, por nivel de conocimiento, así como por tipos de personalidad. Son muchos los criterios posibles y, a nuestro juicio, estos habrían de ser decididos en la mayor medida por el profesor de manera que sean acordes a sus finalidades didácticas.

Como ayuda al trabajo colaborativo también se sugiere la creación de agentes virtuales que supervisen el proceso de cooperación, así como contribuyan a apaciguar discusiones que puedan surgir en la agrupación. Al respecto, queda por comprobar que la intervención de un *chatbot* en una discusión pueda calmar el conflicto.

Otra propuesta que debe ser tomada en cuenta sería la posible creación de un *software* que monitoree los foros de estudiantes (Holmes et al., 2019). Esta herramienta podría aliviar la tensión inherente a la acumulación de mensajes en los foros. Las dudas similares podrían ser agrupadas de tal manera que el profesor pudiera responder de manera simultánea a varias preguntas a la vez. De igual manera, el dispositivo podría realizar una “lectura” de conceptos clave de los *posts* publicados por los estudiantes para redirigirlos a otros de contenido similar.

En consecuencia, ya hemos realizado una demarcación de los dos primeros planos de nuestro análisis. Tras dar este paso estamos en condiciones de plantear una serie de problemas en torno a la enseñanza de las humanidades. Sin embargo, de manera soterrada comenzaremos desde el no evidentemente relacionado pero fundamental problema del entrecruzamiento entre la pedagogía y la tecnología. En este campo, se vislumbraría a nuestro juicio la inherente finalidad que subyace a todo “mejoramiento” técnico.

6. El papel de las humanidades en la era de la inteligencia artificial

En el desarrollo previo, hemos atendido a una reconducción de lo que la *inteligencia artificial* representaría actualmente en el plano educativo y global al tiempo que atendíamos a las problemáticas inherentes a esta disciplina. Del mismo modo, veíamos que la misma tenía diferentes aplicaciones educativas con enfoques muy variados. Asimismo, estas aplicaciones no tenían de por sí incorporadas una intencionalidad de sustitución del humano o de guía al mismo hacia el final de los tiempos. Al contrario, la tecnología educativa parecería constituir un campo abierto de posibilidades con el potencial de dirigirse a múltiples metas.

La *inteligencia artificial*, por tanto, parece no tener un destino predefinido. En este aspecto, la condición de la tecnología parece alinearse con la situación de la educación. Ambas son herramientas con un potencial transformador inmenso pero su aplicación concreta depende del establecimiento de unos fines determinados. Como señala Juan Delval:

“Una reflexión sobre los fines de la educación es una reflexión sobre el destino del hombre, sobre el puesto que ocupa en la naturaleza, sobre las relaciones entre los seres humanos”. (Savater, 2010, p.42).

Esta inicial intuición es la que guio la orientación de nuestra investigación. Inspirados por el tópico arendtiano de la autoproducción del hombre, vimos en la temática de la *inteligencia artificial* su más estrecha realización. Más aún llegamos a apreciar que la educación con inteligencia artificial constituía un modo de paradójica autoproducción cultural del hombre a través de un dispositivo cuya finalidad inicial residía en modelizar nuestra inteligencia.

“Este hombre futuro- que los científicos fabricarán antes de un siglo, según afirman- parece estar poseído por una rebelión contra la existencia humana tal como se nos ha dado, gratuito don que no procede de ninguna parte (materialmente hablando), que desea cambiar, por decirlo así, por algo hecho por él mismo. No hay razón para dudar de nuestra capacidad para lograr tal cambio, de la misma manera que tampoco existe para poner en duda nuestra actual capacidad de destruir toda la vida orgánica de la Tierra. La única cuestión que se plantea es si queremos o no emplear nuestros conocimientos científicos y técnicos en este sentido”. (Arendt, 2018, p. 15)

De esta manera, la educación con *inteligencia artificial* se constituía como un modo de autoproducción de la inteligencia. La educación con *inteligencia artificial* podría inscribirse incluso conceptivamente como uno de esos modos de realización, a través de la tecnología, de los más íntimos sueños metafísicos del hombre. El ser humano desearía librarse de esta prisión terrena a través de la sustitución de sí por algo producido por él mismo.

Asimismo, el sueño de autoproducción que encarnaría la *inteligencia artificial* en su proyección inicial tendría su correlato en los sueños transhumanistas. ¿Qué podemos decir del deseo de transportar el contenido mental a un ordenador con el propósito de lograr la inmortalidad? (Kurtzweil, 2021).

Ya hemos visto, reiteradamente, como estos usos tecnológicos no se sitúan en un afuera del mundo, sino que, en su inmanencia, llegan a reproducir los sueños más típicamente humanos. Los dispositivos de identificación biométrica permitían el acercamiento a una situación de un gobierno capaz de establecer un control casi total de los movimientos de sus gobernados. El auge de la disciplina de ciencia de datos nos llevaba al ideal de acceso a la realidad nuda, a una objetividad suprahumana a través de la acumulación de datos, a una concepción *dataísta*.

Más aún, cada cultura tiene unos contenidos y fines distintos, hecho que se traduce de manera crucial en los variados cauces de la historia educativa (Negrín Fajardo y Vergara Ciordia, 2009). Educar es encauzar a alguien en una dirección determinada. La tecnología es una herramienta que debería prestarse a la realización de unos determinados fines humanos. ¿Qué ocurriría con una educación que se desprendiera por completo de la búsqueda de las finalidades en una mera inclinación hacia la perpetuación de la tecnología? ¿Adonde nos llevaría esa tecnología? ¿Cómo se articulan tecnología y pedagogía para el tema que nos corresponde? ¿No se atisba acaso que ambos campos nos remiten de manera necesaria al problema de las humanidades?

En el apartado siguiente plantearemos una discusión en torno a las figuras predecesoras del actual proyecto de educación con *inteligencia artificial*. Con ello, no dejaremos de ver disonancias en las diferentes proyecciones educativas, así como en las bases teóricas de todos aquellos proyectos que nos han llevado al estado actual. Esto nos conducirá, no obstante, a una crítica unificada de ciertas notas comunes que atravesarían los diversos proyectos educativos en este ideal técnico plasmado.

6.1. Sobre máquinas de enseñanza-aprendizaje

En los apartados próximos vamos a atender a los antecedentes históricos de la aquella situación educativa con inteligencia artificial plasmada en el estado de la cuestión.

6.1.1. Sobre máquinas de enseñanza

En el año 1924 un psicólogo llamado Sidney Pressey presentó en Washington DC, en la American Psychological Association (APA), el modelo de una máquina de enseñanza especializada en administrar *test* de opción múltiple. Esta maqueta no tenía el carácter de mera herramienta puesto que, a su vez, se fundamentaba en el diagnóstico de una necesidad para las futuras generaciones:

Tiene que producirse una “revolución industrial” en la educación en la cual la ciencia educativa y la ingenuidad de la tecnología educacional se combinen para modernizar los severamente ineficientes y torpes procedimientos de la educación convencional¹⁰. (Pressey, 1944, p. 640)

La máquina en cuestión consistía en un método automático de registro de respuestas de *test* de opción múltiple. El aparato presentaba cuatro opciones ante la cual el individuo debía elegir una respuesta. A su vez, tenía dos modos de actuación.

En primer lugar, si elegíamos la opción examen o “test” la máquina registraba todas las respuestas de manera consecutiva y tras la finalización de la prueba otorgaba los resultados finales. Por el contrario, si elegíamos la opción “enseñanza”, entonces el artilugio no dejaba avanzar en las preguntas hasta que no diera una respuesta correcta. A su vez, en un determinado momento, Pressey incluyó en su diseño un dispensador de caramelos con la funcionalidad de liberar una golosina como recompensa una vez el alumno hubiera registrado un determinado número de respuestas correctas (Benjamin, 1988).

En 1926, en un artículo titulado “A Simple Apparatus Which Gives Tests and Scores- and Teaches” Pressey realizó una descripción del aparato para la revista *School and Society*. Ahí también se indicaba la finalidad del artilugio, a saber, se perseguía que los alumnos se concentraran en lo importante, esto es, en la obtención de unos resultados concretos. Del mismo modo, se prometía reducir la plantilla de profesores debido a la optimización de la enseñanza augurada.

¹⁰ La traducción es nuestra.

No obstante, el modelo de Pressey no logró imponerse en las aulas y cayó en el olvido. Puede pensarse que estas máquinas no triunfaron debido a la época en la cual fueron producidas y promocionadas. Las máquinas de enseñanza de Pressey se trataron de popularizar en Estados Unidos en la década de los años 30, un periodo convulso afectado por la Gran Depresión y en el cual el gobierno no estaba dispuesto a aumentar gastos en educación al tiempo que había exceso de profesores (Benjamin, 1988).

No obstante, años más tarde, con principal incidencia en la década de los 50 y de los 60, el célebre psicólogo conductista Burrhus Frederic Skinner, reformularía este modelo incidiendo en este proyecto de automatización de la enseñanza con fundamento en sus teorías sobre el aprendizaje.

De manera evidente, Skinner valoraba el uso del refuerzo que realizaban las máquinas de Pressey pero, sin embargo, a juicio del primero, era necesario que estas máquinas otorgaran nueva información a los estudiantes. Del mismo modo, Skinner rechazaba el *test* de opción múltiple (Aguayo, 2016). A su modo de ver, este tipo de examen caería en el riesgo de confundir a los estudiantes. Su contrapropuesta residía en abogar por el hecho de que los estudiantes construyeran sus propias respuestas para dar con la correcta. A su vez, si el alumno daba con la respuesta correcta, la conducta entonces era reforzada.

Esta serie de intuiciones de Skinner, quedaron plasmadas en libros como *Tecnología de la Enseñanza* del año 1968. En sus propuestas se vislumbraba la idea de una enseñanza programada basada en principios conductistas, como el uso del refuerzo, y en la propuesta de un aprendizaje progresivo hacia mayores grados de complejidad. Se proponía, a su vez, que estas máquinas constituyeran un apoyo y no tanto sustitución de los profesores a cargo (Valero, 2016). De esta intuición nos quedan textos programados como el famoso *Análisis de conducta. Un manual programado* de Holland y Skinner.

En los años 60, las máquinas de enseñanza de Skinner se popularizaron y ocuparon la discusión pública. Al respecto, no podemos dejar de reseñar las notables semejanzas entre la discusión pública en torno a las máquinas de enseñanza de entonces y la que está aconteciendo ahora en el contexto de la enseñanza con *inteligencia artificial*.

Para empezar, tanto Pressey como Skinner abogaban por una maximización de la eficacia del proceso educativo a través del uso de la tecnología. Asimismo, ambos hacen incidencia en el aumento de los resultados de la enseñanza. Esta concepción no está lejos

de aquella promesa de mejorar los resultados del aprendizaje a través del uso de las tecnologías de la *inteligencia artificial* (UNESCO, 2019). De igual modo, encuentra continuidad con las palabras del representante encargado en el reciente acto de inauguración del Aula del Futuro: “Lo que tenemos por delante es el desarrollo de un ecosistema de educación digital de alto rendimiento y mejora de competencias de todos los componentes del sistema educativo” (INTEF, 2021).

