



# DIÁLOGOS CIUDADANOS

INTELIGENCIA ARTIFICIAL, ÉTICA Y PARTICIPACIÓN CIUDADANA

INFORME | EJE 2

## Algoritmos sesgados: Sin dejar a nadie atrás

**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"

# [19/01/2021] – SEMINARIO Y FOCUS GROUP

## 1) Seminario (17:00 -18:00 h) – Impartido por la Dra. Maria Vanrell

Enlace al seminario: <https://www.youtube.com/watch?v=T8i8ZDWB3b4&t=3s>

DIÁLOGOS CIUDADANOS  
INTELIGENCIA ARTIFICIAL, ÉTICA Y PARTICIPACIÓN CIUDADANA

19/01/2021 | 17:00 h | ONLINE

**SEMINARIO**

**Algoritmos sesgados: sin dejar a nadie atrás**

A CARGO DE  
Dra. **Maria Vanrell** | *Centro de Visión por Computador  
Universidad Autónoma de Barcelona*

Organizan:

www.iabcn.cvc.uab.es

## 2) Focus group: moderado por Berta Llos

- Participantes: 12 personas (7 mujeres y 5 hombres)
- Tuits de resumen del seminario (realizados por los participantes del focus group):

## Tuit de la ponencia

Mentimeter

Los algoritmos sesgados son resultado de una sociedad desigual. No es un reto tecnológico, es un reto social.

La inteligencia artificial es tan sesgada como la humana, fijando protocolos éticos podemos avanzar en justicia social.

El peligro de aplicar un algoritmo, un riesgo no controlado como sociedad

Presentación didáctica y clarificadora sobre los algoritmos sesgados. deja abiertos interrogantes sobre como regular y supervisor estas potentes aplicaciones.

Ya que los algoritmos están sesgados... ¿Se podrían crear algoritmos que discriminaran positivamente a determinados colectivos?

La inteligencia artificial, aunque avanzada, no es capaz de discriminar comportamientos racistas / homófobos / etc... ¿cómo lo solucionamos?

Detrás de un algoritmo siempre hay una persona

Los algoritmos de IA son cajas negras, no sabemos qué resultados van a dar a partir de una entrada y el resultado es una aproximación estadística, no garantizan que en un caso concreto acierten o no.

En la IA estamos buscando un espejo de nuestras habilidades y defectos o un apoyo complementario para mejorar y crecer? El problema no es el ordenador, otra vez es quien lo utiliza, sus prejuicios y/o su ignorancia.

## Análisis del focus group:

*\*El foco del debate, según los participantes, no se encuentra tanto entre discernir si el sesgo es del algoritmo o los datos, sino en la definición e intencionalidad de los usos.*

### **OPORTUNIDADES DE DESARROLLO**

La **neutralidad de la tecnología** y, por ende, de los algoritmos, es una cuestión no resuelta entre los representantes de la ciudadanía del presente eje. Si bien la falta de acuerdo emerge sobre todo por terminologías y definiciones, se produce cierto consenso en el sesgo inherente de la persona que trabaja con algoritmos y, por tanto, se manifiesta la necesidad de cuestionar el reconocimiento de desigualdades por parte de la gente que maneja los algoritmos, así como sus intereses y usos en cada caso, más que el carácter bueno o malo de los datos *per se*.

Así, la **importancia de las comisiones de ética**, por una parte, y la **transparencia en el uso de los datos** por otra parte, tanto de la administración pública como de las empresas privadas, emergen como dos de las principales actuaciones en las que, si se consiguen avances, podrían representar una oportunidad de desarrollo en el campo de la IA, y concretamente, cuando se utilizan algoritmos. Un buen ejemplo para reflejar dicha condición se pone de manifiesto: disponer de numerosos datos sobre colores de piel, puede servir para avances en el tratamiento del cáncer de piel, o para investigaciones que tienen que ver con la cosmética, pero también pueden conllevar sesgos racistas. Por lo tanto, lograr cierta transparencia en el uso de los datos, y el obligado cumplimiento de criterios éticos, que podrían estar definidos por las comisiones de cada organización, ayudaría a aumentar la confianza por parte de la ciudadanía.

