

COVID-19 Y DATOS: LECCIONES A TENER EN CUENTA PARA EL DISEÑO DEL HORIZONTE 2030.

Informe elaborado por Emilio Soria Olivas.

Catedrático de Universidad. Departamento de Ingeniería Electrónica (ETSE) – Universitat de València.



IDAL (Intelligent Data Analysis Laboratory) - <http://idal.uv.es>  IDAL
Intelligent Data Analysis Laboratory

València, 8 de junio de 2020

PROSPECT
COMUNITAT VALENCIANA

2030



Introducción.

El COVID19 ha supuesto un auténtico revolcón para la Humanidad, un virus que surge en China a finales del año 2019 ha conducido a un contagio mundial impactando directamente en las diferentes sociedades como nunca antes. En estos días se ha hablado de crisis económicas (por ejemplo la del petróleo de finales de los 70 o el *crash* del 2008) pero ninguna ha sido como ésta en cuanto a cantidad de países implicados y en cuanto a profundidad. Además han aparecido otras diferencias importantes como son:

1. Se han puesto a prueba los sistemas sanitarios de todos los países; tanto en recursos materiales como personales. De estas pruebas de estrés se deberían analizar las conclusiones derivadas del *análisis objetivo de los datos*.
2. Se ha tenido que variar la forma de impartir clase a todos los niveles. Este hecho ha demostrado que nuestro país tiene capacidad tecnológica para poder dar soporte a este tipo de formación y, segundo, que es posible llevarla a cabo.
3. Las empresas se han tenido que adaptar a trabajar de forma remota. No me gustaría hablar de teletrabajo en estas situaciones donde se tiene a toda la familia confinada; bajo mi punto de vista no es teletrabajo. Independientemente de su nombre se ha puesto también de manifiesto que es posible trabajar de forma remota y mantener la operatividad de la empresa con la tecnología existente actualmente en nuestro país.

Esta pandemia conlleva unas consecuencias importantes; todos los países se preparan para un tremendo golpe económico cuyas consecuencias son muy difíciles de predecir y, todavía más, de solucionar dada la total ausencia de precedentes de crisis de esta magnitud. Este informe es una visión muy particular de esta crisis; una visión centrada en los datos y su análisis, llamémosle como queramos; *Data Mining*, *Business Intelligence* o el que está más de moda y que, posiblemente, sea el menos usado en nuestro país: *Inteligencia Artificial (IA)*. Las tecnologías basadas en los datos y el análisis de la información han sido claves en esta pandemia de la misma forma que están siendo claves para el desarrollo económico de los países. El mejor ejemplo, China, ha pasado de ser un país eminentemente agrícola con una industria a merced de países exteriores (el famoso *Made in China* junto con el *Designed in California*) a ser actualmente una de las grandes potencias (hay un serio debate si ya no es la primera) en Inteligencia Artificial¹. ¿Cómo lo ha conseguido? Pues con una hoja de ruta relativamente sencilla pero que conlleva cambios sociales (hay que educar a la población de la bondad de estas tecnologías); económicos (hay que irse adaptando a la era de la información de

¹ K.F. Lee. *Superpotencias de la Inteligencia Artificial*. Deusto 2020

los datos y “saltar” de era) y políticos (hay que situar los datos en el centro ser, de verdad y no en plan lanzar eslóganes políticos, *Data-Driven*).

Antes de centrarnos en el uso de los datos durante el COVID-19 es importante centrarnos en las tecnologías derivadas del dato y por qué su *boom* actual.

Datos e IA: ¡siameses!

Durante los últimos tres años en tecnología han habido tres palabras que han resaltado; uso los términos ingleses que son los más extendidos: *Big Data*, *Data Science* y *Artificial Intelligence*. Y lo que poca gente sabe es que los tres están directamente conectados. Tienen una relación extremadamente fuerte; la actual Inteligencia Artificial no podría existir sin los otros dos; no existe independencia entre ellos. ¿Pero qué tiene que ver los datos con la IA?. La Inteligencia Artificial tiene sus raíces hace muchos siglos, es un objetivo que se ha planteado desde siempre el poder sustituir nuestra capacidad cognitiva por una máquina² sin embargo los avances en los últimos 5 años han sido espectaculares. La aproximación que, actualmente, es la que implementan nuestros sistemas de IA es la que se conoce como *Data-Driven*. Nuestros sistemas de AI son modelos estadísticos extremadamente complejos con un gran número de parámetros (serían los tornillos a ajustar para que todo funcione como toca). Al igual que cualquier otro modelo matemático para ajustar esos parámetros se necesitan datos y ésta es la clave de estas tecnologías actualmente: *sin datos, no tenemos IA*. Y aquí hay que romper otro mito: *la clave de la IA no son los algoritmos, ni las máquinas que procesan los algoritmos ni el software que lo realiza; ¡son los datos!*. Hay una frase que aparece en el libro de *Kai-Fu-Lee* que recoge esta dependencia: *China es actualmente el Arabia Saudí de los datos* y ahí está su valor y una de las razones de su posición actual de potencia dominante en la IA.