En segundo lugar, la prensa no dejó de elaborar todo un arsenal de argumentos a favor y en contra con respecto a la posible implantación de las máquinas de enseñar de Skinner en las aulas. Por un lado, unos periódicos y articulistas temían que las máquinas sustituyeran a los profesores mientras que otros argumentaban en contra del hecho puesto que sostenían que precisamente, las máquinas de Skinner podrían librar a éstos de las tareas más tediosas (Benjamin, 1988). Al respecto, se argumentaba que si la *inteligencia artificial* es capaz de sustituir tu labor como profesor entonces bien sustituido estás.

Por otro lado, otros argumentarían que las máquinas serían aprovechadas por el gobierno para aumentar la ratio de las aulas (Benjamin, 1988) mientras que otros, como el actual INTEF, trataban de reportar la satisfacción de los profesores que habrían usado estas herramientas en las aulas. En todo caso, Skinner y la actual UNESCO incidían en el hecho de que la incorporación de estas máquinas en las aulas otorgaría una atención más individualizada a los estudiantes. Del mismo modo, abundaban los enfoques críticos que ponían en cuestión que las personas pudieran ser enseñadas del mismo modo que las palomas de Skinner.

Lo cierto, es que, finalmente, estas máquinas cayeron en el desuso y terminaron por no incorporarse en el sistema educativo estadounidense. ¿Verdaderamente se logró aplicar esta eficiencia revolución industrial al ámbito educativo? ¿Quedaron sustituidos los profesores por máquinas? ¿A qué se debió este abandono? El proyecto de nuestra investigación es bastante modesto en comparación con las exigencias de las preguntas planteadas.

No obstante, debemos resaltar que esta experiencia quizás nos sirva para apreciar que el problema de la enseñanza no puede ser resuelto. La desaparición de las máquinas de aprendizaje puede ser considerada como una lección de la historia sobre una nueva pregunta que nos surge de nuevo ante el panorama de la cristalización de la *inteligencia artificial* en el plano educativo. ¿Puede la enseñanza automatizarse?

De manera preliminar, podríamos adelantar una crítica a los modelos de enseñanza propuestos por los artilugios descritos. El problema advertido es que, ambas propuestas configuran modos de enseñanza demasiado enfocados en la mera obtención de la respuesta correcta. Estas concepciones de la optimización de la enseñanza caen en el error de identificar el dinamismo del proceso educativo con una serie de resultados ponderables. A su vez, buscan de una manera, en cierto modo, tramposa que los resultados de aprendizaje coincidan con los parámetros de la experimentación científica. Con ello, parece reducirse la enseñanza a la mera aplicación de un modelo. Esta crítica encontraría apoyo en la observable cuantificación del proceso de enseñanza y la primacía otorgada al proceso de examen.

Con ello, podríamos decir que, en la pregunta por la optimización de la enseñanza, lo que hacen tanto la máquina de Skinner como la de Pressey es desplazar la pregunta, al identificar el proceso de enseñanza-aprendizaje con el examen y del mero enfoque en los resultados del proceso examinador concluir que se ha dado por mejorado el proceso de enseñanza a nivel global. Esta inferencia, por otro lado, no resulta demasiado lejanas a concepciones ciertamente actuales. Pero ¿es el proyecto técnico-educativo de Pressey y Skinner identificable por completo con el actual *boom* de la *inteligencia artificial*?

6.1.2. Sobre máquinas de aprendizaje

En unas fechas coetáneas empezó a tomar cuerpo un extraño proyecto que en el futuro amenazaría con cambiar el modelo productivo y educativo de la sociedad.

En el año 1951 un ingeniero y matemático llamado Marvin Minsky construyó aquello que podríamos denominar como “máquina de aprendizaje”. Este artífice dio a luz al primer simulador de redes neuronales el cual fue bautizado bajo el nombre de SNARC (*Stochastic Neural Analog Reinforcement Computer*). Como su nombre indica, esta máquina de aprendizaje se basaba en una modelización de comportamiento eléctrico de las neuronas en el cerebro.

SNARC trabaja como un cerebro vivo. Sus elementos eléctricos respondieron a las señales de manera muy similar a como lo hacen las neuronas del cerebro. La máquina recibe una tarea (en este caso, resolver un laberinto), pero, a diferencia de una computadora, no recibe un programa que le indicará cómo realizarla, en cambio, las neuronas artificiales se inician con conexiones aleatorias. Sin embargo, si una conexión particular acercaba a la máquina a su objetivo, la misma se "refuerza" (dando un valor más alto que hacía más probable que persistiera). Gradualmente, se forma una red de tales

conexiones reforzadas, lo que permite a SNARC realizar su tarea. En otras palabras, SNARC "aprende" cómo hacer algo a pesar de que solo tiene unos cientos de tubos de vacío en lugar de los millones de neuronas del cerebro. (Fajardo de Andara, 2021, p. 43).

Este proyecto, en cierta manera, parecía retomar las intuiciones iniciales de Alan Turing quien había propuesto que para el logro de una verdadera *inteligencia artificial* la vía más adecuada sería la construcción de una *máquina infantil*. La *máquina infantil* sería un dispositivo con la capacidad de aprender al igual que un niño. Si lográramos esto, a través del mero entrenamiento de la máquina ella misma podría lograr el desarrollo de una inteligencia similar e incluso superior a la de un humano adulto.

No obstante, este prototipo de *máquina infantil* ideado por Marvin Minsky tuvo que hacerse a costa del abandono de la lógica computacional simbólica propia de las primeras averiguaciones de Alan Turing.

Mientras que la máquina de Turing habría de guiarse por un modelo de cálculo serial, esto es, de una serie de operaciones sintácticas secuenciales, la SNARC trataría de imitar la lógica asociativa del cerebro humano. El tratamiento de información del humano se construye primariamente bajo la formación de conexiones neuronales que se van consolidando y dando lugar a determinadas respuestas. Este ejercicio de mimesis sería lo que, al fin y al cabo, permitiría el "aprendizaje" en este tipo de dispositivos.

Para realizar esta imitación, la SNARC empleaba aquello que se conoce como arquitectura conexionista de computación. La arquitectura conexionista funciona a través de la creación de nodos, los cuales son receptores eléctricos capaces de enviar respuestas a otros nodos según se hayan alcanzado unos determinados valores cuantitativos. Los valores alcanzados y las determinadas transmisiones en consecuencia van configurando aquello que se conoce como "red neuronal".

Estas redes neuronales, a su vez, funcionan en paralelo. De este modo, la máquina va adquiriendo una estructura determinada según los *inputs* recibidos y, por ello, podemos decir que aprende (Sanguineti, 2008). Esto, a nuestro juicio, representaría a pequeña escala la plasticidad del cerebro la cual ha cobrado una especial relevancia en la reflexión neuroeducativa. Los circuitos neuronales siempre estarían configurándose y su propia estructura variaría según las experiencias y aprendizajes consolidados (Béjar, 2016).

De este modo, este modelo "primitivo", consiguió una primera aproximación a la obtención de lo que sería un modelo de aprendizaje de nuestro cerebro. A pesar de todo,

cabe preguntarse si esta analogía entre circuito cerebral y red neuronal artificial es de algún modo perfecta.

Psicólogos como Manuel de Vega no dejan de reseñar la esencial diferencia que existe entre el modelo computacional de la inteligencia y la inteligencia humana. Tanto es así que, “el sujeto en el razonamiento práctico se apoya en su conocimiento temático, en convenciones lingüísticas (no lógicas) y en heurísticas o reglas de ‘andar por casa’”. (García y de Vega, 1988, p. 7). Nuestro modo de razonar, inscrito en diversas *formas de vida*, no seguiría de manera estricta el modelo de raciocinio que propondría la inteligencia artificial pues “los conceptos, o categorías naturales del hombre de la calle, no son bien definidos”. (*Ibid*, p. 6.).

De igual modo, la arquitectura conexionista necesita tener unos valores fijados de antemano para poder iniciar la circularidad de sus procesos. ¿Puede el ser humano programar un dispositivo capaz de aprender en una multiplicidad de contextos, realizar todas esas asociaciones que configuran nuestra *forma de vida*? De momento, la respuesta histórica es escéptica: todavía no ha podido configurarse una *inteligencia general* que, para nuestro caso, sea similar a la inteligencia de un profesor.

Como es conocido, esta máquina de aprendizaje, prototipo de lo que posteriormente sería el *deep learning* estaría siendo aplicada actualmente en el ámbito educativo. Un dispositivo que tomaría como modelo nuestra inteligencia estaría así pujando por entrenar nuestra inteligencia-modelo. Solo la historia determinará si aquellas máquinas de enseñanza-aprendizaje adquirirán un rostro humano.

En consecuencia, a raíz de nuestra exposición sobre las “máquinas de enseñanza-aprendizaje” podríamos concluir que la historia parece haber probado en el pasado que el proceso educativo se resiste a ser automatizado. De igual modo, apreciamos que también los procesos históricos suelen defraudar a los profetas de prensa.

No obstante, a nuestro juicio, en la experiencia actual, la historia no parece haberse repetido de una manera idéntica. Las oportunidades ofrecidas por la inteligencia artificial parecen no ser identificables ni mucho menos con las posibilidades que planteaba la máquina de enseñar de Skinner. Al respecto, encontramos entre la *inteligencia artificial* y estos lejanos prototipos dos diferencias fundamentales.

Para empezar, las herramientas de la inteligencia artificial presentan un grado de adaptación al estudiante sin precedentes. “La máquina de aprender” permite una personalización de la experiencia de aprendizaje que nunca logró “la máquina de enseñar”. A continuación, a nuestro juicio, la inteligencia artificial ha alcanzado un desarrollo técnico tan grande que permite una pluralidad de fines educativos insospechada. No obstante, a pesar de todo, sí que podrían localizarse sesgos específicos que englobarían estas técnicas.

Estas dos diferencias serán abordadas en los apartados siguientes.

6.2. Una pluralidad técnica ante una pluralidad de fines

En la década de los 60, en Estados Unidos el ingeniero Marvin Minsky emprendió un proyecto colaborativo con uno de los mejores discípulos de Piaget, el psicólogo Papert. Este proyecto sincronizaba las teorías sobre el aprendizaje de Papert con las inquietudes sobre el conocimiento del cerebro infantil de Minsky en pos de su modelización y computerización. Juntos, desarrollaron un lenguaje de programación para infantes denominado *Logo*. Se entendía así que, si se comprendiera mejor el fenómeno del aprendizaje, este podría ser aplicado con éxito a las tecnologías.

La aplicación perseguía que los alumnos aprendieran a programar dando comandos a una tortuga robot física que, por cuestiones de presupuesto terminaría siendo virtual (Solomon y Papert, 1973). El proyecto educativo se fundamentaba en la teoría del aprendizaje de Papert, el construccionismo. Según ésta el aprendizaje tendría lugar constructivamente, a través de la interacción del educando con el medio. El conocimiento, de esta manera, quedaría entendido como el producto del trabajo que uno mismo hace (Obaya, 2003).

Como puede verse, el correlato educativo del desarrollo técnico de Minsky se tradujo en una concepción de la enseñanza constructivista que actualmente, como vimos, tendría su presencia en los ELEs y, a nuestro juicio, continuidad en un reciente proyecto del ministerio con el propósito de crear una asignatura que desarrolle el pensamiento computacional.

