Comentarios al Libro Blanco sobre el enfoque a la excelencia y la confianza

En los puntos 3 y 4 se habla de excelencia científica europea. Parece que se plantea como una línea tutelada, financiada, supervisada por la UE o alguno de sus organismos. Creo que ése no es el papel de las instituciones públicas, nacionales o comunitarias. EEUU es líder en la tecnología y no es gracias a una política tecnológica dirigida desde el gobierno, sino gracias a políticas fiscales sobre mecenazgo, proyectos importantes y facilidad de creación de empresas. Pretender liderar el un Hub de datos sería como convertirse en mediador del mercado vegetal, como si teniendo cocos y raíces de jengibre se pudieran sacar nuevos productos. Los datos, los conjuntos de datos, son tan variados como los granos de arena. Cada empresa sabe qué necesita sin que un organismo público medie su trabajo. Cada empresa tiene derecho a publicar, vender o ceder los datos que quiera y a quien quiera, ya que una vez liberados los datos se pierde su control: los datos son pasivos, no avisan de quién los usa ni para qué. Una entidad puede compartir ciertos datos con sus socios, pero no va a ceder algunos datos a la competencia o a posibles inversores buitres porque le perjudicaría. Un hub de datos o un centro de investigación de excelencia a nivel europeo o de los países miembros, creados por ley, no funciona, no tiene lugar en una economía de mercado. Basta mirar a países occidentales para ver qué centros de investigación son relevantes. La mayoría, universidades, algunos gigantes tecnológicos contra quienes competir sería un suicidio económico, y unos pocos transnacionales como la ESA o el LHC. Ese modelo excepcional no es aplicable de modo generalizado en el caso de la IA. Instituciones como la universidad de Oxford, el Instituto Alan Turing, el instituto Ada Lovelace, el instituto HAI de la universidad de Stanford, el instituto AI Now de la universidad de Nueva York son modelos más asequibles y con una producción científica y un impacto social y en las políticas eficaz y más rentable.

El documento afirma que Europa continuará guiando el progreso en los algoritmos de IA construyendo sobre su propia excelencia. La realidad es que la excelencia en tecnologías tradicionales no se transmite automáticamente a otras, ni a la IA, como demuestran las start-ups de éxito, excelentes en IA y sin tradición previa.

La mayor parte de los estudios universitarios sobre IA se han centrado en programas master. Creo que es importante ampliar, con el impulso necesario, la creación de grados y promover el uso aplicado de la IA en las demás áreas de conocimiento universitarias y profesionales. La enseñanza de IA en la escuela a partir de 10-12 años es uno de los pilares del futuro de la IA en China y EEUU.

Además, la Universidad suele ser uno de los mejores centros de investigación si se dispone de medios económicos. También las empresas, grandes, pequeñas, start-ups o consolidadas. Creo que es importante apoyar tanto la investigación básica como la aplicada. Fomentar la creación de foros, congresos, simposios, etc., que puedan servir también de baremo del cumplimiento y justificación las ayudas. El acceso a éstas debe ser simplificado

para no limitar el acceso, pero exigente en cuanto a resultados. Preferible créditos a bajo interés y largo plazo que ayudas a fondo perdido que aumentan la deuda y fomentan los proyectos sin utilidad.

Los Digital Innovation Hubs (DIH) corren el riesgo de ser el destino mayoritario de las ayudas y de convertirse en organismos burocráticos sin límites presupuestarios. Construir un modelo de IA es de una escala distinta al LHC. El software para crear IA es open-source en su mayoría, está en internet, en papers, o es el objeto de la misma investigación. Ninguno de los gigantes de internet (Apple, Microsoft, Facebook, Amazon, IBM,...) se han creado desde un DIH. Como ejemplo, SpaceX, fundada en 2002, es una compañía privada líder en lanzamientos aeroespaciales. En solo dieciocho años ha logrado una tecnología de reutilización de componentes, y economía de fabricación no lograda por ninguna otra mucho más experimentada. Ni de capital privado como Boeing o Martin Marietta, ni público como NASA o Arianespace. SpaceX o Tesla, la compañía más valorada en bolsa, no nacieron o fueron apoyadas por un DIH, sino con dinero, financiación, y con un líder visionario excepcional, igual que Apple, Microsoft, Amazon o Facebook. Esas iniciativas, con proyectos atractivos y financiación adecuada, son las que atraen talento.