Así pues el origen son los datos por lo que empezaremos con el *Big Data*; datos masivos, correspondientes a datos que varían muy rápidamente, presentan un gran volumen y tienen una alta variedad. Si se piensa en *Facebook*, las diferentes páginas cambian rápidamente añadiéndose texto/imágenes/vídeos (que presentan un gran volumen) y todos ellos con diferente formato. El bajo precio de los dispositivos de almacenamiento y adquisición de datos ha hecho que el volumen de éstos crezca sin parar. Igual hay un lector que recuerde el microordenador *Spectrum* de 16K, muy famoso en la década de los 80, actualmente cualquier dispositivo portátil tiene un millón de veces más memoria y con un mismo precio si hacemos una correspondencia por el nivel de vida. *No hay nada en tecnología que haya evolucionado tan rápido y de forma tan económica*, y, por tanto, tan accesible para cualquier usuario. Además del almacenamiento el procesamiento de datos ha

² N. Nilsson. *The Quest for Artificial Intelligence*. Cambridge University Press, 2013.

seguido la misma evolución³. A estos dos factores se le suma el procesamiento en la nube o *cloud computing*. Grandes compañías, *Amazon, Microsoft, Google e IBM*, ofrecen grandes servidores para ser usados como máquinas de proceso y análisis de datos durante tiempo limitado por un precio muy económico. Este hecho ha supuesto que pequeñas empresas, *start-ups*, disponen de una capacidad de cálculo propia de grandes empresas por un precio muy bajo; este hecho *ha democratizado el hardware para todo el mundo*⁴. Junto con este crecimiento exponencial de las capacidades del *hardware* (y un decrecimiento del mismo tipo en el precio) se ha producido un *todo gratis en el software relacionado con el tema del análisis avanzado de datos*. Se tiene que los programas más avanzados que analizan datos (actualmente *R* y *Python* y sus librerías más avanzadas de IA como puede ser *Tensorflow* uno de los buques insignia de *Google*) son gratuitos. Y hablando de código destacar que los algoritmos más avanzados de IA se encuentran disponibles de forma gratuita para todo el mundo. Pero no solo eso (por si teníamos poco.....) los propios modelos de IA están disponibles para ser usados de forma gratuita; por ejemplo, están disponibles reconocedores de imágenes, segmentadores de vídeo en tiempo real, generadores de lenguaje natural, incluso modelos ya entrenados para ser usados en imágenes clínicas y detectar COVID pero de eso ya hablaremos. Y para cerrar el círculo se pueden modificar estos modelos de una forma relativamente sencilla y rápida para adecuarlos a nuestros propósitos con una cantidad mínima de datos y de recursos. Por ejemplo puedo usar un modelo de segmentación de imágenes y usarlo para determinar la distancia entre personas en un determinado recinto (¿le suena al lector alguna aplicación de esto?) y todo ello con un equipo mínimo de personas en un tiempo muy pequeño y con poco dinero. Siguiendo el tema del gratis total se tiene otro elemento clave para el desarrollo de cualquier área tecnológica: la información. En este punto destacar que, actualmente, los avances de la IA están disponibles *de forma gratuita* en un gran número de sitios web pero quiero destacar tres: a) la página web *arxiv.org* donde se publican los trabajos que se van a enviar en los próximos meses/semanas a las revistas (o congresos) más importantes de IA; b) los sitios web de las grandes tecnológicas donde muestran su poder en estas áreas, como ejemplo se tienen <https://openai.com/> y <https://deepmind.com/research> y c) la cantidad descomunal de *blogs* y *newsletter* donde aparece la información sobre cómo desarrollar modelos de IA así como el código para realizarlo; a modo de ejemplo se tiene <https://towardsdatascience.com/>.

Si lo juntamos todo se tiene que *las tecnologías de IA no son caras, al contrario, actualmente es una de las tecnologías más baratas que existen*.