En este último proyecto, el cual nos vemos en la obligación de reseñar, se observa una presencia de principios educativos similares al proyecto de Minsky y de Papert. Ello queda patente, por ejemplo, en el postulado según el cual “se aprende haciendo” así como en la sugerencia de dejar al alumno “trastear” con la aplicación antes de enseñarle a usarla

(INTEF, 2022, P.73). De igual manera, en ambos proyectos se observa una proyección a desarrollar la competencia específica del pensamiento computacional la cual quedaría definida como “habilidad que permite resolver problemas y comunicar ideas aprovechando la potencia que ofrecen los ordenadores” (INTEF, 2022, p. 10).

No obstante, en este punto de la exposición, debemos pararnos a reflexionar si acaso todas las aplicaciones de inteligencia artificial son herederas de esta concepción papertiana de la enseñanza. Como ya hemos anticipado, de ningún modo esto es así.

Atendemos a que, por ejemplo, en plataformas online para el autoaprendizaje como los MOOC presentes en Coursera.org, al igual que las máquinas de enseñanza de Skinner y de Pressey, se estaría respetando una estructura canónica de transmisión del saber. La concepción normativa quedaría plasmada en los propios fines educativos de la plataforma. Según el diseño de ésta, el reto consistiría en ir adquiriendo una serie de aprendizajes prefijados cuyo contenido se hallaría en una serie de lecciones magistrales pregrabadas. Las destrezas, asimismo, serían comprobadas a través de la superación de diversos exámenes de evaluación automática. Si todas las pruebas son superadas se produce la expedición de un certificado de manera instantánea.

Del mismo modo, en los sistemas de tutorización basados en el diálogo veríamos de manera nítida el seguimiento de un modelo de enseñanza iniciativo que a menudo suele ser entendido como socrático. El papel de la inteligencia artificial es la de llevar al alumno a adquirir ciertos conocimientos con un cierto grado de flexibilidad. El tutor “artificial” lo que hace es guiar al alumno para que llegue por sí mismo al conocimiento, fomentando así una implicación activa con la enseñanza.

Pues bien, esta característica de la realidad de las herramientas educativas con *inteligencia artificial* mencionadas como humanistas o como filósofos que somos no puede dejar de llamarnos la atención. El hecho de que la inteligencia artificial tome tantos caminos distintos nos hace pensar en el papel que juega nuestra disciplina. La tecnología, hablando claro, no “va sola” sino que tiene unas finalidades humanas atribuidas.

¿Qué concepción del ser humano manejamos para el desarrollo de estas tecnologías? ¿Qué ciudadano político queremos construir a través de la escuela? ¿qué valores queremos transmitir? ¿queremos usar estas herramientas para realizar una necesaria inclusión? ¿dónde situaremos el límite entre la personalización y la

segregación¹¹? ¿usaremos estas tecnologías para hacer más accesible el conocimiento en lugares remotos? ¿servirán estas tecnologías precisamente para ahorrar costes al gobierno en educación recortando el número de profesores en las aulas y aumentando la ratio?

Asimismo, una cosa es el enfoque que adopta la máquina y otra el enfoque que adopta el profesor humano. Al respecto, no podemos dejar de preocuparnos por si acaso el profesor tendrá opción de elegir si aplicar o no estas tecnologías. ¿Tendría que lograrse una sincronía entre el estilo docente del humano y de la inteligencia artificial?

Las preguntas al respecto son numerosas. La incertidumbre que se deriva de la pluralidad de fines atribuibles a la *inteligencia artificial* al mismo tiempo conforta. Sin embargo, no podemos dejar de preguntarnos si acaso este tipo de dispositivos presenta algún sesgo específico transversal a todas las diversas inspiraciones pedagógicas que puede adquirir. ¿Qué implicaciones tiene que estas aplicaciones funcionen a través de la recopilación masiva de datos? ¿No hay en este proceder datable y por tanto cuantificador un carácter antagónico con respecto a la enseñanza de las humanidades?

6.3. Aspectos de resistencia computacional de la enseñanza de las humanidades

En la pregunta por el problema de la aplicación educativa de la inteligencia artificial hemos visto diversos caminos y, con ello, diversos modelos de enseñanza que han sido emprendidos por la pluralidad de herramientas. El camino de desarrollo de esta técnica no está del todo claro.

Veámos así que los tres modelos de enseñanza (iniciativo, normativo y constructivista) quedaban igualmente contemplados en las diversas herramientas analizadas. No obstante, a nuestro juicio cada una de las concepciones en cada aplicación de la I.A., en lo correspondiente a la enseñanza de las humanidades, presentaría un cierto grado de resistencia que, cuanto menos, merece ser contemplado.

En primer lugar, las herramientas que utilizan un enfoque normativo presentan la contrapartida de que en ellas se produciría una falta de comunicatividad, no adaptando el nivel del contenido (magistral) al alumnado y presentando al mismo tiempo un déficit de

¹¹ Esta pregunta fue planteada por otro compañero del máster en educación en una clase a propósito de la inteligencia artificial en educación.

diálogo lo cual se traduce en un menor desarrollo crítico, competencia imprescindible para el ámbito humanístico.

En segundo lugar, con respecto a las herramientas que adoptan un enfoque iniciativo, estimamos que las mismas todavía no han alcanzado el nivel requerido de desarrollo del procesador del lenguaje natural. Del mismo modo, se enfrentan a la problemática de tan solo atender a la formalidad de lo dicho, careciendo de la capacidad de comprensión del significado. Asimismo, la guía específica al alumnado está condicionada al enorme grado de abstracción y homogeneización que suponen los dispositivos de personalización inteligentes, a saber, los modelos predictivos basados en una recopilación masiva de datos.

En último lugar, con respecto a las actividades que adoptan un enfoque constructivista no dejamos de reseñar que estas presentan un claro enfoque al ámbito de las Matemáticas y de la Computación. Al respecto, nos mantendríamos escépticos pero abiertos a la posibilidad de enseñanza de humanidades a través de entornos exploratorios de aprendizaje. Pensamos que quizás podría diseñarse un paisaje efectivo para plantear la “resolución” de problemas éticos, pero no creemos que estos puedan suplir la labor total del profesor humanista. En todo caso el diseño de estas herramientas debería tener un carácter *ad hoc*. Pero las competencias humanísticas específicas a nuestro juicio no pueden florecer en un entorno de mero procesamiento de la información.

En consecuencia, a través del estudio de estas herramientas llegamos a la conclusión de que el estudio y la enseñanza de las humanidades parece configurarse como un valorado excedente al paulatino proceso de automatización de las disciplinas. Hay elementos propios, pero quizás no particulares, de la disciplina que se resisten a este tipo de automatismo. Por eso, quizás desde el distanciamiento con respecto a lo que ya existe el pensamiento humanista puede llegar a altas cotas de desarrollo.

El dataísmo evaluador

Para empezar, podríamos pensar que un aspecto fundamental de resistencia de las humanidades con respecto a la disciplina de la *ciencia de datos* sería el método evaluativo. La *inteligencia artificial* funciona esencialmente desde la cuantificación, desde la conversión de las cosas percibidas en datos. A través de la acumulación de datos genera modelos en base a los cuales toma decisiones y, por tanto, califica. ¿De verdad puede evaluarse el pensamiento humanístico basándose en modelos predictivos?

Las humanidades asimismo presentan un tipo de discurso que se asemeja a la problemática de los *datos no estructurados*. Su propio tratamiento se resiste a la categorización en filas y en columnas. La transformación de esta en parámetros estandarizables está siempre sometida a un equilibrio muy precario.

¿Cómo un algoritmo diferenciará si un alumno sabe encontrar el nexo causal a los hechos históricos estudiados y no tan solo enuncia una sucesión de datos históricos? ¿Cómo va a evaluar correctamente un ensayo filosófico un corrector con inteligencia artificial? ¿Puede evaluarnos un dispositivo que no comprenda los conceptos a los cuales estamos aludiendo?

Del mismo modo, el modo evaluativo de las humanidades parece consistir esencialmente en la posibilidad de apreciación del matiz. La coyuntura que nos impone el obligado tratamiento de los datos contra la flexibilidad de la evaluación del profesor humanista se asemeja a la diferencia que existiría entre las variables continuas y las variables discretas.

Una variable discreta es aquella que no pueden tomar ningún valor entre dos consecutivos (del Rio, 2019). Un ejemplo de este tipo de variable sería el número de hermanos. Una persona puede tener dos o tres hermanos, pero, en todo caso, no podrá tener dos hermanos y medio. Por otro lado, una variable continua es aquella que puede tomar cualquier valor dentro de un intervalo. Por ejemplo, podemos pensar en la distancia que existiría de un punto a otro.

Pues bien, el modo de evaluar, así como de enfrentarse a los contenidos de los procedimientos *datables* parece consistir en la imposición de un valor determinado sin la posibilidad de situarse en ningún punto intermedio entre dos polos. Si evaluamos a través del mero registro es posible que se equiparen conductas de aprendizaje del todo punto antagónicas. La mera atención a los requisitos formales implica una pérdida inmensa derivada de la no atención al complejo crisol que conforman los significados.

Más aún ¿puede realizar un trabajo de evaluación un dispositivo que no tenga conciencia de los plexos históricos y sociales de la evaluación?

Autores como Philippe Perrenoud no dejan de reseñar el condicionamiento social y cultural de los dispositivos de evaluación. En la configuración de los programas educativos ya se contendría, de hecho, una decisión de elogiar o excluir y castigar a

determinados alumnos. Por lo cual, aquello que determinamos como excelente o fracasado es algo cultural e históricamente determinado (Perrenoud, 2008).

Dicho lo cual, ¿De verdad puede evaluarse sin una conciencia crítica de lo que es evaluar? ¿No por ello precisamente decimos que el propósito de la evaluación es propiamente formativo?

Contemplando que la evaluación es formativa y es parte integrante de un proceso de socialización secundaria como es la enseñanza, no queda sino cuestionar su posible carácter de objetividad. La evaluación forma parte de un conjunto de interacciones humanas y de juegos de significados que los programadores no han podido todavía datificar de manera íntegra.

Los dispositivos requieren de una conversión y recolección masiva de la realidad en datos. Su ontología funciona meramente en estos términos. Pueden recopilar gestos de atención, tiempo atencional en la pantalla, puntuaciones en determinados exámenes. Es capaz de calcular el tiempo dedicado a cada tema y hacer estadísticas del progreso del alumnado. Podrían dar resultados predictivos de la nota que se obtendrá en el examen si se sigue con este modo de aproximación a la asignatura.

Pero ¿qué ocurre con las mentiras piadosas de los profesores? ¿no hay en la enseñanza precisamente un juego de mentiras necesarias que ayuda a los alumnos a prosperar y a consolidar los aprendizajes? Podríamos hablar por de la estrategia de realizar correcciones más severas en el primer parcial. O, por ejemplo, de hacer creer a un alumno que puede aprobar cuando con el nivel demostrado hasta entonces todos los modelos predictivos imaginados hubieran sentenciado que éste no sería capaz de ningún modo de aprobar la asignatura.

