

SÍNDIC
EL DEFENSOR
DE LES
PERSONES



IOI - Europe
International Ombudsman Institute
European Chapter

Intel·ligència artificial i drets humans

reptes, rols i eines
de les institucions
d'ombudsman

2 i 3 de març de 2020
Disseny Hub Barcelona

Amb la col·laboració de:



Barcelona
Supercomputing
Center
Centre Nacional de Supercomputació

Amb el patrocini de:



Dilluns, 2 de març de 2020

Inauguració 5

Rafael Ribó, president europeu de l'IOI i síndic de greuges de Catalunya

Pablo Martín, sotssecretari del Ministeri de Ciència i Innovació

Àngels Chacon, consellera d'Empresa i Coneixement

Laia Bonet, tercera tinenta d'alcaldia de l'Ajuntament de Barcelona

Peter Tyndall, president de l'IOI i Ombudsman de la República d'Irlanda

Realitats i límits de la intel·ligència artificial 15

Moderator: Andreas Pottakis, director de l'IOI i ombudsman de Grècia

Ulises Cortés, catedràtic i investigador titular de la Universitat Politècnica de Catalunya, Barcelona Supercomputing Center (BSC)

Montserrat Pardo Bayona, directora d'affers corporatius de Microsoft España

Alfonso González Herrero, director de Government & Regulatory Affairs, IBM Spain

Oliver Smith, director d'estratègia i cap del Departament d'Ètica a Alpha Health

Carina Lopes, cap de la Digital Future Society Think Tank

De la societat analògica a la digital 41

Moderator: Genís Roca, expert en internet

José María Lassalle, director del Foro de Humanismo Tecnológico d'ESADE i professor de Filosofia del Dret d'ICADE

Daniel Innerarity, catedràtic de Filosofia Política de la Universitat del País Basc; Ikerbasque Foundation of Science; European University Institute of Florence

Renata Ávila, directora executiva de la Fundación Ciudadano Inteligente

David Cabo, director de la Fundació Civio

Dimarts, 3 de març de 2020

Drets, reptes: Aproximació sectorial 53

Modera: Esther Paniagua, periodista independent especialitzada en tecnologia

Ciberseguretat Ángel Gómez de Ágreda, coronel d'aviació, cap de l'Àrea d'Anàlisi Geopolític, DICOES (Secretaria de Política de Defensa, Ministeri de Defensa)

Recerca en salut i bases de dades Josep M. Argimon, director gerent de l'Institut Català de la Salut

Programes socials i no-discriminació Liliانا Arroyo, experta en transformació digital; Institut d'Innovació Social (ESADE)

Bioètica Itziar de Lecuona, subdirectora de l'Observatori de Bioètica i Dret, càtedra UNESCO de Bioètica de la UB

Infància i adolescència José María Sánchez Bursón, consultor

Drets, reptes: Tractats internacionals i definicions 75

Modera: Miquel Molina, director adjunt de *La Vanguardia*

Gregor Strojín, president del CAHAI (Consell d'Europa)

Martha Stickings, analista de polítiques, Agència de Drets Fonamentals (FRA)

Geraldine Mattioli-zeltner, assessora a la Comissària Europea de Drets Humans, Consell d'Europa

Peter Bonnor, jurista principal de l'Ombudsman Europeu

Reptes, rols i eines de les institucions d'ombudsman 89

Modera: Joan Manuel del Pozo, síndic de la Universitat de Girona

Rob Behrens, ombudsman parlamentari i del sistema sanitari del Regne Unit

Elisabet Rynning, ombudsman parlamentària de Suècia

Quirine Eijkman, comissària i vicepresidenta de l'Institut Neerlandès de Drets Humans

Eric Houtman, ombudsman federal belga de l'Energia i membre de la NEON

Conclusions i eines 105

Modera: Catherine De Bruecker, vicepresidenta europea de l'IOI i ombudsman federal de Bèlgica

Ulises Cortés, catedràtic i investigador titular de la UPC, Barcelona Supercomputing Center (BSC)