Otro punto a tener en cuenta en todo el ecosistema del dato es la irrupción de las tecnologías 5G que van a permitir transferencias de datos a una velocidad/volumen nunca vistos anteriormente. Proyectos que actualmente se están desestimando (relacionados muchos de ellos con

³ A. Agrawal, J. Gans, A. Goldfarb. *Máquinas Predictivas. La sencilla economía de la inteligencia artificial*. HBR, 2020.

⁴ M. Boden. *Inteligencia Artificial*. Turner, 2017.

procesamiento en tiempo real de los datos) se podrán implementar y desarrollar en muy poco tiempo.

¿Y cómo llevamos a cabo un proyecto de IA?; para su aplicabilidad se necesitan tres elementos:

1. *Una pregunta a resolver.* Es la clave, aquí entra el factor humano, el experto es el que tiene que definir qué problema se va a resolver.
2. *Unos datos asociados.* Tienen que estar relacionados con la pregunta que se quiere resolver.
3. *Un análisis avanzado de datos.* Última etapa, aquí tendremos desde técnicas clásicas estadísticas hasta las técnicas más avanzadas de IA (dependiendo de la calidad y la cantidad de los datos).

Como se aprecia, lo comentado es extremadamente general de tal forma que su aplicabilidad es enorme, no hay campo del conocimiento o actividad productiva donde no se pueda aplicar. Pondré tres ejemplos para no extenderme demasiado pero seguro que al lector se le ocurre alguno más:

1. *Sanidad.* Uno de los campos donde se está empezando a aplicar la IA de forma cada vez más extensa. Recordemos que es uno de los sectores que más recursos consume y que es necesario optimizar porque, por lo que apuntan las cifras, dejará de ser sostenible dentro de unos años. Estas técnicas de IA permiten, por una parte optimizar la calidad de vida del paciente y, por otra, optimizar los tratamientos; no es una utopía, podemos llegar a una mejor sanidad y con un menor coste. Algunas preguntas a resolver: ¿qué pacientes tienen problemas con un determinado tratamiento?; ¿qué necesidades voy a tener el próximo trimestre?; ¿qué dosis de fármaco debo administrar a este paciente?; ¿tiempo de estancia en el hospital?.
2. *Agricultura/Ganadería.* Otra de las áreas importantes en nuestra comunidad. Es una de las áreas claves para la humanidad (nuestra alimentación depende de ellas) pero donde no se aplica análisis de datos de forma generalizada. Su aplicación (bajo el término de *agricultura de precisión*) mejoraría enormemente la rentabilidad en este sector económico. Algunas preguntas a resolver: ¿cuándo debo regar el campo?; ¿qué características van adquiriendo mis cultivos con el tiempo?; ¿qué abono y en qué cantidad debo usar?; ¿cuáles son las características de alimentación de mis animales para optimizar mi producción?
3. *City Brains*⁵. El concepto de *Smart Cities* está superado, más que nada porque se han quedado en “*cidades con sensores*” pero no se han llegado a generalizar modelos

⁵ <https://www.alibabacloud.com/solutions/intelligence-brain/city>.

predictivos que mejorarán la vida del ciudadano de forma real. Este concepto viene, ¡como no!, de China y es muy recomendable la visión de este vídeo⁶. Algunas preguntas interesantes a resolver en este campo serían: ¿se prevén atascos para mañana?; ¿cuál es la duración de los semáforos en la próxima hora?; ¿debo desviar el tráfico?; ¿qué niveles de contaminación tendré mañana?; ¿a qué hora debo regular el alumbrado de acuerdo por zonas y características del tiempo?.

Antes de ver cómo podemos plantear la implementación de estas tecnologías pasemos a ver cómo esta pandemia nos ha mostrado nuestras debilidades en el tema del almacenamiento, procesado y análisis de los datos.

COVID-19 y Datos.