En el estudio con modelos predictivos de notas podemos imaginarnos fácilmente una situación en la cual el estudiante viviera a crédito. El alumno ofrecería datos a la máquina para que esta hiciera pudiera elaborar estadísticas sobre su progreso. Trataría con todas sus fuerzas obtener una buena estadística para posteriormente probar en la realización del examen efectivo que su progreso así ha transcurrido. En ese caso, el examen de hecho no se postularía como mero resultado del esfuerzo, como mera foto fija del dinámico proceso del aprendizaje, sino que, por el contrario, parecería una suerte de deuda pendiente, de probación efectiva de la profecía algorítmica. ¿De verdad estudiar bajo la amenaza de una nota puede llevar a aprender?

La particular mirada de los algoritmos reproduce un tipo de enfoque demasiado centrado en lo cuantitativo. A su vez, tiene unos compromisos ontológicos muy determinados que no tienen por qué coincidir con los del profesor.

El profesor no mira al alumno y ve números, si somos bien pensados. Sin embargo, existe un riesgo real de que el profesor delegue su función evaluadora en los dispositivos mencionados. Si el profesor no es consciente de que el algoritmo se encuentra situado, puede recaer en una concepción dataísta y entender que con la acumulación de datos del alumno se lograría una superior nitidez sobre su progreso en la asignatura. El docente no puede delegar su factor de conocimiento pues, como señala Han en *La salvación de lo bello*, en el dataísmo:

“El hombre debe regirse por los datos. Abdica como productor del saber y entrega su soberanía a los datos”. (...) “Ahora el saber es producido maquinalmente”. (Han, 2016, p. 66).

De igual manera que ocurría con su antecedente, la máquina de enseñar de Pressey y de Skinner, la función de la enseñanza parecería entonces remitirse al examen, hacia el mero carácter demostrativo o, como parece apuntar el informe de la UNESCO, hacia el vano incremento de los resultados del aprendizaje.

No obstante, quizás la aplicación masiva de estas tecnologías lleve a un replanteamiento de los habituales modos de evaluación. Ello podría ser debido a que la inteligencia artificial paradójicamente estaría automatizando precisamente aquello que habitualmente suele ser evaluado en un examen (Holmes et al, 2019). Quizás este factor pueda llevarnos a un replanteamiento de los modos clásicos de evaluación. Como advierte Perrenoud:

“El sistema clásico de evaluación fuerza a los docentes a preferir los conocimientos aislables y clasificables numéricamente frente a las competencias de alto nivel (razonamiento, comunicación) difíciles de encerrar en una prueba escrita y en tareas individuales”. (Perrenoud, 2008, p. 87)

El panorama, de nuevo, no está claro. La incorporación de la inteligencia artificial en los dispositivos de evaluación podría llevarnos tanto a una reformulación de la tradicional evaluación cuantitativa como a una pérdida de todo aquello que no sea ponderable a través de la ciencia de datos.

El procesador humano

Si recordamos nuestro estado de la cuestión podremos recuperar la imagen de aquel procesador humano que era planteado por Newell. El ser humano era planteado, al igual que el ordenador, como un sistema de procesamiento de la información.

A nuestro juicio filosófico, esta propuesta no puede ser compartida. Para nosotros la información no es algo directamente disponible a los sentidos. Aquello que entendemos como mera pieza de información, en su atemporalidad y asepsia es algo ya culturalmente construido. Consideramos que, el procesar de los sistemas no se parece al comprender ni al proceso creativo de la memoria que documentan las fuentes neuroeducativas (Mora, 2017).

La ausencia de comprensión de la *inteligencia artificial*, de hecho, se manifiesta en la constante recaída de la misma sobre el plano de la literalidad. Ésta no alcanzaría a recrear una comprensión del significado sino una mera transposición de signos.

Para ilustrar este aspecto advertido, vamos a realizar una breve reseña de una proeza técnica recientemente constituida. El año pasado, la plataforma *Netflix Is A Joke* publicó una serie de cortos de diversos géneros cuya peculiaridad residía en haber sido escritos de manera íntegra por dispositivos de inteligencia artificial.

En el caso reseñado, tras el visionado de más de cuatrocientos mil horas de archivos sobre el género de comedia *stand up*, se dejó al mecanismo componer su propia comedia basándose en los patrones recolectados. La inteligencia artificial llegó a constituir chistes como el siguiente:

“Internet mad at me. Well, I’m mad at internet. I accept cookies but received no pastry treats. Web is so fat it is worldwide. Maybe put away cookies and eat cyber salad¹².” (Netflix Is A Joke, 2022)

¹² Hemos decidido conservar el fragmento original en el cuerpo del texto para no alterar el sentido humorístico de la *inteligencia artificial*. En cualquier caso, adjuntamos a continuación una traducción de este.

“Internet está enfadado conmigo. Bueno, más bien estoy enfadado con internet. Acepto *cookies*, pero no recibo dulces de pastelería. La red está tan gorda que está en todo el mundo. Quizás guarde las *cookies* y coma ensalada cibernética.”

El monólogo íntegro está disponible en el enlace siguiente: <https://www.youtube.com/watch?v=j-zTrfiQMSY>

Como puede apreciarse, el juego de la *inteligencia artificial* consiste en la imitación de la formalidad del chiste, siendo los juegos de humor preferidos aquellos que seleccionan la polisemia de un término y dan una connotación descontextualizada de éste. En otras palabras, la principal gracia de la *inteligencia artificial* proviene del sinsentido que permite la vinculación entre palabras similares fuera de su contexto.

En este caso, el algoritmo está “entreteniéndose” con la polisemia del término “cookies”. De igual manera, también observamos que la *inteligencia artificial* es capaz de jugar con la literalidad de los signos con proximidad fonética. En un determinado momento del monólogo dice a un miembro del público “i think you should leaves” tras lo cual el personaje pasa a convertirse en hojas.

En cualquier caso, merece ser tomada en cuenta la brillantez de los programadores. Estos han sido capaces de aprovechar el déficit de la inteligencia artificial, a saber, la carencia de comprensión y la recaída sobre el plano de la literalidad para hacer recaer éstas sobre el plano humorístico.

De este modo, podemos ver que la inteligencia artificial tiene problemas de comprensión y, en todo caso, una capacidad de *transferencia* muy limitada, si es que podemos decir que existe. Los campos de relaciones se establecen entre similitudes fonéticas, y poco más, tras lo cual se establecen juegos de sentido. La inteligencia artificial es capaz de vincular referencias y hacer gracias con la literalidad.

Sin embargo, la transferencia de nuestros alumnos es mucho más amplia. Su capacidad de trasposición de unos aprendizajes a otros contextos no se limita a una aplicación descontextualizada de un mismo significante. Los alumnos no viven necesariamente en la ironía del sinsentido.

La comprensión humana, de esta manera, implica el recurso a las facultades de *transferencia* y *relevancia*. La facultad de *transferencia*, como ya hemos señalado, consiste en la capacidad de trasponer unos aprendizajes adquiridos en un determinado contexto a otros de naturalezas distintas. Por otro lado, la facultad de *relevancia* implica la posibilidad de seleccionar lo importante en un campo abundante de información y de contenidos. Estas competencias suelen ser atribuidas a lo propiamente humano, a lo más característico de una *inteligencia general*.

Al respecto, podemos señalar que encontramos ambas estrechamente relacionadas. Ello es debido a que la *relevancia* o selección de lo importante implicaría el reconocimiento de una finalidad, de saber seleccionar lo que nos es adecuado para un determinado propósito. Además, vemos lo preciado que es la transformación de un conocimiento a través de su aplicación a otro campo. Podemos entender así que la cualidad de la *transferencia* del conocimiento está muy vinculada con la innovación.

Vemos estas facultades puestas en marcha de una manera especial en aquellas famosas “preguntas de pensar” propias de las humanidades. Son habilidades requeridas para la constitución de un ensayo, para la elaboración de un adecuado comentario de texto e incluso para ser capaz de manejarse en la red sin ser engullido por el exceso informativo característico de la misma.

Para realizar una argumentación, necesitamos jerarquizar la información y establecer relaciones que rebasan la intención comunicativa inicial. Para constituir un comentario de texto hemos de ser capaces de seleccionar las ideas principales y de lograr conexiones con lo expresado para aumentar la comprensión, clarificar lo escrito. Para navegar en la red con provecho hemos de saber encontrar lo que estamos buscando, conocer conceptos análogos e incluso ideas que relacionen nuestra búsqueda.

De hecho, quizás, precisamente, incidiendo en el aspecto innovador que implica la facultad de *trasferencia*, así como en el carácter sistemático que va aparejado a la selección de lo relevante, obtengamos una vacuna contra los dos típicos defectos algorítmicos que ya habíamos mencionado: las *cámaras de eco* y las *burbujas de filtro*. Transferir un conocimiento a otro lugar implica un reconocimiento de su carácter poliédrico, y por tanto una amplitud de la forma de mirar al mundo, cosa que no ocurre en las *burbujas de filtro*. Encontrar lo relevante implica además un distanciamiento crítico con respecto a la información disponible, más allá de la constante repetición de las *cámaras de eco*.

Quizás podemos replantearnos por qué estas facultades entonces son tan valoradas en la era actual y propuestas por algunos de ser reforzadas en el currículum (Holmes et al., 2019). Es posible que no se trate de que estas habilidades jamás podrán ser adquiridas por las máquinas de tal modo que mediante su reforzamiento nos convirtamos en insustituibles. Quizás simplemente se proponga su fortalecimiento porque a nosotros, personalmente, nos viene bien para movernos en un panorama regido por máquinas.

Si no queremos que la ciencia de datos totalice por completo nuestra manera de mirar del mundo, de cierta manera vamos a tener que incluir un tipo de reflexión alternativa que sostenga otra mirada sobre las cosas. Las cosas han de poder ser contempladas desde su historización y no desde la aséptica y atemporal información. Del mismo modo, han de poder ser atendidas desde la posibilidad de matiz, desde un carácter poliédrico como el que ofrece el arte y no desde un inmóvil y unívoco dato. Hemos de ser capaces de superar esta datificación de la educación y acercar a los alumnos hacia los plexos de sentido que constituyen al fin y al cabo el conocimiento.

El profesor robot

Podríamos pensar que la herramienta de inteligencia artificial no contiene en sí misma una reflexión sobre su propia aplicación. Este ámbito solo puede ser quizás logrado desde una suerte de distanciamiento con la misma a través de una reflexión propia de disciplinas como la Filosofía, la Literatura y la Historia.

En comunicados como los de la UNESCO (2019) vemos plasmados sueños de acabar con brechas educativas a través de la *inteligencia artificial*. También observamos el proyecto de aumentar los resultados de la enseñanza y del aprendizaje sin conocer del todo en qué consisten estos resultados. Percibimos la amenaza, incluso, de incluir de manera ineludible todas estas tecnologías por el hecho de haber llegado a un punto de no retorno.

Atendemos a que continuamente se habla de dar soluciones a problemas educativos a través de la *inteligencia artificial* pero no se definen de una manera exhaustiva los problemas. Se mencionan algunos como que ésta contribuya a solucionar dilemas éticos como mejorar la equidad a través de la personalización de la educación y facilitando el acceso a lugares remotos, así como facilitando la calidad de la enseñanza.