Llegar a acuerdos con el sector privado se sugiere como uno de los medios para alcanzar la excelencia, punto 4.E. La realidad es que la iniciativa dinámica de creación de riqueza proviene en Europa del sector privado. La investigación pública sin su aplicación por el sector privado carece de sentido.

En el punto 5 al hablar de la confianza se omite hablar de la equidad (fairness) y de los valores culturales y religiosos. Este artículo https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3547922 asegura que la equidad computacional o procedural no es lo mismo que el valor de equidad defendido en Europa. Quizás la forma de incorporar esa equidad sea mediante la supervisión humana durante la operación de los sistemas inteligentes. De ese modo queda también cubierta la responsabilidad (liability) en todas las fases de la operación.

Entre los problemas de confianza está el de las crecientes asimetrías derivadas del poder otorgado por la acumulación de datos. Es un problema tanto si la acumulación se produce en el sector privado como en el público. No se cita ninguna medida para atacar este problema.

Los deep fakes son también una sorprendente omisión en el documento. Las deep fakes son una aplicación perversa de la IA que mina la confianza en la sociedad, elemento esencial de la convivencia, igual que la falta de equidad. Es un problema a caballo entre la seguridad, la IA y la defensa de los valores de la democracia y de la fama de las personas y entidades. Pueden usarse como arma de guerra. Al mismo tiempo muestran la ínfima robustez de los sistemas inteligentes, lo que es una reafirmación de que la IA es un software sobrevalorado.

Siempre he defendido que los sistemas de corrección ética son necesarios cuando afectan a personas, pero creo que también son necesarios cuando afectan a instituciones. El tejido

económico no puede quedar al arbitrio de una IA no regulada. Esto es especialmente necesario en el sector público, donde los procedimientos posiblemente automatizables con sistemas inteligentes aplican tanto a personas jurídicas como físicas. La justicia es un derecho tanto de unas como de otras. Tomado separadamente, el sector público es muchísimo mayor en todos los países que cualquier otro sector económico. La concesión de permisos, subvenciones, créditos, investigación de impuestos, etc., supone un riesgo. La vida, la economía, la política está directamente afectada por algún organismo público. Creo que es el sector estratégico más importante y donde el impacto sería mayor. Y quizás donde tendría más impacto y en al que se debería prestar especial atención a la ética. Por eso la responsabilidad y la regulación deberían ser mayores. A modo de ejemplo, Indonesia plantea sustituir el nivel inferior de los cuatro de su administración pública por robots https://www.thetechnolawgist.com/2019/12/04/indonesia-impulsa-un-plan-para-sustituir-a-funcionarios-por-robots-administrativos/

El documento asegura que la falta de requerimientos en la IA dificultan el trazado de las decisiones tomadas en los sistemas inteligentes y que sería difícil obtener compensación por los daños. Creo que, dada la complejidad de los sistemas simples y de sistemas complejos con múltiples subsistemas inteligentes, la probabilidad de encontrar una explicación en una controversia es imposible, igual que es imposible en un juicio saber qué pensaba el autor de un crimen, o porqué un perro bien entrenado decide en una ocasión hacer algo que provoca un daño. En consecuencia, pretender recoger pruebas para intentar explicar lo no explicable es un gasto sin sentido, dar esperanzas infundadas de una justicia tecnológicamente fundamentada y unos costes forenses inabordables. La explicabilidad es una característica que se requiere en los sistemas inteligentes éticos, pero este documento la ignora. Así como también que los sistemas basados en deep learning no son explicables. La explicabilidad, además, consiste en que cualquiera pueda entender cómo se ha tomado una decisión, algo que con la tecnología actual es difícil o imposible. Pretender explicar el comportamiento de un sistema a partir del entrenamiento es como pretender explicar el comportamiento de un animal amaestrado: el entrenamiento no justifica el comportamiento posterior. Así, en vez de centrarse en las condiciones previas, sería más productivo centrarse en las consecuencias y quienes son responsables, sea cual sea la tecnología empleada.