De la mano de las comisiones éticas, y con la preocupación generalizada de no contribuir a la reproducción de las desigualdades a través de los algoritmos, otras propuestas emergen a lo largo del diálogo conjunto. En primer lugar, se insiste en distintas formas de supervisión, disponer de sistemas que puedan controlar el trabajo de los algoritmos se hace imprescindible, y más después de conocer la función actual de los algoritmos a través de redes neuronales, en donde es más difícil identificar donde puede encontrarse una discriminación, ya que el comportamiento está distribuido en todas las neuronas artificiales, y por este motivo, se apunta con mayor urgencia, al avance en términos de supervisión humana. En la lógica de mejora de los sesgos que pueden implicar una discriminación, se sugieren propuestas como la de añadir sesgos en el sentido contrario, para intentar revertir las desigualdades; se emplaza para el debate la viabilidad de una actuación de este tipo.

Por último, un aspecto que podría, según varias de las personas representantes, ayudar al desarrollo del uso de los algoritmos con una perspectiva de justicia social, consistiría en la alfabetización de la ciudadanía en términos de IA. Promover procesos de **participación ciudadana** como el que se encuentran los mismos participantes en el momento de la discusión, de forma más generalizada y organizado por parte de la administración pública. 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.

### **LÍMITES**

El principal límite que, con consenso, no se puede cruzar en el uso de los algoritmos, es la **falta de regulación**. Son muchas las evidencias que se ponen de manifiesto a lo largo del diálogo conjunto y que apuntan al miedo que produce pensar en el control de los datos sobre la ciudadanía que se

puede tener, sobre todo por parte de empresas privadas, al no estar regular como pueden tratarse o no los datos, y cuáles deben ser disponibles o no. Se cuestiona de esta forma la inexistencia de metodologías que garanticen la equidad de los datos, ya no solo en el testeo sino también en la autorización. Tiene que quedar muy claro quien autoriza el uso de los datos, qué bases se consideran públicas, y con qué intencionalidad se pueden utilizar o no.

Se subraya en varias ocasiones la **desprotección de la ciudadanía** ante la IA, y concretamente en este caso con el uso de algoritmos, y parece que los documentos existentes como la “Carta de derechos digitales” del gobierno español no es muy esperanzadora por lo que refiere a empoderamiento de la ciudadanía. En este sentido se apunta que la concepción de las administraciones sobre el poco control que pueden tener de los datos de la ciudadanía, ya que estamos constantemente cediéndolos en nuestras acciones diarias, y eso acaba naturalizando que se cedan, con lo que las empresas, según varias opiniones del grupo, se aprovechan de esta lógica. Por lo tanto, los límites, en este sentido, no pasan solo por regular los usos, sino que consisten en abordar la índole política, que afecta a la legislación y, en consecuencia, al replanteamiento más crítico al que se deben someter los documentos que se producen, como es el caso de la “Carta de los derechos digitales”. Con todo, existe transversalmente una preocupación sobre como de atrás queda la administración con relación al ritmo en el que avanza la tecnología, y se pone de manifiesto cómo esta **falta de adaptación de ritmos** puede tener consecuencias en contra la ciudadanía.

Y finalmente, ligado a los argumentos anteriores, se identifica cierta **desconfianza en las intencionalidades de los grandes organismos**, en cuanto al uso de los datos. La premisa que debería regir sobre su uso cuando promueve la mejora para la calidad de vida, o cuando se trata de fomentar el bien común, queda en duda, al repasar entre el grupo de participantes, algunas de las actuaciones más flagrantes, en términos de discriminación o incluso vulneración de derechos, que han llevado a cabo importantes organizaciones. En este sentido, las diferencias entre administración pública o academia y las empresas privadas se identifica como uno de los límites más peligrosos, y es que, aunque se manifieste la importancia de regularizar o de trabajar para la transparencia, en el sector privado los participantes consideran que es mucho más difícil la intervención, y cuestionan si existe alguna forma de que haya más control y supervisión en lo que este sector refiere.

## [26/01/2021] – DEBATE

### Participantes:

- Dra. **Maria Vanrell**, investigadora del Centro de Visión por Computador (CVC) y profesora de la Universidad Autónoma de Barcelona (UAB)
- Dra. **Marta Ruiz Costa-Jussà**, investigadora de la Universidad Politécnica de Cataluña (UPC)
- Sra. **Mayte Hidalgo**, responsable de estrategia y ética de la IA en everis
- Sr. **Daniel Santanach**, coordinador de la estrategia de IA de Cataluña (Catalonia.AI) en el Departamento de Políticas Digitales y Administración Pública de la Generalitat de Catalunya.
- Sra. **Judith Membrives**, mentora e investigadora de Experiencia de Persona Usuaría e Innovación en el UXLab de Barcelona Activa (*portavoz del focus group*).

