



DIÁLOGOS CIUDADANOS

INTELIGENCIA ARTIFICIAL, ÉTICA Y PARTICIPACIÓN CIUDADANA

DEBATE DE CIERRE

Conclusiones finales

Autores:

Berta Llos, Profesora asociada de Ciencias de la Educación de la UAB

Dr. Carlos Sierra, director de comunicación del CVC

Núria Martínez, técnica de comunicación del CVC



Fundació "la Caixa"

Este documento recoge los aspectos más recurrentes a lo largo de los diferentes ejes de trabajo realizados durante la segunda edición del ciclo de diálogos ciudadanos “[Inteligencia artificial, ética y participación ciudadana](#)” organizado por el Centro de Visión por Computador (CVC), en colaboración con la Fundación “la Caixa”, con el fin de identificar en qué punto se encuentra el estado de la cuestión, y definir cuáles deberían ser los próximos pasos.

La principal conclusión fue que la Inteligencia Artificial (IA) en sí misma no es buena ni mala y, como sucede con cualquier tecnología, el debate social debe centrarse en sus usos. Por ello, una de las principales preocupaciones identificadas fue la necesidad de introducir la ética tanto en los debates, como en los quehaceres tecnológicos. En este sentido, se reivindica la urgencia de incorporar comités de ética en cualquier organización pública o privada que utilice IA. La Administración Pública catalana ya dispone actualmente de algunos recursos en esta línea, como el Observatorio de Ética del Departament de Polítiques Digitals; aun así, los dilemas éticos alrededor de algunos de sus usos están presentes y, por lo tanto, hay que asegurar que los usos que permiten llevar la IA a la práctica deben guiarse por el criterio de la búsqueda del beneficio común, y no para el interés de unos pocos o para la reproducción de las desigualdades sociales. Para ello, la solución debe ser holística y transversal, con un énfasis específico en algunos ámbitos:

1. **Educación y formación**: la educación es una de las prioridades tanto desde la Academia, como desde la Administración pública y buena parte del sector privado. Por un lado, es necesaria una alfabetización de la ciudadanía en términos de IA, asegurando que alfabetizar no implica solo la introducción de conocimiento técnico sino, sobre todo, el desarrollo de un pensamiento crítico, conectándolo con el día a día de las personas, independientemente del nivel formativo, del contexto laboral o educativo en el que se encuentren. En este sentido, se da especial importancia a la introducción del pensamiento crítico a través de la IA en las etapas educativas obligatorias, por lo ya existen algunos programas diseñados específicamente para tal fin, y se reivindica no olvidarse de los sectores de población con más edad o con menos oportunidades de aprendizaje en este ámbito. Por otro lado, se manifiesta la importancia de la formación ética para las personas que se dedican o se van a dedicar a la tecnología, así como el fomento de equipos multidisciplinares en entornos de IA, a través de incorporar perfiles de ciencias sociales o humanidades. Una posible solución sería la incorporación de una asignatura de ética en las carreras técnicas: Física, Ingeniería, etc.
2. **Regulación**: aunque la sensibilización, concienciación ciudadana, y formación en cualquier ámbito son soluciones esenciales, y tienen sentido a medio plazo, se requiere,

para el corto plazo, de otro tipo de soluciones. Una tiene que ver con la regulación y la protección o privacidad de la ciudadanía. A lo largo de los distintos ejes, se ha subrayado en varias ocasiones la sensación de desprotección que siente la ciudadanía ante la IA. En este sentido, se pone sobre la mesa que la regulación debe venir principalmente desde la Administración pública, a la cual se critica por llegar tarde casi siempre, aunque, por otro lado, se argumenta que los procesos de aprendizaje son lentos, y hasta que no se interiorizan, no se puede regular. Aunque se reconoce el esfuerzo legislativo tanto a nivel autonómico, como estatal o incluso desde la UE, a partir de redacción de documentos, protocolos y líneas estratégicas de trabajo, se ha puesto de manifiesto en distintas ocasiones que la legislación no debe tratar solo el establecimiento de derechos, sino ayudar a la ciudadanía a fortalecerse y ejercerlos. Así, aunque la falta de regularización es una necesidad detectada continuamente, “¿Quién vigila al vigilante?”, se manifiesta como reto a partir del cual seguir trabajando para la mejora.

