

Entrevista a Federico Ferrero



“

A partir de su trabajo como investigador en el campo de la tecnología educativa desde una perspectiva sociocultural, **Federico Ferrero** comenzó a preguntarse por la construcción de sistemas informáticos destinados a la evaluación de docentes y estudiantes. Es por ello que su interés no está únicamente en comprender los resultados arrojados por los algoritmos que procesan gran cantidad de datos educativos, sino también en debatir, desde los saberes que aportan las ciencias de la educación, con diferentes especialistas provenientes de otros campos de estudio como son la programación, la estadística, etc. Entre muchos otros, uno de los tópicos que genera frecuentes intercambios es el reconocimiento de diferentes tipos de sesgos y las consecuencias que implican en las prácticas sociales de los agentes que estos sistemas de evaluación involucran.



El trabajo del pedagogo en tiempos de algoritmos y big data

Entrevista a Federico Ferrero.

¿Por qué estudiar big data en educación?

Big data (macrodatos) hace referencia a la gran cantidad de datos que hoy están disponibles a partir del uso de entornos virtuales. A ello, se suma la capacidad técnica que hoy se tiene para llevar adelante el procesamiento de una enorme cantidad de información que proviene de fuentes muy diversas.

Cuando hacemos uso de entornos virtuales, motivados por diferentes razones –ya sean particulares o institucionales–, vamos dejando huellas que no son otra cosa que sellos que revelan el flujo y la acumulación de información. Como los sedimentos de las capas geológicas, esta información está disponible para ser procesada por técnicas de inteligencia artificial o de minería de datos, lo que permite el desarrollo de ciertos sistemas de cálculo o algoritmos cuyo fin es promover una forma definida para la toma de decisiones. La importancia de esta cuestión se puede ponderar cuando se considera la penetración que estos algoritmos empiezan a tener en campos muy diversos de nuestras sociedades como el de los negocios, el de la salud o el de la gestión de los sistemas penitenciarios. Por supuesto que esto también ocurre en lo que aquí nos convoca y que está vinculado a la educación. Por ser toda esta materia una forma de valoración, resulta claro que aquello que se está poniendo sobre la mesa es una categoría clásica de la pedagogía: la evaluación. Es por ello que comienzan a aparecer una serie de reflexiones en torno a estos sistemas de procesa-



miento de datos y sus significados para la enseñanza y los aprendizajes. Por ejemplo, tenemos sistemas de evaluación de la “productividad” docente, del aprendizaje de los estudiantes, de la consideración y la recomendación de clases para los alumnos y de las razones para actuar sobre ellos. Lo que resume todo lo planteado, y lo que debemos sopesar en lo referente al uso de macrodatos, es la controvertida idea referida a la posibilidad de que en la educación se pueda dar una predicción de carácter determinista, reduciendo todo el arte a una cuestión técnica. Por ello, la apuesta desde lo pedagógico es tratar de pensar de forma crítica el sentido de estos sistemas algorítmicos que, insistimos, son modos de evaluación. Para este ejercicio de reflexión es necesario aglutinar a programadores, estadísticos y educadores.

¿Qué son los algoritmos?

La definición de algoritmo es compleja porque cuando uno se le aproxima desde una mirada sociocultural no es posible reducir la definición de esa idea a la forma matemática clásica que nos dice que es “un conjunto estructurado de pasos que tienen una expresión formal y que permiten analizar un conjunto de entradas (*inputs*) sobre las que se aplican ciertas operaciones para obtener determinados resultados (*outputs*)”. La idea de algoritmo implica la perspectiva de la automatización de procesos. Por eso, se construyen y escriben programas: para automatizar procesos y liberar de cargas cognitivas a los seres humanos, y que, así, puedan dedicarse a otras actividades. Pero, abordándolo socioculturalmente, se hace necesario trascender esta idea matemática y considerar cómo todo ese proceso formal está inserto en determinados contextos de uso y que, este hecho, tiene efectos concretos en la vida cotidiana de los sujetos que están implicados. Por ello, decimos que los algoritmos son complejos, que no son solamente una expresión formal de texto de código, sino que llevan implícito en su estructura la intervención de los sujetos que participan en su diseño, en su desarrollo y en su revisión. Incluso, se debe tener en cuenta que quienes programan y diseñan un mismo conjunto de algoritmos tienen objetivos distintos. Puede ocurrir, además, que hayan sido creados en el pasado y que se los reutilice en nuevos textos de programa. Otra característica