Propongo que haya una regulación específica sobre sistemas inteligentes, y quien quiera protegerse podrá recopilar pruebas para un eventual conflicto. Quien las recoja y quiera adherirse a un esquema de compliance será bienvenido y mejorará la confianza, pero esa certificación no debería significar exoneración en caso de demanda, ya que lo que se debe proteger es a las personas físicas y jurídicas. La supervisión humana (human oversight) previa a la operación no garantiza la completitud de los posibles errores, por eso siempre debe estar complementada con la correspondiente responsabilidad sobre los resultados (liability). Por tanto, creo que poner el foco en una sobrecarga de la preparación es limitar la innovación y posiblemente debilitar la responsabilidad.

Creo que es insuficiente limitar la regulación a las aplicaciones de “alto-riesgo”, ya que hay muchas otras amenazas por descubrir. Nadie podía adivinar en el conocido caso de la publicación de los datos anónimos de uso de Netflix que podían ser identificados los

usuarios. Hoy día sabemos que las técnicas de anonimización no sirven https://www.vice.com/en_us/article/dygy8k/researchers-find-anonymized-data-is-even-less-anonymous-than-we-thought _Otras técnicas como privacidad diferencial son más eficaces en la prevención de filtración accidental de datos, personales, industriales, corporativos o de cualquier tipo. No estoy convencido de que el enfoque basado en riesgo sea el enfoque adecuado para determinar qué deberá estar sujeto a regulación.

Se mencionan otros casos de alto riesgo como el reclutamiento y la identificación biométrica, pero no se indica una regla clara de cuáles son esas excepciones, quién es la autoridad que debe resolver los conflictos o imponer la regulación ¿algún observatorio? ¿un sitio público de demandas?

El mantenimiento y posible inspección de datos así como de los algoritmos y pruebas puede eventualmente plantear problemas de privacidad y de propiedad industrial. Cuestión que es ignorada en el Libro Blanco.

En el apartado de robustez y precisión se habla de la “ex-ante due and proper consideration of the risks”, pero no se sabe qué es una consideración debida y apropiada, ni quién dice que lo es.

Un sistema de sellos de calidad por tecnologías o por sectores o por riesgos facilita mucho la comprensión de los usuarios y la confianza, como actualmente la identificación de lugares con cámaras de vigilancia, el símbolo de pago por tarjeta sin contacto o los usos posibles del menaje de cocina con las distintas tecnologías de calentamiento (gas, vitro, inducción...). Los signos visibles, fácilmente interpretables e intuitivos son una ayuda. Un ejemplo de signo confuso, equívoco y de dudosa utilidad, excepto para la recaudación, es el etiquetado eco de los coches en España.

En resumen creo que este Libro Blanco es poco concreto, incompleto y que plantea unos objetivos inalcanzables, una regulación regulación insuficiente, y unas cargas y una intervención pública excesivas y unas expectativas fuera de las previsiones de los expertos en la materia.

La IA no es más que software sobrevalorado y, a la postre, lo que hay que proteger es a las personas, las entidades, los valores, la cultura y afrontar las responsabilidades en caso de incidentes. Exactamente lo mismo que hace el sistema judicial, pero en nuevos escenarios. No creo que sea necesario poner más cargas procedurales, burocráticas o legislativas, sino ampliar lo que hay a los nuevos escenarios. Y no creo que deba haber una intervención pública activa en la IA, sino fomentar el desarrollo, la formación y la financiación de iniciativas mediante líneas de crédito, programas de investigación y facilitar el mecenazgo. La experiencia y éxitos de la ciencia en general y de la IA en particular en nuestro entorno cultural liberal y de mercado nos dice que el papel estatal debe ser el de facilitador, no el de mediador y mucho menos interventor.

Aniceto Pérez y Madrid, Filósofo de las Tecnologías, Especialista en Deep Learning, Especialista en Ética de Datos, Ingeniero Superior de Telecomunicación, PDD por el IESE y Empresario.
aperezymadrid@gmail.com