3. **Colaboración entre Academia, Industria y Administración**: tanto en el terreno educativo como en el legislativo, se detecta repetidas veces la clara ventaja de enfrentar las soluciones conjuntamente. En este sentido, se identifica la carencia de una red colaborativa entre Administración, Academia e Industria, y en donde se incluya también la participación ciudadana. Además, se comparte la idea de que, aunque la IA puede ser una herramienta que ayude a superar algunos de los límites, debe ser la propia sociedad quien los establezca. Aun así, se manifiesta que, aunque entre Academia y Administración los mecanismos de relación son más conocidos y utilizados, entre empresa y universidad el intercambio es más complejo. Se habla por una parte de desconfianza, y por otra de falta de cultura de relación.
4. **Derechos humanos e inclusión**: independientemente del nivel de colaboración que se consiga entre los distintos sectores de la sociedad, un criterio unánime en todas las discusiones ha sido el compromiso que tiene que haber en IA para la no vulneración de derechos humanos. Si bien la carta de derechos digitales de la ciudadanía es conocida entre los participantes, esta no lo es entre el común de la sociedad y se identifican paralelamente una cantidad de realidades en donde ciertos grupos de población son discriminados o excluidos a través de la IA. Se habla, en esta línea, de algoritmos sesgados, pero también de bases de datos sesgadas, y finalmente, de personas que al no tener una mirada interseccional, retroalimentan la discriminación en los mecanismos de IA. Así, se reivindica que generalmente se pone siempre el foco en el sesgo de los algoritmos, cuando en realidad, el sesgo está en los datos, y en todo caso los algoritmos

umentan dichos sesgos. Este aspecto vuelve a tener como punto de retorno la educación, y pone de relieve el esfuerzo obligado para avanzar hacia el desarrollo de dispositivos más inclusivos.

5. **Impacto medioambiental:** Otras cuestiones que tienen que ver con el bien común para la sociedad, y que tuvieron un eje específico para ser tratadas, pero que han ido saliendo de manera transversal a lo largo de la edición, es la huella medioambiental. En este sentido, coincide la preocupación común de que el avance en digitalización no implique un impacto negativo en materia medioambiental. Si bien es cierto que existe unanimidad en considerar que los avances en IA no deben pararse porque los beneficios son evidentes en todos los ámbitos, también se pone sobre la mesa que hay que valorar la cantidad de recursos que se gasta y energía que se consume para lograr tales avances. Así, los participantes opinan que, para un óptimo escenario de sostenibilidad de los datos, habrá que saber moverse en un sistema de balanzas, por ejemplo, si se calcula que el gasto de carbono es mayor que lo que se quiera automatizar, el impacto medioambiental sería excesivo. Se sugiere además que el próximo paso es no solo pensar en el propio consumo, sino en el tipo de recursos que se utilicen para este consumo, ya que según si se utilizan energías renovables o no, el impacto para el planeta también es distinto. Por último, se proponen varias iniciativas para regular la cuestión del impacto medioambiental sobre todo en el sector Industrial, se habla de posibles multas para aquellas Industrias que no se comprometan con la disminución de la contaminación, o de inspecciones obligatorias que se tuvieran que pasar en términos de huella medioambiental, aparte de los ya existentes en algunos sectores, comités de ética que son de mucha utilidad, según los participantes, en este sentido.
6. **Desinformación y transparencia:** la reivindicación de hacer a la ciudadanía partícipe y activa en los procesos que tengan que ver con usos de la IA, descrita anteriormente, proviene de una preocupación sobre el desconocimiento general que existe alrededor de la IA, que genera a la vez desconfianza e incomodidad social. Una de las intervenciones que mejor recoge dicha preocupación es la siguiente: si se consigue que la ciudadanía tenga un papel activo y cierta agencia en los procesos, se podría ir cambiando la perspectiva de la IA como una caja negra, a partir de socializar las intenciones y usos de los instrumentos, como en este caso de los algoritmos, y avanzar hacia un modelo de sociedad que no deje al margen a sus ciudadanos de los procesos de cambio y transformación a los que los someta. En este sentido, se da importancia a usar canales que sean de acceso a toda la población y, sobre todo, a asegurar las formas