importante a considerar en este abordaje social de la idea de algoritmos es que, como todo acto humano, están embebidos en contextos particulares de uso. El estudio aislado de un algoritmo como un dispositivo matemático no basta y se requiere considerar su inserción en contextos más amplios para poder valorar los efectos que tienen sobre la vida de las personas a las cuales esos algoritmos involucran. Por otra parte, es importante saber que, como cualquier tecnología educativa, los algoritmos son expresiones que están comprometidas políticamente. Detrás de ellos hay asunciones, hay prioridades, hay ciertos enfoques que hacen que un algoritmo no sea un objeto neutro. Su constitución lo define como una manera de ver el mundo y de modelizar fenómenos. Esto nos exige un estudio que pueda poner en evidencia esas asunciones y compromisos. Es en ese sentido que la pedagogía, en el caso de algoritmos vinculados a la educación, puede ayudarnos a tratar de leer cuáles son los enfoques desde donde se diseñan.



Para analizar un algoritmo referido al aprendizaje, se deben tener en cuenta las múltiples dimensiones y las complejas relaciones que se dan en el escenario educativo. Muchas veces sucede que se hace un análisis formal sencillo y se dejan afuera facetas que, desde la teoría educativa, son consideradas cruciales para entender los algoritmos de manera más apropiada. La pedagogía debe ayudarnos a ver cuáles son las relaciones que se deben tener en cuenta –o que se han supuesto– al momento de modelizar, por ejemplo, el aprendizaje.



Algoritmos y sesgos en los resultados

Cuando se utilizan los sistemas algorítmicos en educación, se constatan, en función de otros saberes, sesgos en los resultados. O sea que ciertos grupos, cuyos datos se procesan, son “beneficiados” por sobre otros. Ante este estado de situación, la reflexión de carácter pedagógico nos puede ayudar en la identificación, el análisis y la elucidación de estos sesgos. Esto nos va a permitir comprender modos, o formas, que nos parecen inadecuados en función de ciertos horizontes de posibilidades deseables. La pedagogía es un discurso prescriptivo que nos hace pensar que hay un “deber ser”, un proyecto sobre el cual descansamos y en el cual confiamos. Entonces, cuando estos resultados aparecen sesgados respecto de este mundo axiomático, se nos plantea una discusión ética. En estos casos, se hace inevitable el análisis sobre el significado que tiene el uso de este tipo de sistemas. Lo que parece crucial es poner en evidencia que la categoría de sesgo tiene una particular pertenencia de carácter pedagógico, lo cual no solo permite, sino que obliga a los pedagogos a ingresar en el estudio de esta cuestión.

Algunos ejemplos pueden ayudar a comprender e ilustrar lo que sostengo. Es el caso de las plataformas de video: hay estudios que muestran que aquello que hacen los algoritmos de recomendación es radicalizar la sugerencia que formulan de tal forma que, finalmente, se obtienen videos que son “conservadores” con respecto al tema que se está buscando. El algoritmo te lleva a resultados sesgados en tanto que “blindan” cierta perspectiva sobre el contenido con el cual se inició la búsqueda. Por supuesto que hay discusiones sobre las agencias que están detrás, en relación con la publicidad y la propaganda política, y que sostienen económicamente el funcionamiento de estos algoritmos. Es interesante considerar que cada vez confiamos más en los resultados de estas búsquedas en plataformas de video como formas de acceder a las fuentes de información, pero, lejos de cualquier neutralidad, lo que está operando allí es una política de recomendación que es subliminal y modela nuestra manera de apreciar el mundo.



¿Qué tipos de sesgos pueden identificarse en los algoritmos?