Moderado por: Dr. **Carlos Sierra**, director de comunicación del CVC.

- Biografías de los ponentes: <http://iabcn.cvc.uab.es/algorithmes-esbiaixats/>
- Vídeo del debate: <https://www.youtube.com/watch?v=jkkQgPrbwvs&t=2s>



## Aportaciones del debate:

### **Intervenciones iniciales:**

Maria Vanrell (representante de la academia): La tecnología no es neutra, como nada lo es si está hecho por la humanidad. No hay que tener miedo a los avances. La tecnología digital hay que recibirla igual que la tecnología en otros ámbitos. Ejemplos como la tecnología farmacéutica, y urgencia en la introducción de protocolos sobre los usos.

Daniel Santanach (representante de la administración pública): Es importante diferenciar un sesgo de un error. La Inteligencia Artificial tendrá tantos sesgos como los que tienen las personas. Algunos se pueden neutralizar, pero otros son innatos. Aportación literaria sobre los sesgos innatos que no controlamos, a través de la obra de Daniel Kahneman "Pensar rápido, pensar despacio" y la de Yuval Noah Harari "Sapiens: de animales a Dioses". Aun así, según el portavoz, la administración no tiene que preocuparse por el sesgo sino por sus efectos. Por eso la prioridad es avanzar hacia la justicia, por una parte, con lo que afecta a la parte reguladora (leyes y normas) y, por otra parte, poniendo a la disposición de la ciudadanía instrumentos o herramientas para que se puedan utilizar, y por tanto asegurar transparencia en los procesos. Se pone, además, como práctica que lleva a cabo la administración, el ejemplo de la creación reciente del observatorio ético del Departamento de Políticas Digitales

Marta Ruiz (representante de la industria, también academia): Los avances tecnológicos van más rápido que lo que podemos asumir los humanos, y además gran parte de la sociedad no está familiarizada con la IA, por tanto, es esencial la educación y formación en tecnología, la información sin educación no sirve para nada. Por otra parte, hay que plantearse como queremos que sea la

tecnología, ya que todos los datos que producimos sirven para entrenar a nuestros sistemas, y la producción de datos no se puede parar. Por último, una ventaja de la tecnología frente los humanos, que destaca la portavoz, es que a un algoritmo le puedes hacer un “reset”, puedes reentrenarlo y, por tanto, ¿podría por ejemplo llegar a ser más equitativo un algoritmo que los propios humanos? Hay que valorar las oportunidades, dice.

Mayte Hidalgo (representante de la industria): Estamos inmersos en un cambio de paradigma, pasamos de un ecosistema de toma de decisiones tradicional, a un sistema de decisiones automatizadas. El reto más grande es pensar como limitan algunos aspectos éticos, por ejemplo, cuando iniciemos cualquier proceso de IA se tendría que incorporar un capítulo ético obligatoria, para asegurar que no haya ninguna parte de la población que salga perjudicada. No hay que caer en la reproducción de las desigualdades. Por otra parte, somos generadores de datos y ya que actualmente son, irrevocablemente, una expresión de la humanidad, hay que aprovecharlo y que sirvan para el bien común. En este sentido, hay que educar para el entendimiento sobre el uso de los datos. También, hay que eliminar los pánicos producidos por los medios de comunicación. Así, esta generación de datos no enriquece solo al sector empresarial, sino que también ayuda a la riqueza comunitaria.

Judith Membrives (representante de la ciudadanía): Existe un sentimiento de desprotección por parte de la ciudadanía por varias razones. Por una parte, existen unas relaciones de poder en el diseño de las tecnologías y la sensación es que hay intereses políticos o económicos detrás. Por otra parte, se está de acuerdo con el empoderamiento de la ciudadanía, pero la alfabetización es muy lenta y mientras tanto habría que disponer de más estructuras de regularización. Finalmente, sería interesante, considera la ciudadanía, incorporar más la parte de la ética en los procesos, por ejemplo, desde la perspectiva de los derechos humanos.

### **¿El algoritmo o los datos? ¿Dónde hay que poner el foco?**

Se identifican posturas que argumentan no poder separarse una cuestión de la otra, se pone el foco en la transparencia de los datos, en la urgencia de disponer de datos abiertos para que todo el mundo pueda saber cómo se ha alimentado el algoritmo, y que por tanto se pueda opinar, participar y criticar. Por otro lado, las representantes de la industria consideran que el sesgo está en los datos, y en todo caso el algoritmo puede aumentar este sesgo. Por ende, hay que poner el esfuerzo en regular el tratamiento de los datos. En este sentido, la portavoz de la ciudadanía apunta que el problema es que la supervisión humana se hace post-test, a partir ya de unas consecuencias, y se reivindica que este control tendría que estar desde antes.