No obstante, no se ofrecen los contrapesos que pueden suponer la incorporación de todas estas tecnologías. Por ello, ¿no es necesario parar todas estas elucubraciones mediante una reflexión acerca de nuestras esperanzas al respecto? ¿No sería pertinente realizar un análisis pausado de este frenético avance tecnológico? Podría pensarse que estuviéramos encomendando a la *inteligencia artificial*, en su alteridad inhumana, una resolución de casi todos nuestros problemas. ¿Podemos permitirnos delegar de esta manera?

Podría pensarse si acaso los problemas educativos que la *inteligencia artificial* pretende resolver merecen la pena que sean resueltos. Es probable que al eliminar ciertos problemas también puedan eliminarse ciertas virtudes de la enseñanza. Por ejemplo, ¿qué ocurriría si solucionáramos el problema de la autoridad? Quizás ello implicaría aniquilar toda individualidad a los alumnos y por ello su inherente desafío a la figura al mando como es necesario a todo adolescente.

Hay problemas que quizás no merezca la pena resolver porque acabar con su propia existencia como problema sería acabar con la vida. Quizás precisamente enseñar consista en ser capaces de aceptar y tolerar la contradicción, así como la tensión perpetua por la continua realización y el progresivo perfeccionamiento de una tarea que jamás puede ser concluida.

Ser profesor en cierta manera puede suponer ser capaz de tolerar el vacío. ¿Puede una *inteligencia artificial* hacer esto?

Al mismo modo, podemos cuestionar este panorama sustitutivo a nivel global. ¿De verdad es rentable sustituir al ser humano por una máquina? El humano es un ser muy barato de producir y de mantener. No requiere unos materiales de producción demasiado extraños. Del mismo modo, aunque tenga que descansar, él solo es capaz de encontrar motivaciones intrínsecas que le mueven a la acción, no debe ser programado explícitamente para realizar su tarea. Con ello, dudamos ante el escenario planteado por Nick Bostrom en *Superinteligencia*. No creemos que en un futuro nos enfrentemos a un panorama de progresiva disminución de la población humana debido a su inutilidad frente a las máquinas.

Asimismo, el profesor humano constituye un tipo de *inteligencia general* que no ha sido logrado de ningún modo por una inteligencia artificial. El profesor suele ser capaz de relacionar los contenidos de maneras insospechadas, de moverse en los plexos de sentido.

De igual modo, debemos de pensar en qué representaría la figura del robot para nosotros, sea en versión física, con su *hardware* de robot o en una versión meramente computerizada y virtual como mero *software*. El robot, a nuestro juicio, y como parece vislumbrarse en aquella magistral novela de Ishiguro, *Klara y el sol*, parece representar una versión funcional de lo que sería el ser humano. Para nosotros, el robot consiste en una interpretación del ser humano que entiende al mismo como si estuviera concebido

para un plan específico. Es el logro final de la autoproducción del hombre tal y como queda patente en aquel fragmento de Hannah Arendt reproducido. El robot es, al fin y al cabo, una versión del hombre dispuesto con un sentido intrínseco y extrínseco a él. Es el hombre sin soportar aquel estado de indeterminación de tener que forjarse a sí mismo un destino.

¿Qué es un robot profesor? El robot profesor sería aquel profesor que no tiene problemas, que ya ha encontrado el modo de resolución de la enseñanza y que simplemente puede centrarse en el desempeño de una tarea.

El robot profesor no comprende, sino que tan solo ejecuta. El robot profesor podemos serlo todos, pero su sueño está prometido a través del uso de la tecnología. Sería la constitución de una *inteligencia general*, tal y como prometería el proyecto inicial de la disciplina de la *inteligencia artificial*. Pero esta inteligencia general tendría de antemano establecidas sus funciones, su sentido en la vida. Sería la constitución del ser humano sin ningún grado de indeterminación, en una optimización perfecta.

Ese sueño se plasma en las inquietudes transhumanistas, en todo aquel deseo por sustituir todo aquello que nos ha sido dado por algo producido por nosotros mismos, algo cuyo sentido nos pertenezca y no sea una construcción de bases de arena movediza que en cualquier momento puede quedar en el pasado y pasar al olvido. El sueño del profesor robot puede consistir en la búsqueda de algo estable, que no cambie con el tiempo. Algo concluido, algo así como un fin de la historia. Una resolución definitiva al problema de la enseñanza.

No obstante, debido a las diferencias existentes entre la inteligencia humana y la *inteligencia artificial*, estimamos que la labor del profesorado no puede ser sustituida por el *software*. A nuestro juicio, solo deberíamos optar a tener unos buenos asistentes artificiales que complementen nuestra labor, pero no a profesores artificiales.

Usando unos parámetros wittgenstenianos que quedan expresados de manera nítida en la magnífica obra *El quinteto de Cambridge*, podríamos plantear que la inteligencia lo que hace, en un sentido general, es servir a un modo de vida. Por ello, no tiene sentido hablar de la inteligencia de las máquinas.

Por ejemplo, si pensamos en la existencia de vida inteligente extraterrestre, quizás el uso de la palabra “inteligencia” constituye lo que en Filosofía se conoce como *error*

categorial. Esto es debido a que la inteligencia en cierto modo parece consistir en estar proyectada a fines. No podríamos imaginarnos en aquello que consistiría la inteligencia alienígena porque no conocemos el modo de vida de los alienígenas. No sabemos cómo serviría ese dispositivo que nosotros llamamos “inteligencia” a sus usos sociales. Por ello, quizás ni siquiera lo podríamos conceptualizar.

La inteligencia de la máquina, podríamos decir, y tomando en cuenta la definición conductista de inteligencia que toma cuerpo en la prueba de Turing, sería semejante al grado de comprensión que tiene un alumno al copiar los apuntes de otro para hacer un examen. Asimismo, el robot no tiene sociabilidad, no puede desarrollar un modo de vida en el cual tome cuerpo su inteligencia. El robot tan solo realiza una serie de cálculos muy complejos que debido a su rapidez nosotros estimamos y extraemos una especie de “magia” que puede equiparar a nuestra inteligencia.

De igual modo, como hemos apreciado, esta magia puede trasladarse a una suerte de búsqueda de objetividad en los criterios evaluativos. Pero la enseñanza no puede ser automatizada. La enseñanza tiene que ver con una serie de lazos sociales. Y esto es algo que solo otros seres humanos pueden hacer.

6.4. ¿Humanidades?

La noción actual de humanidades es un producto histórico. Podríamos encontrar una cierta continuidad con la misma en el proyecto de enseñanza de las artes liberales emprendido por Varrón en el siglo I a.C. (Vita Proupech, 2004). No obstante, autores como Savater, no dejan de señalar el origen renacentista del término, aunque se advierte en el mismo una esencial divergencia con la concepción vigente.

La designación humanidades, en el Renacimiento hacía referencia a la disciplina cultivada por aquellos estudiosos que se enfocaban en textos de origen humano y no divino (Savater, 2010). La finalidad del humanismo renacentista se regía por la búsqueda de una formación holística del ser humano y, por ello, los *Elementos de geometría* de Euclides constituían una formación humanística en la misma medida que *El Banquete* de Platón (*ibid.*)

El término humanidades ya no alude a una concepción renacentista. Por el contrario, actualmente, cuando nos referimos a las humanidades nos referimos a toda esa serie de disciplinas que en el siglo XIX se separaron de las ciencias naturales y de la física (Rojas, Castro, 2013). Humanidades hoy son disciplinas como la filosofía, las ciencias

sociales, las artes, la literatura y la lingüística y la antropología (Cordua, 2012). Estas se diferencian por aportar una formación y metodología específica con respecto a las ciencias en un sentido estricto.

Las competencias esencialmente atribuidas a esta serie de disciplinas suelen estar constituidas por facultades como el pensamiento crítico, las habilidades comunicativas, la aproximación creativa a los problemas. Es en este punto de la historia cuando determinadas organizaciones, ante el auge automatizador prometido por la *inteligencia artificial*, reclaman la necesidad de incorporar en el currículo un mayor peso a este tipo de habilidades. Así la UNESCO sostiene que:

“Un mundo impactado de manera creciente por la IA requiere una pedagogía que, más que enfocarse a aquello en lo que las computadoras son buenas (memorización y computación), ponga el énfasis en habilidades propiamente humanas como pensamiento crítico, comunicación, colaboración y creatividad, y en desarrollar la habilidad de colaborar las herramientas de IA omnipresentes en la vida, el aprendizaje y el trabajo”. (Unesco, 2021).

No obstante, en este tipo de planteamientos advertimos un problema. Cuando se habla de humanidades rara vez se da una definición acerca de las mismas y, de hecho, en muchas ocasiones se utilizan de manera indistinta para hacer referencia a aquellas habilidades que serían competencia exclusiva del humano. Pareciera así que las disciplinas humanísticas se vinculan al problema de la *inteligencia artificial* por su proximidad fonética con la cuestión de lo humano.

Esto no puede dejar de llevarnos a sospecha. Cuando nos referimos a las humanidades en el contexto de auge de la *inteligencia artificial* ¿estamos remitiéndonos a una aliteración con función propagandística? Podríamos pensar que el uso de la palabra “humano” a la hora de hablar del auge tecnológico supone un recurso retórico con el propósito de evitar el terror ante la proyección que genera el imaginario ficticio de un futuro distópico derivado de la incorporación de estas técnicas.

De igual manera, podemos advertir que no está del todo claro que esta serie de habilidades atribuidas a las humanidades jamás puedan ser alcanzadas por una *inteligencia artificial*.

En los últimos diez años con la imposición generalizada del *deep learning*, esto es, de las *redes neuronales*, los dispositivos de inteligencia artificial están siendo capaces de llegar a notables grados de creatividad e incluso de realizar autónomamente labores de

programación. El más desafiante desarrollo de la inteligencia artificial parece consistir en aquello que se conoce como *foundational models*. En un análisis de este tipo de modelos, suele sostenerse que los algoritmos estarían siendo capaces de captar chistes e incluso de relacionar situaciones concretas con refranes populares (The Economist, 2022).

De igual manera, ¿no se encuentran este tipo de habilidades en disciplinas que no son propiamente humanísticas también? Las ya necesarias habilidades del manejo de datos requieren de la aplicación del pensamiento sistemático. Las demostraciones matemáticas precisan de la aplicación de la creatividad.

En este caso, ¿por qué parece haber un impulso romántico por parte de los humanistas hacia la delimitación de todas aquellas habilidades y competencias que serían infranqueables por una *inteligencia artificial*? ¿Estamos aprovechando el intimidante progreso técnico para autoafirmarnos como humanistas?

Pudiera parecer que los humanistas estuviéramos leyendo el auge de la *inteligencia artificial* como una victoria definitiva frente a los científicos. Pudiéramos estar pensando que, los científicos en su búsqueda del mejoramiento técnico perpetuo habrían logrado su propia obsolescencia. También pudiéramos creer así que entonces podríamos vivir en un panorama vital gobernado por ingenieros encargados de mantener las instancias y de humanistas con la posibilidad de dedicarse a mejores reflexiones.

Pero este enfoque nos resulta del todo punto irresponsable. En primer lugar, porque reproduce de manera cristalina el tópico de *las dos culturas*. Tal y como retrataría el físico y escritor Snow (1988) estaríamos creyendo en la existencia de una insalvable rivalidad e incomunicabilidad entre las disciplinas humanísticas y científicas. Por ello, en segundo lugar, estimamos que esta reflexión también resulta irresponsable por invitar a los humanistas a desatenderse de las cuestiones científicas.