Me voy a referir a los sesgos que identificamos en los algoritmos cuyas reglas son definidas “de antemano” por los desarrolladores y son respetadas de manera exacta al ser ejecutadas (a diferencia de los más recientes algoritmos que “aprenden” autónomamente). Cuando trabajamos la idea de sesgo, se tiende a tratar de identificarlos como si fuesen propios del algoritmo. Y nuestra hipótesis de trabajo es que, en realidad, en el propio algoritmo están reflejados sesgos que provienen de otros elementos de la unidad de análisis con la que trabajamos, en este caso, la Teoría de la Actividad, que es nuestro enfoque teórico. En esta teoría, lo que se utiliza es una unidad de análisis que se juzga suficiente y pertinente en donde un sujeto aprende o conoce sobre un objeto o trata de dominar una situación problemática con la mediación instrumental y, con la mediación de otros sujetos, de la comunidad; todo esto, asentado en una base que considera la división del trabajo y las reglas que allí funcionan. Al ser lo suficientemente amplia esta unidad de análisis, se hace posible el estudio de los algoritmos no de manera aislada, sino en su relación con los otros elementos de la unidad de análisis. Animados por esta unidad de análisis, más allá del instrumento mediador, intentamos determinar los loci donde poder identificar las fuentes de sesgos que terminan reflejados en el algoritmo, pero que provienen de otros lugares. Lo que tratamos de verificar allí es la idea de “cognición distribuida”. La cognición se distribuye en cada uno de los elementos de la unidad de análisis, que es una idea clásica para el enfoque sociocultural. Movidos por este interés –donde no queremos hacer un análisis aislado sobre el algoritmo propiamente dicho, sino ver cómo impactan los otros elementos y las relaciones en el propio algoritmo–, lo que podemos identificar son ciertos tipos o clases de sesgos, que incluyen: sesgos teóricos, sesgos metodológicos, sesgos por interpretación errónea de los datos, sesgos por descontextualización de los sistemas y sesgos por entrenamiento del algoritmo.

Los sesgos teóricos están implícitos en el proceso de modelización de un fenómeno. Allí, el sujeto que se encarga de programar, desarrollar o evaluar, toma decisiones para modelizar el fenómeno sobre el cual está trabajando. Esas decisiones implican ciertas



prioridades, ciertos enfoques donde algunas dimensiones tienen más peso que otras. Todas esas decisiones que, en el mejor de los casos, son explícitas y, en el peor, quedan implícitas y están funcionando a pesar de que los propios diseñadores no se anotan de ello. Un sesgo teórico frecuente es el uso privilegiado de lo que es fácilmente medible antes de lo que es realmente importante de ser considerado. Nos hace pensar que las cuestiones teóricas en la educación ayudarían a evitar estos sesgos teóricos que simplifican fenómenos que, sabemos, tienen una complejidad muy grande, tal como ocurre con la enseñanza y el aprendizaje.



Los sesgos metodológicos son los vinculados a las decisiones que toman los diseñadores de los sistemas, influenciados por sus saberes particulares en el campo de la estadística, la inteligencia artificial, la minería de datos; hay ciertas métricas y ciertas técnicas que requieren determinadas condiciones para elegir si usarlas o no. Todos esos saberes, que son muy específicos del campo de la estadística y de la programación, si no se consideran pueden producir sesgos al elegirse técnicas que, tal vez, no son adecuadas para tratar ese conjunto de datos, etc.

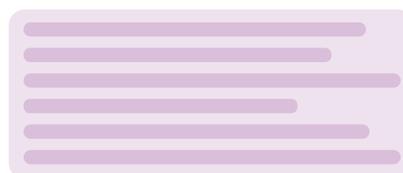
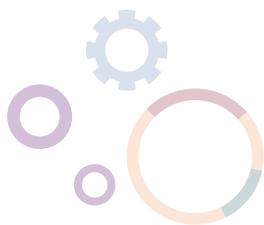


Otra posible fuente de sesgo está en la interpretación de los resultados de los sistemas algorítmicos que estudiamos. Muchos de estos sistemas funcionan con principios de razonamiento inductivo, que es la forma en la que se construye conocimiento estadístico. Es la idea de que uno puede obtener conocimiento con una cierta probabilidad y sin certeza absoluta –a diferencia de los razonamientos de tipo deductivo, en los que se puede concluir con total certeza–. Estos razonamientos de tipo inductivo nos vinculan con resultados que, en realidad, están atados a cierto margen de error y a ciertos intervalos de confianza. Entonces, si uno desconoce o ignora que este saber tiene esta calidad, podríamos interpretar erróneamente los resultados del trabajo estadístico que está implicado allí. También, se pueden citar casos de interpretación no adecuada de estimadores o parámetros. El caso clásico es la correlación estadística que, cuando se establece, es leída en términos de relaciones causales, lo cual la propia estadística no nos permite establecer: que se den fenómenos simultáneos no quiere decir que tengan un vínculo de causa-efecto. Así como muchos otros ejemplos vinculados a cómo interpretamos los resultados que nos arrojan estos sistemas algorítmicos y que son responsabilidad de quienes leemos esos resultados: de los administradores en los sistemas educativos, de los que hacen política educativa... El conocimiento en estadística es muy útil para saber cuándo se está utilizando ese conocimiento de manera adecuada.