Lo que sigue no quiero que se considere un comentario político partidista porque, independientemente de quién gobernaba, ha sido general en todas las administraciones españolas el tremendo desastre en la gestión del dato relacionado con este tema. Se ha comprobado la poca importancia que se le ha dado a tener datos limpios, de calidad, centralizados, con posibilidades de hacer predicciones con ellos de forma rápida, con posibilidad de compartirlos para que la comunidad de científicos de datos existente en España realizara modelos para ayudar a tomar decisiones. Con datos limpios y disponibles para ser cruzados entre sí se podrían haber establecido muchos sistemas predictivos de ayuda a la decisión funcionando *en tiempo real*. Hemos asistido a un espectáculo lamentable de cifras que implicaban muertos, cifras que implicaban usos de recursos sanitarios, cifras reflejadas de encuestas realizadas sin ningún control, modelos epidemiológicos muy precisos pero que fallaban de forma sonrojante cuando salían las verdaderas (¿?) cifras y que rápidamente se escondieron debajo de la mesa. Y todo ello teniendo un impacto en la sociedad terrible. Aquí no se critica la metodología utilizada ni la capacidad de los investigadores ya que todos ellos han tenido el mismo problema: la falta de cantidad y calidad de los datos. A modo de ejemplo los modelos epidemiológicos tienen una serie de parámetros a ajustar (como los modelos de IA) que usan los datos para obtenerlos, si los datos son erróneos los parámetros obtenidos serán falsos y, por tanto, cualquier predicción no será correcta. Debemos plantearnos si nos podemos permitir la gran heterogeneidad en los datos (por ejemplo diferentes hospitales diferentes formatos), que aparezcan trabas al acceso de los datos (no puede ser que existan “dueños” de esos datos); tener diferentes sistemas para cada una de las Consellerías y, sobre todo, una escasez casi absoluta de análisis de los datos que se tienen.

⁶ https://www.youtube.com/watch?v=v4_2QuS4Xns

¿Qué podríamos haber hecho con unos buenos datos, con la cantidad suficiente y estando disponibles para implementar modelos?; algunas ideas:

1. *Herramientas de predicción de necesidades sanitarias.* No todos los hospitales han sentido la misma carga en estos días; si hubiéramos tenido datos se podrían haber planteado sistemas predictores de necesidades de recursos y, por lo tanto, se podrían haber distribuido los recursos (tanto materiales como personales) de la mejor manera. Por ejemplo con los datos de estos días se podrían haber planteado modelos de estimación de tiempo en planta/UCI que podrían haber ayudado a la optimización de las UCIs. Si hablamos de otros elementos clave, los respiradores, podríamos haber planteado un modelo predictor de la necesidad de estos sistemas. Aparte están los sistemas de predicción de diagnóstico que podrían haber ayudado en los sistemas de unidades de emergencias de los hospitales; se podrían haber implementado sistemas basados en imágenes/texto (historias clínicas) que podrían haber ayudado a nuestros clínicos. Aquí también entrarían los diferentes modelos epidemiológicos que hemos visto estos días pero que no han podido usarse por la falta de datos correctos.
2. *Control del estrés en la población/profesionales clínicos.* Los primeros días de confinamiento conllevaron un alto estrés para muchas personas; herramientas de escucha en redes sociales podrían haber dado un buen termómetro a las autoridades sobre lo que estaba pasando y cómo se percibía por la población. Atención especial merecen en este punto los profesionales clínicos que soportaron en primera persona esta pandemia; conocer su nivel de estrés podría haber ayudado en gran manera a darles soporte a ellos.
3. *Análisis automático de la información textual.* Estos días han surgido una cantidad enorme de análisis/estudios científicos que podrían haber ayudado a los clínicos en su trabajo. Un buscador *inteligente* que analizara información textual en tiempo real podría haber filtrado tratamientos usados con éxito en otras partes del mundo o bien destacar determinadas reacciones a los fármacos usados. Otra aplicación de este análisis de la información textual es un servicio de noticias para la población que, automáticamente, fuera seleccionando todas aquellas cuyo asunto principal estuviera relacionada con el COVID-19.
4. *Difusión de la información.* La ausencia de información veraz en estos días ha supuesto en muchas personas un problema. Hemos visto cuadros de mando (*dashboards*) muy buenos⁷ y que podríamos haber desarrollado para nuestra población; la información hay que darla en crudo pero también hay que facilitar mecanismos para que se pueda interpretar. Si a

⁷ <https://coronavirus.jhu.edu/map.html>

esta herramienta de visualización se le añaden modelos de predicción la población sería más consciente de la difusión de esta pandemia y sería más responsable. En este grupo se pueden añadir las herramientas para evitar las noticias falsas (*fake news*) que pueden llegar a tener un efecto demoledor sobre la población llegando, incluso, a provocar disturbios si no se paran a tiempo.