Por tanto, a nuestro juicio, esa desatención ante las cuestiones técnicas por parte de los humanistas trae de manera implícita una confianza ciega en su propio funcionamiento y, en consecuencia, en una posible delegación sobre la *inteligencia artificial*. Si no ¿cómo iban a desatenderse de la reflexión sobre estas cuestiones? Si no nos preocupamos por conocer el funcionamiento de los dispositivos, en esta desatención de la ciencia, ¿no precisamente se nos impone la ciencia con una mayor fuerza?

En palabras de George Steiner:

Hasta que los estudiantes de humanidades no aprendan seriamente un poco de ciencia, hasta que la gente que estudia lenguas clásicas o literatura española no estudie también matemáticas, no estaremos preparando la mente humana para el mundo en que vivimos. Si no entendemos algo mejor el lenguaje de las ciencias no podemos entrar en los grandes debates que se avecinan. A los científicos les gustaría hablar con nosotros, pero nosotros no sabemos cómo escucharlos. Este es el problema. (Steiner, 2005).

Tal vez precisamente, el progreso técnico ya esté dando con una insalvable dimensión interdisciplinar del conocimiento. Son diversos los proyectos que actualmente abogan por una superación de las barreras existentes en las disciplinas. En el plano educativo podemos atender a proyectos de enseñanza multidisciplinarios como serían la STEAM (*Science, Technology, Engineering, Arts and Mathematics*). De igual forma, en el ámbito universitario podríamos apreciar la composición curricular de grados como el de *Diseño y desarrollo de videojuegos* en el cual los estudiantes tienen que profundizar en habilidades informáticas, lógicas y matemáticas, pero también de diseño y dibujo.

Más aún, en el plano de las humanidades, ya estaríamos, desde hace varios años, atendiendo a la proliferación de las *Humanidades digitales*. Este campo de conocimiento habría sido iniciado en 1974 por el jesuita Roberto Busa al haber hecho uso del cómputo para elaborar un índice de concordancias de las obras completas de santo Tomás de Aquino y autores relacionados (Galina Russell, 2011).

De este modo, las humanidades digitales estarían constituidas por un conjunto de disciplinas humanísticas en las cuales el investigador incorpora herramientas y análisis propios de las disciplinas computacionales. Por ello, el trabajo del humanista digital suele distinguirse por su carácter práctico, así como por una interpretación del mundo que cuenta con el uso de herramientas informáticas (Rojas Castro, 2013).

Por ello, cabe plantearse si acaso, al incidir en la recuperación de las humanidades no estaremos poniendo de manifiesto la necesidad de la recuperación de un enfoque interdisciplinar, que incluya dentro de sí a las humanidades como herramienta esencial, en esta venidera era de la *inteligencia artificial*.

Conclusiones

De acuerdo con lo expuesto, llegamos a las siguientes conclusiones:

- 1) La pregunta por el vínculo entre tecnología y pedagogía no nos devuelve una respuesta uniforme. Ello es debido a que la construcción de las herramientas se produce en orden a fines y la inteligencia artificial no es ajena a ello. Diferentes concepciones de la enseñanza coexisten en la diversidad de herramientas educativas inteligentes. Quizás, por ello, los humanistas debemos perseguir implicarnos de una manera significativa en la construcción de estas herramientas.
- 2) Existe una amenaza de que la concepción dataísta inunde el ejercicio de enseñanza-aprendizaje. Si así fuera, su alianza con los dispositivos de IA sería muy fuerte. Esto afectaría a los modos de evaluación y, en consecuencia, a la íntegra formación de nuestros estudiantes.
- 3) A pesar de haber cuestionado el propio estatuto de “las humanidades” sí que encontramos en estas disciplinas elementos específicos de resistencia con respecto al proceder de la ciencia de datos. Especialmente, apreciamos que las disciplinas humanísticas normalmente trabajan el factor comprensión, así como la posibilidad de movilidad en los plexos de sentido. Así, las humanidades no pueden computarse. Sin embargo, aceptamos de buen humor que el auge tecnológico presente hace imposible la existencia de una disciplina “pura” para enorme beneficio de las humanidades y las ciencias.
- 4) Nos hemos visto obligados a dejar de lado muchos temas relacionados de enorme relevancia. En especial, nos apena no haber podido profundizar en la pregunta por el inminente cambio del mercado de trabajo a causa de la incorporación de la IA a nivel industrial. Asimismo, lamentamos no haber podido contemplar como aquella cuestión podría influir en la configuración de la enseñanza.
- 5) Ni Pressey y Skinner lograron ponerse de acuerdo en las finalidades de un aparato tan rudimentario como la “máquina de enseñar”. Ante esta multiplicación de posibilidades, la reflexión humanista de profesores y alumnos resulta del todo punto necesaria.
- 6) La experiencia de las máquinas de enseñar nos lega la conclusión de que el problema de la enseñanza no puede ser resuelto, la enseñanza no es automatizable. Asimismo, nos proporciona una lección de sano escepticismo con respecto al furor causado por estos artilugios.

- 7) Existe un riesgo considerable de que el profesor pueda delegar en su función si quiere hacerlo. La I.A. sitúa a los profesores en una situación de enorme responsabilidad. Se puede simular ser profesor delegando a los *bots* la asignación de tareas, de resolución de dudas puntuales e incluso de la evaluación.
- 8) Si la concepción dataísta dominara podría imponerse una concepción de la enseñanza excesivamente centrada en el examen. El proceso educativo quedaría reducido a la mera demostración obviando el factor de socialización que entraña la escuela.
- 9) Consideramos que la figura del robot representa una versión funcional del humano. El robot en su literalidad no se enfrenta al problema del sentido, sino que directamente se encomienda a una tarea. En consecuencia, proponemos una elusión de la robotización de la enseñanza.
- 10) Quizás sería interesante llevar a cabo una planificación de la enseñanza más vinculada al factor comprensivo. Sin embargo, no ha de obviarse el factor memorístico que entraña la comprensión humana y viceversa pues, sostenemos que, el ser humano no es un procesador de la información.
- 11) Por último, apreciamos que el panorama actual ofrece un buen escenario para la superación de las dos culturas. Actualmente, siendo responsable, es imposible no ser interdisciplinar.

Bibliografía

- Aguayo, L. V. (2016). “Máquinas de enseñanza de Skinner”. *Facultad de Psicología, Universidad de Málaga*.
- Almenara, J. C. (2006). “Bases pedagógicas del e-learning”. *RUSC. Universities and knowledge society journal*, 3(1).
- Alok Jha (Anfitrión). (2022-presente). *Babbage: Artificial intelligence enters its industrial age* [Podcast]. Acast. <https://www.economist.com/foundational-AI-pod>
- Anderson, C. (2008, 23 junio). “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete”. *Wired*. <https://www.wired.com/2008/06/pb-theory/>
- Arendt, H. (2018). *La condición humana* (5.a ed.). Paidós.
- Artiles Rodríguez, J., Guerra Santana, M., Aguiar Perera, M. V. y Rodríguez Pulido, J. (2021). “Agente conversacional virtual: la inteligencia artificial para el aprendizaje autónomo”. *Pixel-Bit*.
- Béjar, M. (2014). “Una mirada sobre neuroeducación”. *Padres y Maestros/Journal of Parents and Teachers*, (355), 49-53.
- Benjamin, L. T. (1988). “A history of teaching machines”. *American psychologist*, 43(9), 703.
- Bostrom, N. (2016). *Superinteligencia*. Teell Editorial, S.L.
- Bringsjord, Selmer, y Naveen Sundar Govindarajulu. (2020). Artificial Intelligence. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2020/entries/artificial-intelligence/>
- Burch, S. (2005). “Sociedad de la información/Sociedad del conocimiento”. *Palabras en juego*, 56.
- Carbonell, J. R. (1970). “AI in CAI: An artificial-intelligence approach to computer-assisted instruction”. *IEEE transactions on man-machine systems*, 11(4), 190-202.
- Card, S. K., Moran, T. P. y Newell, A. (1983). *The Psychology of Human-Computer Interaction*. CRC Press.
- Casti, J. L. (1998). *El Quinteto de Cambridge*. Taurus.

- Comisión Europea. (2018, 25 abril). *Inteligencia artificial para Europa* [Comunicado de prensa]. <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52018DC0237&from=ES>
- Comisión Europea. (2021, 21 abril). *Una Europa Adaptada a la Era Digital: la Comisión propone nuevas normas y medidas para favorecer la excelencia y la confianza en la inteligencia artificial* [Comunicado de prensa]. https://ec.europa.eu/commission/presscorner/detail/es/ip_21_1682
- Cordua, C. (2012). “La crisis de las humanidades”. *Revista de filosofía*, 68, 7-9.
- Enciclopedia Herder. (2009). *Inteligencia artificial*. https://encyclopaedia.herdereditorial.com/wiki/Inteligencia_artificial#:~:text=Marvin%20Minsky%2C%20uno%20de%20los,ser%20humano%20calificamos%20como%20inteligente%20BB
- Fajardo de Andara, C. (2021). “Marvin Lee Minsky: pionero en la investigación de la inteligencia artificial (1927-2016)”. *Publicaciones en Ciencias y Tecnología*, 15(1), 41-50.
- Gabriel, M. (2022). *El sentido del pensamiento*. Ediciones de Pasado y Presente.
- Galina Russell, I. (2011). “¿Qué son las humanidades digitales?”. *Revista Digital Universitaria*.
- Galindo, C. (2018, 25 julio). Inger Enkvist: “La nueva pedagogía es un error. Parece que se va a la escuela a hacer actividades, no a trabajar y estudiar”. *El País*. https://elpais.com/elpais/2018/07/17/eps/1531826084_917865.html
- García, M. D. V. y de Vega Rodríguez, M. (1988). “Un estudio experimental del razonamiento cotidiano en tareas de silogismos: una aproximación pragmática”. *Cognitiva*, 1(1), 33-62.
- González, B. (2021). *Luces y sombras de la nueva ley europea sobre inteligencia artificial*. UOC (Universitat Oberta de Catalunya). <https://www.uoc.edu/portal/es/news/actualitat/2021/192-ley-europea-inteligencia-artificial.html>
- Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A. y Louwerse, M. M. (2004). “AutoTutor: A tutor with dialogue in natural language”. *Behavior Research Methods, Instruments, & Computers*, 36(2), 180-192.
- Han, B. (2016). *La salvación de lo bello*. Espapdf.