Cierta forma de sesgo de resultados se produce con la extrapolación de sistemas desde las condiciones iniciales de desarrollo y uso hacia su aplicación, de manera casi lineal, en nuevos contextos. Esto carece de legitimidad porque, si pensamos en la unidad de análisis, nuevos son los sujetos, nuevas son las comunidades que están incluidas, nuevas son las reglas que allí funcionan, nueva es la división del trabajo que está involucrada. Es el problema de la “escalabilidad”: cómo algoritmos que funcionan bien en ciertos contextos, escalan y se generalizan en muchos casos, y terminan generando problemas para que los datos sean analizados de manera adecuada.



Una fuente de sesgo importante para identificar es el producido en el entrenamiento del algoritmo. Lo que sucede es que una de las fases del desarrollo incluye el "aprendizaje" (no me refiero al concepto técnico utilizado en *machine learning*) del propio algoritmo donde opera un proceso de ajuste para minimizar el error posible, para decirlo de una manera simple. Para realizar este proceso de ajuste se utilizan bases de datos que permiten mejorar el algoritmo. En este caso, la calidad de los datos nos interesa porque pueden estar sesgados ellos mismos, trasladándose esto al funcionamiento del algoritmo. Un caso clásico es el de los algoritmos de reconocimiento facial que, en ocasiones, se han entrenado con rostros caucásicos. Entrenados de ese modo, cuando se los pone a funcionar, no pueden reconocer rostros que no sean blancos, entonces, aparecen las discusiones sobre discriminación, racismo y algoritmos. En realidad, es un tema vinculado a los sesgos que provienen del entrenamiento. Otro caso es el de los autos autónomos. En el último tiempo, ha salido un trabajo de investigadores del MIT [Massachusetts Institute of Technology] donde describen el resultado de un ensayo en el cual hicieron un videojuego donde se les presenta a los jugadores una situación ficticia de conducción de un auto en la que deben decidir a quién salvar ante un choque inminente. Lo interesante que estos investigadores encuentran es que, dependiendo de las regiones geográficas, las decisiones que toman los jugadores son diferentes.



Por ejemplo, si bien muchos de ellos coinciden en salvar a grupos antes que a individuos, mujeres antes que hombres, en países donde hay grandes diferencias económicas, los jugadores suelen preferir salvar a personas que parecen hombres de negocios antes que a gente que vive en la calle, etc. Estas decisiones –que van quedando en estas bases de datos y en las que han participado más de dos millones de personas en todo el mundo– muestran que los propios datos