Una intervención desde la voz de la academia sobre lo problemático que puede ser el hecho de los intereses económicos privados en el uso de los datos abre una discusión sobre la necesidad de asegurar una transparencia. También desde la ciudadanía se reivindica que la mayoría de las compañías privadas no quieren explicar cómo funcionan sus algoritmos. Se ofrecen varios ejemplos en esta línea, por parte de la industria y la administración pública, como el de incorporar en los data sets unos “chips” como los de los móviles, para que quede en abierto como son los datos, de donde vienen, etc. También se comparten experiencias de ciudades como Ámsterdam o Helsinki en donde ya se cuenta con datos en abierto para toda la ciudadanía, y el portavoz de la administración explica que en Barcelona se están planteando opciones semejantes.

Por otra parte, se manifiestan propuestas que podrían plantearse para la preocupación en los usos de las empresas privadas. Una sería obligar a pagar impuestos a las grandes corporaciones para el uso de datos, y a partir de las ganancias de estos impuestos, invertir en educación tecnológica, por ejemplo. Si bien la administración está de acuerdo en poner un coste a los datos para quien quiera

usarlas, apunta que no hay que preocuparse tanto por las corporaciones más grandes, por las que todo el mundo conoce, ya que estas son las más controladas y tienen que ser más transparentes, pero hay otras empresas menos importantes, que también recogen muchos datos y aquí sí que se necesita actuar.

La cuestión de la transparencia lleva a la importancia de la educación, a partir de una intervención del portavoz de la administración en donde señala que aparte de la transparencia es importante que la ciudadanía confíe en el algoritmo, y para que estos dispositivos de inteligencia artificial generen comodidad a la gente, se necesita educar en este ámbito. Se habla de un material didáctico que se está creando desde una universidad catalana para la difusión de conocimiento sobre los algoritmos, a todo tipo de población. En este sentido, se expresa también que hay una cuestión de prioridades políticas, y se asocia por parte de la industria, en este caso, el hecho de no enseñar tecnología o programación en los centros educativos como una decisión política que tiene sus consecuencias graves en la falta de alfabetización de la población. La portavoz de la ciudadanía se suma a ser crítica con las prioridades políticas, pero piensa que en el sistema educativo no se puede hablar de incorporar IA aún, si los niveles de desigualdad social son tan grandes que no hay un ordenador en cada casa. Por tanto, un primer paso puede ir más encaminado al desarrollo de un pensamiento crítico, más que a la introducción de un conocimiento técnico sobre la creación de algoritmos. Aun así, se subraya la importancia de generar debates públicos y que se visibilicen en medios de comunicación, por ejemplo.

### **Mientras dura el lento proceso de educar a la sociedad, ¿qué hacemos?**

La administración pública se posiciona en este sentido claramente a favor de la creación y actualización de leyes que regulen todos los procesos, y reivindica la falta de competencias que se tiene en Cataluña para reformar la legislación. Por otra parte, pone como buen ejemplo la carta de derechos digitales, la cual también se ha hecho en versión contextualizada del contexto catalán. Se explica que esta carta está abierta a la ciudadanía a través de una plataforma virtual, para que se pueda participar y proponer. Una intervención aquí apunta que las declaraciones de intenciones están bien pero que hay que traducirlas a consideraciones éticas. Y otra aportación cuestiona como se pueden materializar las grandes líneas que se abordan en la carta.

La academia sugiere, en relación con la carta y sus posibilidades, u otras formas de regularización, que no se trata de saber de dónde vienen los datos, sino que la ciudadanía pueda conocer a qué conclusiones llegan sobre sus datos, y que perfil se construye de ellos a partir de estas conclusiones. En este sentido, la administración pone el ejemplo de una iniciativa impulsada pero que de momento no se ha podido poner en práctica por restricciones a nivel estatal, que consistiría en una etiqueta de cada persona sobre sus datos, en la que el ciudadano podría consultar el origen de los datos y editar la etiqueta en caso de haber errores.