- Han, B. (2014). *Psicopolítica*. Herder.
- Harley, S. (2020). *Menos Tech y más Platón: Por qué la tecnología necesita a las humanidades*. Lid editorial.
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Center for Curriculum Redesign.
- INTEF. (2021, 10 noviembre). *Acto de inauguración del Aula del Futuro del INTEF. Resumen* [Vídeo]. YouTube. <https://www.youtube.com/watch?v=37smnLnGko0>
- INTEF. (2022). *Escuela de Pensamiento Computacional e Inteligencia Artificial 20/21: Enfoques y propuestas para su aplicación en el aula. Resultados de la investigación*
- INTEF. (2019). *Informe Resumen: El impacto de la Inteligencia Artificial en el aprendizaje, la enseñanza y la educación*.
- International Center for Scientific Debate BARCELONA. (2017). *Inteligencia artificial: sueños, riesgos y realidad*. https://www.biocat.cat/sites/default/files/sinopsibdebate_artintelligence_es.pdf
- Ishiguro, K. (2021). *Klara y el sol*. Anagrama Océano.
- Jeff Orlowski (director). (2020). *El dilema de las redes sociales* [Documental]. Netflix.
- Kurzweil, R. (2021). *La singularidad está cerca*. LOLA BOOKS.
- Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación. Boletín Oficial del Estado, 340, de 30 de diciembre de 2020, 122868-122953. https://www.boe.es/boe/dias/2020/12/30/pdfs/BOE-A-2020-17264.pdf*
- Linares, F. M. (2019, 2 febrero). “Así usa China la inteligencia artificial para controlar a sus ciudadanos”. *La Vanguardia*. <https://www.lavanguardia.com/tecnologia/20190202/46161322564/china-inteligencia-artificial-ciudadanos.html>
- Mavrikis, M., Gutiérrez-Santos, S., y Poulouvasilis, A. (2016). “Design and Evaluation of Teacher Assistance Tools for Exploratory Learning Environments”. *LAK '16: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 168–172. <https://doi.org/10.1145/2883851.2883909>
- Mora, F. (2017). *¿Cómo funciona el cerebro?* Alianza.

- Negrín Fajardo, O. y Vergara Ciordia, J. (2009). *Historia de la educación*. Editorial Universitaria Ramón Areces.
- Netflix Is A Joke. (2022, 17 febrero). *The Second Stand-Up Comedy Special Written Entirely By Bots* [Vídeo]. YouTube. <https://www.youtube.com/watch?v=j-zTrfiQMSY>
- Neurorights Foundation. (2022). *The Neurorights Foundation*. <https://neurorightsfoundation.org/>
- Padilla, R. D. M. (2019). “La llegada de la inteligencia artificial a la educación”. *Revista de Investigación en Tecnologías de la Información: RITI*, 7(14), 260-270.
- Parra, S. (2020, 24 febrero). “El colegio en China gracias a la inteligencia artificial”. *MuyInteresante.es*. <https://www.muyinteresante.es/tecnologia/articulo/asi-es-el-colegio-en-china-gracias-a-la-inteligencia-artificial-191582559478>
- Perrenoud, P. (2008). *La evaluación de los alumnos. De la producción de la excelencia a la regulación de los aprendizajes. Entre dos lógicas*. Colihue.
- Pressey, S.L. (1944) *Psychology and the new education*. Harper
- Rickert, H. (1922). *Ciencia cultural y ciencia natural*. Espasa Calpe.
- del Rio, A. Q. (2019, 4 septiembre). “2.5 Variables discretas y continuas | Estadística Básica Edulcorada”. *Estadística Básica Edulcorada*. <https://bookdown.org/aquintela/EBE/variables-discretas-y-continuas.html#:~:text=Si%20las%20observaciones%20corresponden%20a,valor%20dentro%20de%20un%20intervalo.>
- Rodríguez Chávez, M. H. (2021). “Sistemas de tutoría inteligente y su aplicación en la educación superior”. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*. <https://www.ride.org.mx/index.php/RIDE/article/view/848/2978#citations>
- Rojas Castro, A. (2013). “Las Humanidades Digitales: principios, valores y prácticas”. *Janus. Estudios sobre el siglo de oro*.
- Ruiz Herrero, J. (2016). “Escribiendo a máquina: la evaluación automática de textos a debate”. *Revista Didáctica de Español como Lengua Extranjera*. https://marcoele.com/descargas/23/ruiz-evaluacion_automatica.pdf

- Sanguineti, J. J. (2008). *Filosofía de la mente*. Philosophica, Enciclopedia filosófica on line.
<https://www.philosophica.info/>
- Savater, F. (2010). *El valor de educar*. Ariel.
- Searle, J. R. (1980). "Minds, brains, and programs". *Behavioral and Brain Science*. vol. 3, nº 3, 417-457.
- Snow, J. P. (1988). *Las dos culturas*. Ediciones Nueva Visión.
- Solomon, C. J. y Papert, S. (1976, June). "A case study of a young child doing Turtle Graphics in LOGO". *Proceedings of the June 7-10, 1976, national computer conference and exposition* (pp. 1049-1056).
- Steiner, G. (2005). Conferencia que pronunció en la Cátedra Ferrater Mora de la Universidad de Gerona, citado por http://www.lainsignia.org/2005/octubre/dial_003.htm
- The Economist. (2022, 9 junio). *How smarter AI will change creativity*.
<https://www.economist.com/leaders/2022/06/09/artificial-intelligences-new-frontier>
- Unamuno, M. (2021). *Amor y pedagogía*. Verbum.
- UNESCO. (2019). *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*.
<https://unesdoc.unesco.org/ark:/48223/pf0000366994/PDF/366994eng.pdf.multi>
- UNESCO. (2021). *AI and education. Guidance for policy-makers*.
https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p::usmarcdef_0000376709&file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_c6ac8eec-8f44-4967-9057-5ea2d1e7196a%3F%3D376709eng.pdf&locale=es&multi=true&ark=/ark:/48223/pf000376709/PDF/376709eng.pdf#AI%20in%20education_pages.indd%3A.14150%3A1021
- Vera, G. G. H. (2007). "La formación Humanística en el mundo contemporáneo". *El Hombre y la Máquina*, (29), 4-7.
- Vita Proupech, S. F. (2004). "Sobre el problema de la enseñanza de las humanidades en la Educación Secundaria". *Didáctica (lengua y literatura)*.

VICE News. (2019, 29 junio). *China's Vanishing Muslims: Undercover In The Most Dystopian Place In The World* [Video]. YouTube.
<https://www.youtube.com/watch?app=desktop&v=v7AYyUqrMuQ>

ANEXO 1. GLOSARIO DE TÉRMINOS TÉCNICOS

***Blackbox* o caja negra de la inteligencia artificial.**

Término que alude a la dificultad por parte de los programadores de explicar por qué los algoritmos generados han tomado una decisión. Esta situación es propia de todas aquellas técnicas que incorporan un proceso de aprendizaje automático. Hablamos así de la existencia de una caja negra debido a que resultaría extremadamente difícil hacer un seguimiento de los pasos llevados a cabo por los algoritmos.

Esta situación de ininteligibilidad de los procesos puede acarrear problemas éticos y legales. Una gran parte de la legitimidad de las decisiones consiste en su capacidad de fundamentación.

Nos encontramos así, problemas con la atribución de la responsabilidad ¿Qué ocurre si un robot que realiza diagnósticos médicos se equivoca? ¿Y si un algoritmo de evaluación automática de los alumnos proporciona predicciones de resultados erróneas?

También nos encontramos con problemas relacionados con la cuestión de protección de los datos. ¿Cómo podemos permitir la recolección de todos esos datos de nuestros estudiantes si desconocemos el proceso por el cual van a ser tratados esos datos?

Asimismo, ¿y si se incorporaran estos procesos automáticos a la hora de seleccionar a alumnos para un centro? ¿Cómo podremos justificar a los admitidos si desconocemos los criterios empleados?

Ciencia de datos.

Según algunos, denominación alternativa para la *inteligencia artificial*. Hay expertos que sostienen que este término resultaría más adecuado a las funciones reales que desempeña esta disciplina. De hecho, habría quien sostendría que esta denominación implicaría un menor rechazo social que la alternativa, así como una mayor predisposición a su aceptación.

La “ciencia de datos” aludiría a un mero campo de conocimiento destinado al tratamiento de los datos y no implicaría de suyo el empleo de una metáfora que parece retrotraernos a un panorama de sustitución de los humanos por máquinas. De igual modo, esta denominación parece presentar una mayor vinculación fonética con una de las problemáticas de fondo en esta era de la inteligencia artificial: el problema del dataísmo.

Cuarta revolución industrial.

Este término fue acuñado por Klaus Schwab, fundador del foro económico mundial. Alude ese cambio del modelo productivo y social derivado del crecimiento exponencial de las tecnologías que comenzaría con la creación, en los años 80, de los ordenadores particulares y de internet.

Dentro de todos estos cambios tecnológicos y sociales algunos auguran que la inteligencia artificial cobraría un papel. El experto en inteligencia artificial Andrew Ng, ha sostenido que “La inteligencia artificial es la nueva electricidad”.

Este tipo de declaraciones, sostenemos, suelen hacer referencia al potencial de estas tecnologías para interferir en el mercado de trabajo y en nuestra vida cotidiana.

Datos estructurados y datos no estructurados.

Esta clasificación de los tipos de datos es de extrema importancia para el entendimiento de lo que sería la disciplina de la inteligencia artificial. Esto es debido a que de la posibilidad del acceso a los datos depende el futuro desarrollo de la inteligencia artificial.

Los datos estructurados serían aquellos caracterizados por ser fácilmente organizables y tratables puesto que se prestan fácilmente a una disposición de tablas y columnas semejante al archivo de una biblioteca. Por ejemplo, son datos estructurados las fechas, las direcciones, las ventas obtenidas, los exámenes de múltiple opción...

Por el contrario, los datos no estructurados son aquellos que por su propia forma de ser resultan complicados de analizar. Como su nombre indica, no presentan propiamente una estructura o esqueleto que hagan que tengan un modo de organización predefinidos. Por esto, su clasificación y tratamiento son más complicados. Son datos no estructurados las imágenes, los vídeos, determinados textos, los sonidos, los ensayos...

La mayoría de los datos que emplean las empresas, en torno a un ochenta por ciento, se corresponde con datos no estructurados. Es por ello por lo cual, ante esta explosión de la ciencia de datos, aquellas que logran un superior tratamiento de este tipo de datos obtienen ventaja competitiva.

Deep learning.

Es un tipo de *machine learning*. Puede analizar vídeos, imágenes y datos no estructurados de formas que el *machine learning*, en el sentido general previamente descrito, no puede. A través de múltiples ejemplos, y sin especificar sus características, el algoritmo, si ha sido correctamente entrenado, es capaz de identificar objetos con una precisión superior al humano.

En este nivel se considera que la “máquina” sería capaz de razonar y de tomar decisiones. Un ejemplo de *deep learning* sería el siguiente.

Se proporcionarían imágenes de árboles al dispositivo sin introducir las características específicas de los árboles. Tras haber visionado las fotografías suficientes, el algoritmo es capaz de inferir por sí mismo qué es un árbol y qué no. La contrapartida de este método reside en que para llegar a esta posibilidad se necesita proporcionar un número elevado de material.

Su funcionamiento se inspira en una modelización del cerebro humano utilizando lo que se conoce como redes neuronales profundas. Las redes neuronales consistirían en varias capas de neuronas artificiales capaces de transformar la información de manera no lineal y bastante abstracta.

Giro traicionero.

Posibilidad planteada por Nick Bostrom en su libro “superinteligencia”. Sería un escenario en el cual la IA se volviera contra nosotros. Esto podría ocurrir en el supuesto de que la IA se volviera tan inteligente que pudiera engañarnos a la hora de conseguir sus objetivos. Pudiera ocurrir que, por ejemplo, la IA ya no buscara una recompensa, sino que tomara el control del mecanismo de recompensa.