más transparentes de hacer llegar la información. De esta manera, se afirma que el hecho de que los ciudadanos sepan que está pasando con sus datos, que puedan ver como aprenden los algoritmos, si se produce algún sesgo, etc., puede revertir cierta sensación de interés económico y político que existe, cuando se trata de la utilización de datos personales. Si bien las inciertas intencionalidades que podrían tener los grandes organismos juegan en contra en estos procesos, tanto desde la Administración pública como desde la Academia tienen muy claro que la pedagogía y la transparencia deben ser apuestas firmes. En esta línea, se plantea el escenario del Open Data que, si bien ha generado cierto desacuerdo en algunos aspectos a lo largo de los ejes por sus discutidas ventajas y peligros, hay unanimidad en considerar que posibilitan opinar, participar y criticar todos los usos o finalidades de la IA, por parte de la ciudadanía.

7. **Nuevos empleos**: otra de las cuestiones transversales que ha emergido a lo largo de la edición, pensando en los nuevos escenarios que se presentan, es la cuestión del empleo. Existe un acuerdo entre los participantes sobre que esta cuarta revolución Industrial lleva consigo la eliminación de algunos trabajos habituales hasta el momento, pero que la misma Industria 4.0 generará también una amplia oferta de trabajos, y permitirá nuevas formas de ocupación. Se han propuesto distintas iniciativas para abordar algunas necesidades urgentes en este sentido, el diseño de programas de formación que acompañen a aquellos que pudiesen quedar fuera del mercado laboral, así como estrategias para evitar pérdidas económicas en según qué perfil de activos actuales, o incluso la aplicación de un salario mínimo universal a corto plazo, mientras se llevan a cabo las formaciones para nuevos empleos. Por parte de la Administración pública, se hace referencia al Plan de Capacitación Digital como estrategia vigente para que los trabajadores se vayan adaptando a estos cambios. Aunque a una parte de los participantes es un tema que les preocupa, en general la actitud es optimista, y la mayoría consideran que no se trata de una cuestión de pérdida de empleos, sino de una modificación y transformación de estos.
8. **Falta de inversión y fuga de cerebros**: finalmente, otra de las limitaciones que se ha puesto de manifiesto de manera transversal, sobre todo por parte de la Academia, pero también de la Administración pública, es la cuestión de la falta de inversión en España, en comparación a otros países con desarrollo de la IA y el escenario actual de fuga de cerebros. Por un lado, preocupa que la Industria siempre pueda tener presupuestos mayores para invertir en lo que sea más eficiente, que el propio Sector Público, por lo tanto, las personas más talentosas es fácil que acaben trabajando para las grandes

plataformas digitales y redes sociales y, por otro lado, preocupa también la pérdida de talento de nuestro país por el hecho de no poder ser competitivos en cuestiones de salarios o condiciones de trabajo. Así, se reivindica que necesitamos producir un ecosistema I+D+i, que se centre en conservar el talento que tenemos y se nos va.

9. **Supervisión humana**: aunque se hayan recogido en este documento los retos que sugiere la IA, teniendo en cuenta el momento actual, las oportunidades de desarrollo que conlleva, han sido a lo largo de la edición, extensamente compartidas. Para ello, el principio inquebrantable de cualquier actuación es que los mecanismos de IA sean complementarios a la tarea humana, es decir que sirvan de ayuda más que para sustituir. Hay cierto consenso con la idea que cuando se utilizan los dispositivos no se pretende que estos hagan diagnósticos, y por lo tanto tiene que ser siempre la persona con su criterio quien tome las decisiones. Se acuerda que la IA debe ubicarse más en el campo de la sugerencia que de la decisión, y que, por lo tanto, se trate siempre de una herramienta para el ser humano, no al revés.
10. **Moratoria en el desarrollo de la tecnología**: En varios de los ejes surgió la idea de una posible moratoria para evitar el desarrollo de determinadas aplicaciones de la IA mientras no se haya legislado sobre sus potenciales consecuencias para la ciudadanía. Sin embargo, es una idea que no suscitó una gran aceptación. El sentir mayoritario es que es un hecho que la tecnología va más rápida que la legislación, pero que la solución no puede ser parar el desarrollo, sino buscar formas de hacer avanzar más rápidamente la legislación.