5. *Control de distancias en espacios/transportes públicos.* El mantener una distancia de seguridad ha sido crítico en las diferentes etapas de esta pandemia. Existen modelos⁸ ya desarrollados que podrían haberse usado para ser implementados en transportes/recintos públicos para controlar si se mantenía la distancia de seguridad, ¿el coste del modelo?...cero euros.
6. *Desarrollo de Chatbots avanzados.* Actualmente el campo del procesado del lenguaje natural está sufriendo un incremento de aplicaciones y mejora en sus modelos como nunca antes se había visto. Estos sistemas podrían haberse usado como elementos de apoyo en llamadas al 112 para complementar a los operarios. Existen *Chatbots* que podríamos calificar como elementales que se han usado en esta crisis pero se ha desestimado usar otros más avanzados y que no tenían (como se viene comentando a lo largo de este informe) un coste desorbitado.

Cualquiera de las aplicaciones que he comentado más arriba no costarían más de dos semanas a un buen equipo de científicos de datos con unos buenos datos en cuanto a calidad/cantidad. Este equipo podría depender directamente de Consellería lo que daría más validez a sus actuaciones de cara a la ciudadanía que, en definitiva, son los que tienen que dar sus datos (nadie vería un interés económico en sus acciones al no entrar organismos privados). Alguien puede pensar que es muy fácil dar soluciones *a posteriori* pero lo que planteo lo han acometido otros organismos, países, anteriormente a nosotros y podríamos haber copiado las soluciones:

- Un equipo de investigadores del Hospital Infantil de Boston han desarrollado una aplicación para rastrear la propagación del coronavirus (*HealthMap*), el sistema integra datos de búsquedas en *Google*, redes sociales y *blogs*. Estas fuentes de información no se suelen utilizar pero son muy útiles para evaluar la respuesta pública [A. Johnson, How Artificial Intelligence is Aiding the fight Against Coronavirus, Datainnovation, March 13, 2020](#)).
- *Alibabá* entrenó un sistema de IA para reconocer coronavirus con una precisión que se afirma es del 96%. Según la empresa, el sistema podría diagnosticar un coronavirus en 20 o 30 segundos, mientras que la misma operación normalmente

⁸ <https://www.pyimagesearch.com/2020/06/01/opencv-social-distancing-detector/>

le llevaría a un médico experimentado 10 o 15 minutos. El sistema ha ayudado al menos a 26 hospitales chinos a revisar más de 30.000 casos ([C. Li, How DAMO Academy's AI System Detects Coronavirus Cases, Alizila, March 10, 2020](#)).

- El Centro Internacional de Investigación sobre Inteligencia Artificial (IRCAI) de Eslovenia, ha puesto en marcha una vigilancia "inteligente" de los medios de comunicación sobre el coronavirus denominada *Corona Virus Media Watch*⁹, que proporciona actualizaciones de las noticias mundiales y nacionales basadas en una selección de medios de comunicación.

Solo muestro una pequeña muestra de cómo se han usado los datos en otros países y recomiendo al lector buscar en la web donde encontrará gran cantidad de ejemplos¹⁰

¿Y ahora?

Tenemos una sociedad y una economía por reconstruir y todos nuestros pasos deben estar bien fundamentados porque tenemos una oportunidad única de aprovechar esta crisis (al igual que hizo China) de salir muy fortalecidos pero hay que ser muy cuidadosos con las decisiones tomadas. Se hace necesario unos buenos cimientos y, sobre todo, preguntar el *por qué* y el *para qué* se hacen las cosas. Sin estas dos preguntas, básicas, podemos perder muchos recursos económicos y personales (fuga de talento) que, ahora mismo, no nos podemos permitir. Posible caminos para posicionarnos en este campo que será determinante en los próximos años:

1. ***Plantear una Agencia del Dato.*** Estamos en la era del dato y la información; no es posible que se tengan los datos en diferentes Consellerías con formatos diferentes y sin posibilidad de análisis en tiempo real. A modo de ejemplo, ¿se ha intentado cruzar los datos de dos Consellerías? seguro que no. Sin una buena gobernanza del dato es imposible sacarle partido a dichos datos. Volviendo a nuestras preguntas; *¿por qué?* porque sería un elemento innovador a nivel nacional y permitiría unificar los datos para usarlos para los ciudadanos evitando que esos datos se conviertan en *rehenes* de determinadas personas/organismos. *¿Para qué?* para optimizar los servicios de todo tipo y hacer más fácil la vida a los ciudadanos al tiempo que se consigue un gran ahorro así como una transparencia como nunca antes se ha tenido.
2. ***Desarrollar un sistema de formación en IA/Datos.*** Uno de los factores de la explosión en China ha sido la formación en IA y datos. Actualmente hay grados en las Universidades de la Comunidad relacionados con estos temas pero se hace necesario inculcarlo desde la primaria y cimentarlo en la secundaria. Volviendo a las preguntas *¿por qué?* Porque de la misma manera que se plantea el inglés como herramienta indispensable el conocer las herramientas asociadas a la IA lo es ya además que seremos referentes en este sentido.