Inteligencia artificial.

Se dice que la denominación “inteligencia artificial” fue acuñada por el informático y matemático John McCarthy en una de sus conferencias en el año 1956. De manera previa a esta fecha, esta técnica se habría llamado simulación computerizada.

Parafraseando a Marvin Minsky, *inteligencia artificial* haría referencia a aquella disciplina capaz de hacer que los ordenadores desempeñen actividades y tomen decisiones de las que, si fueran realizadas por humanos deduciríamos actividad inteligente.

No obstante, el proyecto de inteligencia artificial ha tomado varias derivas en su historia. Habitualmente, distinguimos dos proyectos distintos que han tomado desarrollo histórico a raíz de lo que fue el proyecto inicial de la inteligencia artificial de Alan Turing, Marvin Minsky, John McCarthy, Allen Newell o Herbert Simon. Así, distinguimos entre inteligencia artificial débil o estrecha e inteligencia artificial fuerte o general.

Inteligencia artificial débil o estrecha.

Este es el tipo de inteligencia artificial que se ha logrado actualmente. Es la más empleada y está disponible en las técnicas como el *machine learning* y el *Deep learning*. Se caracteriza por ejecutar acciones determinadas como detectar el correo basura, jugar al ajedrez o reconocer la voz de una persona. Asimismo, el hecho de realizar una función específica no capacita a este sistema para realizar otras actividades que serían supuestas a un humano si, por ejemplo, supiera detectar el correo basura.

Inteligencia artificial fuerte o general.

Tipo de inteligencia artificial propia del proyecto inicial de *inteligencia artificial*. Implicaría la posibilidad de autoconocimiento y de resolución de múltiples problemas en una variedad de contextos, así como de resolver problemas que los sistemas desconocían en el momento en el cual fueron creados. Este aspecto podríamos relacionarlo con la cuestión de la capacidad de *transferencia*, facultad que es muy propia de los humanos e implica la posibilidad de transportar aprendizajes obtenidos en un campo a otro tipo de experiencias bien distintas.

Machine learning.

Serie de métodos, considerados inteligencia artificial débil, utilizados para detectar patrones en datos. Los programadores de *machine learning* construyen sistemas capaces de mejorarse de forma automática a través de la experiencia.

El algoritmo de *machine learning* funciona en tres pasos que se repiten de forma recurrente. En primer lugar, realiza una búsqueda de patrones a través del análisis de los datos. A continuación, tomando como referencia los patrones obtenidos genera un modelo. Después, con fundamento en este modelo, el algoritmo es capaz de realizar una serie de acciones.

Sin embargo, el proceso no termina aquí, puesto que los resultados de las acciones del algoritmo generarán nuevos datos que serán a su vez registrados por este para dar

comienzo de nuevo al proceso descrito en el apartado previo. Como se ve, el funcionamiento del algoritmo permite que éste vaya mejorando el modelo en el cual se fundamenta su toma de acción. Con ello, quizás podríamos decir que este dispositivo efectivamente aprende pues es capaz de modelar su “conducta” tomando como referencia “experiencias” de datos pasadas.

Aunque programar estos algoritmos sea complicado, esto en cierta medida puede ahorrar grandes esfuerzos. Programar un algoritmo capaz de aprender supone que en un momento inicial no debemos tener un conocimiento tan específico de la situación para la cual queremos aplicar el algoritmo, el algoritmo solo se irá adaptando. Un ejemplo de *machine learning* sería el algoritmo que detecta el correo basura o el que elabora las recomendaciones de plataformas de compras o de series para los usuarios.

Existirían tres tipos de *machine learning* según su modo de aprendizaje (Holmes et al, 2019).

Con el tipo de ML con aprendizaje supervisado daríamos al algoritmo muchos datos ya categorizados. El algoritmo identificaría la función que permite conectar los datos a las etiquetas. De ahí construiría un modelo que puede aplicarse a datos similares. Facebook usa este tipo de *machine learning* para identificar y etiquetar automáticamente a las mismas personas en nuevas fotografías.

Con el aprendizaje no supervisado se proporciona al algoritmo una enorme cantidad de datos que no han sido categorizados o clasificados. El algoritmo analiza los datos sin categorizar en busca de patrones escondidos. Este funcionamiento tiene lugar, por ejemplo, en programas para identificar diferentes letras.

El aprendizaje por refuerzo es bastante poderoso. En los dos casos anteriores el modelo que se genera a raíz de los datos está cerrado, pero si los datos cambian entonces el análisis se emprende otra vez. Lo que hace el aprendizaje por refuerzo es ir mejorando continuamente el modelo basándose en el *feedback*. Se van evaluando los datos y asignando a correctos o incorrectos y así el algoritmo se va perfeccionando.

Máquina infantil de Turing.

Como vía de acercamiento a la aspiración al logro de una inteligencia artificial general, Turing propuso la construcción de una “máquina infantil” o “niño computacional”. Esta idea sostenía que sería más sencilla la constitución de una

inteligencia similar a la humana a través de la creación de una máquina con las habilidades de un niño. Si a esta máquina infantil, dotada de la capacidad de aprender, la fuéramos enseñando progresivamente a desenvolverse con el entorno, entonces sería capaz de obtener el mismo tipo de inteligencia que un humano adulto, así como sus habilidades.

Esta propuesta se oponía a la aspiración de crear una máquina que reuniera de manera directa las capacidades de un adulto. De igual modo, se contempla la fundamental dimensión que representa el aprendizaje en el desarrollo de la inteligencia, así como lo crucial del trato con el entorno.

Prueba de Turing o *juego de la imitación*.

Es una prueba clásica fue diseñada por Alan Turing con el objetivo de mostrar si una computadora tiene inteligencia. Como su nombre indica, esta prueba se basa en una suerte de “juego de imitación”. Una persona seleccionada debe establecer una comunicación a ciegas y realizar una serie de preguntas establecidas. Si evalúa que las respuestas dadas por el agente comunicativo son inteligentes, esto es, podrían ser ejecutadas por un ser humano, entonces se deduce que la máquina, presente en el otro lado, es inteligente.

Esta prueba es a menudo criticada por el hecho de que definiría la inteligencia en unos parámetros conductistas. Esto es, inteligencia en la prueba de Turing tal solo quedaría entendida como conducta inteligente al margen de aquello que denominaríamos “estados mentales”. Esta problemática es debatida con brillantez en la novela “El quinteto de Cambridge” de John L. Casti.

Redes neuronales artificiales.

Sistemas de inteligencia artificial que simulan unidades neuronales siguiendo un modelo de cómo estas actuarían en el cerebro. Por ello suelen organizarse por capas o nodos interconectados y tienen programada la capacidad de aprendizaje autónomo. Un ejemplo de funcionalidad inteligente desarrollada a través de redes neuronales artificiales sería el sistema de reconocimiento por voz.

En este proyecto se pueden observar retazos y aportaciones del proyecto inicial de inteligencia artificial general, esto es, de la aspiración de dar con un modelo computacional de la mente. Hemos de recordar que ya en el año 1951, Marvin Lee Minsky construyó el primer simulador de redes neuronales.

Relevancia.

Aspecto de discriminación de la información verdaderamente importante. Es un problema severo en la actualidad discriminar la información relevante en nuestro severo y necesario tratamiento con los datos. Debido a la inmensa cantidad disponible en la red y en los distintos soportes, resulta del todo punto dificultoso hallar procesos discriminadores efectivos.

No obstante, el ser humano parece tener una capacidad especial para hacer este procedimiento selectivo. Asimismo, esta competencia junto con la de transferencia parece implicar de suyo un tratamiento interdisciplinar de la información. Este aspecto es algo que las máquinas no todavía no tendrían y, por ello, suele plantearse que debería potenciarse en los seres humanos.

Singularidad tecnológica.

Momento en el cual la inteligencia artificial sería más inteligente que los humanos. Kurtweil propone que a través de las combinaciones con la misma por parte del humano sería posible trascender nuestra condición biológica y, en consecuencia, llegar a una era posthumana pero, por fortuna, no necesariamente al fin de la especie humana.

Superinteligencia.

Concepto propuesto por el transhumanista Nick Bostroom en un libro homónimo. El término alude a la posibilidad de alcanzar el logro de una inteligencia superior a la de los mayores genios de la historia a través del uso de medios técnicos. Se plantea que ésta podría ser alcanzada a través de diversas vías como mejoras genéticas, a través de una organización social en colmena e incluso mediante computadoras.

Si la superinteligencia fuera alcanzada por una computadora podríamos enfrentarnos a un escenario en el cual los humanos quedarán sometidos a los propios fines de la misma. De igual modo, la superinteligencia aplicada al humano parece basarse en un argumento antagónico al evolucionismo clásico. La adaptación cerebral ya no seguiría criterios de supervivencia, sino que sería elegida por los científicos. En parámetros transhumanistas, se plantea que el ser humano decida los pasos de la evolución de su especie centrándose en el factor “inteligencia” y no en la mera adaptación al entorno.

Transferencia.

Capacidad de trasvasar una serie de conocimientos adquiridos en un campo o en un contexto concreto hacia otro sin necesidad de haber sido reforzado en ese otro. Este tipo de competencia, al parecer, en estos momentos sería algo específicamente humano.

➤ **Bibliografía del glosario.**

Bostrom, N. (2016). *Superinteligencia*. Teell Editorial, S.L.

Caja negra vs. de cristal: la IA que funciona contra la que se explica. (2020, 7 febrero).

MIT Technology Review. <https://www.technologyreview.es/s/11839/caja-negra-vs-de-cristal-la-ia-que-funciona-contra-la-que-se-explica>

Comisión Europea. (2018, 25 abril). *Inteligencia artificial para Europa* [Comunicado de prensa]. <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52018DC0237&from=ES>

(s. f.). *Descubre los principales beneficios del Machine Learning*. Iberdrola. Recuperado 10 de mayo de 2022, de [https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico#:~:text=El%20Machine%20Learning%20es%20una,elaborar%20predicciones%20\(an%C3%A1lisis%20predictivo\)](https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico#:~:text=El%20Machine%20Learning%20es%20una,elaborar%20predicciones%20(an%C3%A1lisis%20predictivo)).

G.J. (2019, 6 agosto). *¿Qué son los datos estructurados?* Brandwatch.

<https://www.brandwatch.com/es/blog/datos-estructurados/#:~:text=Ejemplos%20de%20datos%20estructurados&text=El%20resultado%20de%20los%20cuestionarios,se%20recopila%20de%20manera%20uniforme>

Holmes, Bialik y Fadel (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Center for Curriculum Redesign.

Kurtzweil, R. (2021). *La singularidad está cerca*. LOLA BOOKS.

Rouhiainen, L. (2018). *Inteligencia artificial: 101 cosas que debes saber hoy sobre nuestro futuro* (Alienta). Alienta Editorial.

UNESCO. (2019). *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*.

<https://unesdoc.unesco.org/ark:/48223/pf0000366994/PDF/366994eng.pdf.mult>

Fajardo de Andara, C. (2021). “Marvin Lee Minsky: pionero en la investigación de la inteligencia artificial (1927-2016)”. *Publicaciones en Ciencias y Tecnología*, 15(1), 41-50.