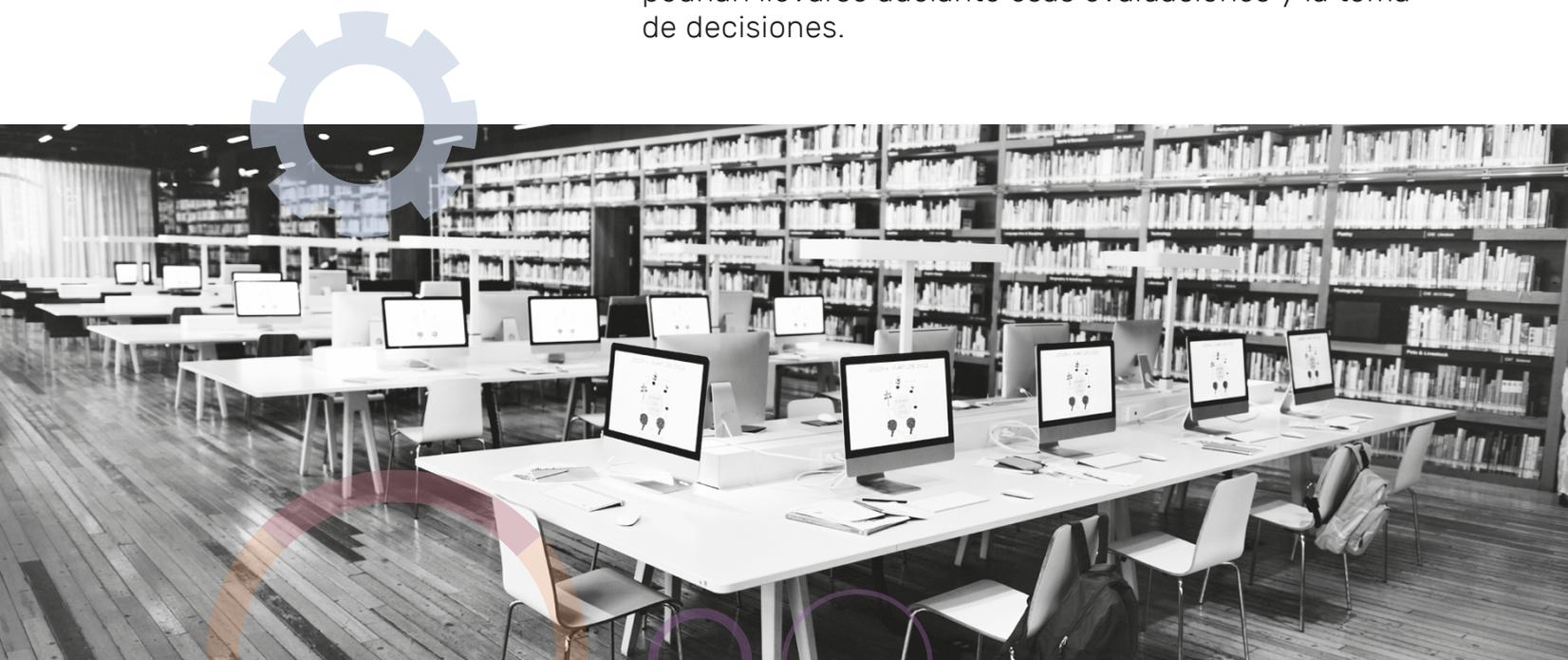
–que van quedando en estas bases de datos y en las que han participado más de dos millones de personas en todo el mundo– muestran que los propios datos que provienen de nuestra actuación en los entornos virtuales están sesgados y que responden a ciertas formas de ver el mundo. Por lo tanto, en el caso de que los algoritmos dedicados al desarrollo de autos autónomos utilicen estos datos, es evidente que implicaría un sesgo en la forma en la que estos vehículos funcionarán. Incluso, empieza a haber discusiones sobre “colonialismo moral”, porque las reglas morales de determinados grupos pueden ser vendidas en el paquete algorítmico, pero pueden resultar extrañas o ajenas a los otros “consumidores” de este tipo de sistemas. Lo que se muestra con claridad en este trabajo del MIT es que los datos previamente sesgados tienen efecto sobre el entrenamiento de estos algoritmos, de modo que la fuente de los sesgos es preexistente al propio algoritmo y está en manos de la gente que deja su “huella” al usar entornos virtuales. Por eso, si nos atenemos a esta clasificación, es evidente que no es el algoritmo en sí mismo el que está sesgado, sino que está en los sujetos que diseñan, en los que leen los resultados de estos algoritmos, en aquellos que escalan hacia otros contextos y en la posición de la comunidad cuyos datos se utilizan para entrenar el algoritmo. Son estos los lugares desde donde emergen las formas sesgadas que se terminan condensando y cristalizando en el algoritmo en sí, pero pueden incluirse como nuestras responsabilidades. Poner en evidencia esto es una forma de reconocer el lugar que nos toca a quienes participamos en la configuración de estos sistemas para que sea posible intervenir sobre estos espacios problemáticos; porque, además, son problemas que están en nuestras manos.

Los algoritmos y las transformaciones en las prácticas de los agentes

Si tenemos en cuenta esta clasificación de sesgos, podemos pensar en diferentes momentos de la producción de estos sistemas algorítmicos y, allí, concentrarnos en diferentes formas de sesgos; por ejemplo, metodológicos o teóricos. En el momento de implementación de estos sistemas, es necesario revisar cómo las lecturas e interpretaciones erróneas de los resultados pueden tener efectos sesgados, incluso,



considerar el uso descontextualizado, porque los sucesos pueden ocurrir en situaciones muy diferentes. También, es interesante analizar la cuestión del tiempo que organiza la vida de un sistema de actividad que se dedica a la toma de decisiones usando algoritmos, porque lo que se ha mostrado es que, cuando un sistema algorítmico en funcionamiento valora ciertas prácticas, ciertos usos, lo que aparece cada vez más patente es que el accionar de los propios sujetos se va acomodando a los requerimientos para satisfacer al sistema tecnológico. Esto implica la imposibilidad de que surjan respuestas novedosas porque los sujetos, en sus prácticas, buscan satisfacer el funcionamiento del algoritmo tal cual está. En educación han sido estudiados fenómenos de este tipo como el efecto Pigmalión, el efecto performativo de las etiquetas. Son sistemas que, desde la educación, podemos trabajar para analizar cómo la evaluación se ancla en ciertos cánones, invisibilizando otras formas en las que podrían llevarse adelante esas evaluaciones y la toma de decisiones.