⁹ <http://coronaviruswatch.ircai.org/?country=All&dashboard=news>.

¹⁰ <https://covid19challenge.mit.edu/> <https://academicdatascience.org/covid> <https://www.oecd.ai/covid>

¿Para qué? Para incubar un ecosistema de IA desde la escuela que permita a nuestros hijos incluirse en el desarrollo que se puede hacer en estos temas desde la Comunidad.

3. **Fomentar la creación de start-ups basadas en IA.** Actualmente existen incubadoras privadas de este tipo de iniciativas pero una apuesta pública por este campo permitiría tener una participación en ellas de tal forma que habría un retorno al sector público además que crearía una imagen de marca innovadora en la Comunidad. *¿Por qué?* Porque este tipo de empresas son las que actualmente más beneficios económicos están generando al tiempo que parte del conocimiento se puede quedar en la administración. *¿Para qué?* Para aumentar la capacidad tecnológica de la Comunidad teniendo efecto llamada sobre otras regiones apareciendo, si se hace bien, una realimentación positiva.
4. **Fomentar la IA en regiones deprimidas .** Por lo comentado anteriormente el lector habrá deducido que *la IA es una de las tecnologías más democráticas que existen y se pueden desarrollar en cualquier lugar.* Este hecho permite que se puedan plantear *polos de talento en IA* en cualquier lugar con las consecuencias que este hecho tiene para lo que se conoce como *España despoblada*. Volviendo a las preguntas; *¿por qué?* Porque tenemos un serio problema de despoblamiento en zonas rurales y estos polos tecnológicos pueden servir de atracción de talento tanto nacional como internacional. *¿Para qué?* Para volver a generar vida en pueblos abandonados gracias a las últimas tecnologías.
5. **Fomentar la IA en aplicaciones propias.** Hay sectores de la IA que empiezan a despuntar claramente y donde un posicionamiento claro y definido podría ser un referente para otras regiones/países. Entre estas aplicaciones se encuentran la Agricultura/Ganadería y la Salud. Establecer un ecosistema de IA centrado en estas aplicaciones tendría mucho impacto económico para la Comunidad; *¿por qué?* Porque el área de la Salud Pública es la que más se puede optimizar usando estas técnicas y porque las grandes empresas de IA están virando a este campo por la gran cantidad de beneficios que se puede conseguir (las grandes farmacéuticas han hecho una apuesta clara por estas técnicas). Si hablamos de la Agricultura/Ganadería tenemos la misma situación con la ventaja adicional que podría instalarse en una zona deprimida como se comentaba en el punto anterior. *¿Para qué?* Para posicionarnos en dos campos que tirarán del campo de la IA en los próximos 5 años de forma clara.
6. **Ser pioneros en la Inteligencia Humana-Artificial.** Actualmente hay un profundo debate de los problemas de la IA y como ésta encaja con los humanos. Se está investigando en problemas éticos, filosóficos, de sesgo y explicabilidad en los modelos, etc. La Comunidad tiene investigadores trabajando en estos campos desde hace años. Está claro que mientras estos problemas estén abiertos no se tendrá una aplicabilidad de estos modelos al 100%. *¿Por qué?* Porque hay importantes preguntas por responder en el tema de la IA y serán claves para su aplicación; *¿para qué?* Para situar a la Comunidad en uno de los puntos esenciales tecnológicos en los próximos años.

A modo de conclusión; está claro el papel que van a jugar los datos y la IA en los próximos años y estamos a punto de perder el tren. Tenemos que funcionar de forma coordinada, usando todos los recursos, asumiendo el papel de líder el sector público (dado el tema tan sensible de los datos) y actuando en todas las líneas (formación/desarrollo/innovación). Si actuamos descoordinados, fomentando actividades privadas pero no públicas y empujando una actividad de investigación teórica (y no de desarrollo aplicado) podremos perder el tren de la industrialización de la IA donde se están invirtiendo tantos esfuerzos actualmente.