Se plantea, en contraposición, que los modelos de datos en abierto pueden ser peligrosos. Un ejemplo de ello es la producción de *fake news*, así como otros elementos a nivel ético que se deberían reflexionar. En este sentido se apunta que no hay que olvidar el contexto de sociedad capitalista en el que vivimos y, en consecuencia, los datos como un activo más a explotar. Finalmente, pues, se llega al consenso que para la tecnología no es una posibilidad, hay que asumir que existe y, por ende, priorizar tanto los avances legislativos como la apuesta en educación.

## ENLACES DE INTERÉS

- **Crónica del debate** | *Algoritmos y sesgos: aún hay tiempo para minimizar sus consecuencias y construir una IA más justa y representativa:*  
<http://www.cvc.uab.es/outreach/?p=2950>
- **Vídeo del debate:** <https://www.youtube.com/watch?v=jkkQgPrbwvs&t=2s>
- **Galería de imágenes del debate:**  
<https://www.flickr.com/photos/189365245@N08/sets/72157718056335808/>
- **[EN LOS MEDIOS] Expertos piden educar a la ciudadanía sobre los sesgos de los algoritmos** | Artículo publicado en La Vanguardia | 26/01/2021 | Disponible en:  
<https://www.lavanguardia.com/vida/20210126/6200597/expertos-piden-educar-ciudadania-sobre-sesgos-algoritmos.html>



## **ANEXO:** Preguntas y temas a tratar en el debate (extraídas del focus group del 19/01/2021)

---

### **CIUDADANÍA**

- ¿Cómo se puede concienciar a la población de la importancia que tiene la cesión de datos?
- ¿Como se puede hacer extensivo el conocimiento de los algoritmos y de cómo influyen en nuestras decisiones?
- ¿Se incorporará al marco de Competencia digital de los ciudadanos, conceptos y criterios sobre IA?
- ¿Qué compromiso (público o privado) de retorno a la sociedad se tiene con las investigaciones que se llevan a cabo utilizando los datos de todos/as?
- ¿Tenéis ejemplos de participación ciudadana o de educación/alfabetización de la población en este ámbito, para empoderarlos en sus acciones diarias?
- ¿Conocéis la “carta de los derechos digitales”? ¿Qué opinión tenéis al respecto?

### **REGULARIZACIÓN**

- Para la administración: ¿Qué papel debe asumir como ente regulador?
- Administración: ¿Cómo garantizamos parámetros de justicia social en los procesos administrativos basado en inteligencia artificial/algoritmos? ¿Mediante qué criterios o protocolos?
- ¿Hay algún sistema para testar si los sistemas de IA cumplen los objetivos establecidos?
- ¿Se piensa incorporar los criterios y regulaciones de "Seguridad en el Diseño" en el desarrollo de aplicaciones de IA?

### **ÉTICA**

- Para la industria: ¿Cómo implicarse para diseñar algoritmos para el bien común?
- Industria: ¿Cómo garantizamos un compromiso ético en el desarrollo de modelos de negocio basados en algoritmos? ¿Mediante qué mecanismos?
- ¿Dentro de qué campo se define el estudio de la ética en IA? ¿Se está esperando de profesionales técnicos, una opinión ética, y en caso afirmativo, qué voz y poder se le da a esa opinión?
- ¿Como piensan adoptar las soluciones de "inteligencia artificial", bajo qué criterios éticos y de control?
- ¿Podrías dar algún ejemplo dónde el uso del algoritmo sirva para el desarrollo, teniendo en cuenta la justicia social y, por tanto, donde no se reproduzcan o incluso se promuevan ciertas discriminaciones?

### **DATOS**

- Academia: ¿Cómo podemos hacer que los datos con los que se entrenan los algoritmos sean más neutros?

## OTRAS CUESTIONES

- A la academia: ¿Por qué no potencia la transdisciplinariedad de una manera real? no sólo retórica en este ámbito y en todos los demás, claro.
- A la academia: ¿Hasta qué punto es necesaria la colaboración con empresas para tirar hacia delante la investigación tecnológica?
- A la administración: ¿De qué manera se puede reequilibrar el avance?
- A la industria: ¿Dado que el desarrollo de productos se nutre de la investigación pública y el contenido abierto, hasta qué punto es justo que se evite la transparencia con respecto a los datos y los algoritmos?
- ¿A que estarían dispuestos a renunciar para el bien de la evolución tecnológica (Poder, beneficios, libertad, ...)?
- ¿Qué poder le damos al algoritmo, y qué poder queremos darle?
- ¿Tendría sentido añadir sesgos en el sentido contrario para revertir desigualdades?