La proyección hacia el futuro es interesante y problemática a la vez porque muestra cómo estos sistemas pueden marcar un "deber ser". La pregunta es si podremos cuestionar estos mandatos. En un sistema que evalúa el trabajo docente, por ejemplo, el sistema valora ciertas dimensiones, y lo que se demuestra es que la respuesta futura de estos docentes es tratar de





producir prácticas y experiencias de manera que sean evaluadas positivamente. Estos mecanismos, que implican una fuerza performativa de la propia tecnología, lo que probablemente hagan es eliminar formas de respuesta innovadoras porque lo que hay allí es un proceso de ajuste a, solamente, lo que el sistema valora como positivo.

Otra de las cuestiones para analizar desde el punto de vista pedagógico es la capacidad de agencias que tienen estos sistemas. Si trabajamos con algoritmos cuyas reglas son definidas a priori por los desarrolladores (es decir, el programador interviene en la definición de ciertas dimensiones que lo constituyen), tenemos la relación de un sujeto que construye ese elemento mediador y que lo utiliza. Allí, claramente, las posiciones están diferenciadas. Pero, en otros casos, el propio algoritmo se encarga, en forma autónoma, de encontrar correlaciones; y, en ese proceso, lo que sucede es que la posición del elemento mediador parece replegarse hacia la posición de sujeto, tomando capacidad de agencia. En educación, lo que tenemos es que, en ciertas ocasiones, el instrumento mediador suele replegarse hacia un lugar de objeto de conocimiento. En un sentido, los medios se transforman en fines y ocurre en las escuelas lo que todos conocemos como pérdida de significatividad social de los aprendizajes. Las discusiones que vienen aparejadas se refieren a quién es el responsable en la toma de decisiones. Porque si bien suele haber intermitencias en donde algunas cuestiones las deciden los sujetos, hay otras veces en las que las decisiones sobre las reglas que se van generando son tomadas por el propio algoritmo (esto es el *machine learning*).

Estos algoritmos, que no están supervisados (en este caso, tampoco me refiero al concepto técnico que se usa en *machine learning*, sino a la supervisión y al control externo que ejerce un programador) y que encuentran correlaciones entre datos e identifican patrones, llevan implícitos la dificultad de trazabilidad de todas estas decisiones que van tomando. Por lo tanto, el algoritmo aparece como una caja negra de difícil inteligibilidad. El lugar del sujeto, analizando e interviniendo allí, es muy complicado. Esto pone sobre la mesa el grado de autonomía en la toma de decisio-



algoritmos que toman decisiones de manera autónoma, sin la intervención de sujetos –al menos para ciertas actividades más simples– y que, en realidad, lo que hacen es reconfigurar la organización del trabajo. Si pensamos en el caso docente, podríamos preguntarnos cómo se podría reconvertir el trabajo si estos sistemas se instituyeran en este sentido. Hay un montón de regulaciones que es necesario discutir entre los agentes del Estado, las empresas que se dedican a su desarrollo, los sindicatos que forman parte de la discusión; pero, claramente, este desarrollo tecnológico lo que hace es poner en evidencia que el control sobre la propia actividad, a veces, podría estar tomado por otros agentes de decisión. La idea de agencia también está muy discutida, y la idea de máquinas y ética también es campo de controversia. ¿En qué sentido un sistema de estas características tiene capacidad de agencia? Incluso, si pensamos en los algoritmos para los autos autotripulados y pensamos en el caso de que uno de ellos esté involucrado en un choque, la responsabilidad, ¿de quién sería? Son todas preguntas que nos llevan a considerar que la tecnología soluciona ciertos problemas, pero genera otros sobre los cuales tenemos que pronunciarnos.

”



Entrevista realizada en el Instituto Superior de Estudios Pedagógicos (ISEP).
Córdoba, marzo de 2019.