Daniel Marco, director general d'Innovació i Economia Digital

Rafael Ribó, president europeu de l'IOI i síndic de greuges de Catalunya

Intel·ligència artificial i drets humans

reptes, rols i eines de les institucions d'ombudsman

Dilluns, 2 de març de 2020

Inauguració

Pablo Martín, sotssecretari del Ministeri de Ciència i Innovació

Àngels Chacon, consellera d'Empresa i Coneixement

Laia Bonet, tercera tinenta d'alcaldia de l'Ajuntament de Barcelona

Peter Tyndall, president de l'IOI i Ombudsman de la República d'Irlanda

Rafael Ribó, president europeu de l'IOI i síndic de greuges de Catalunya

Rafael Ribó

President europeu de l'IOI i síndic de greuges de Catalunya

Benvinguts a les jornades de l'IOI sobre intel·ligència artificial i drets humans.

Confesso de principi que soc un ignorant sobre intel·ligència artificial, però m'espanta perquè trobo poca gent experta en drets humans i intel·ligència artificial. I la nostra tasca, la dels ombudsman, són els drets humans. Em deia la nova secretària d'estat de Digitalització que s'ha de construir una matriu de riscos. I hi ha qui diu que hem de buscar algorismes ètics, i jo aquí ja em començo a perdre. També hi ha qui parla d'algorismes il·liberals. I sobretot tots em diuen que els algorismes provoquen biaixos, especialment biaixos d'igualtat de gènere. Contra la igualtat de gènere. I, per tant, és evident que hem de fer una pressió, unes polítiques, unes mesures, una sèrie d'accions concretes, que pel que fa als marcs legals més solemnes són urgents.

Parlant amb l'ambaixador espanyol a la delegació del Consell d'Europa, per a nosaltres l'òrgan més sagrat de defensa dels drets humans, amb el text més sòlid de la planeta en drets humans, que és el Conveni europeu, em deia: "Hem de reformar els tractats". Però penseu que el Consell d'Europa crea el Comitè Ad Hoc sobre Intel·ligència Artificial el 2019. Aquí a la cantonada. Setembre de 2019. Les estructures nacionals de drets humans també hi estan treballant. Veritablement, cal alguna cosa que ens permeti a tots, començant per les administracions, no només la matriu de riscos, sinó mesures com les que, per exemple, anuncia el Govern de Catalunya, l'anomenat CIDAI.

Els ombudsman debatem constantment sobre les millors pràctiques. I, com sabeu, no m'agrada aturar-me en les millors pràctiques. Prefereixo parlar dels drets humans. I al mig, entre aquestes qüestions, apareix la paraula ètica. La primera vegada que em vaig reunir amb una gran empresa de GAFAM em van dir: "Estem creant un comitè d'ètica". I jo els vaig dir: "No, hauríeu de crear un comitè de drets humans". Hem de saber què és l'ètica, perquè de manera comparada hi ha molts enfocaments d'ètica, i hem d'abordar l'ètica realment des del que es descriu, el que realment hi ha en la definició, fins i tot dins de les portes dels drets humans. Els defensors ens trobem amb problemes com la targeta de crèdit per a prestacions socials, els deures fiscals d'aquestes empreses, els subministraments de serveis essencials a través d'algorismes, els comptadors intel·ligents, les obligacions per a tots els ciutadans europeus ... Molt biaix a través d'algorismes i sobretot discriminació intersectorial, vulnerabilitat múltiple a través de la discriminació intersectorial. Realment hauríem reformar els tractats. Els ombudsman, si volem jugar un paper atacant la mala administració, buscant casos de bones pràctiques, però sobretot amb referència als drets humans.

Jo faria tres apel·lacions molt solemnes: la primera, a tots els meus col·legues ombudsman. Crec que avui fem el primer pas com a col·lectivitat europea i potser mundial per endinsar-nos en aquest món de la intel·ligència artificial i els drets humans. I la primera apel·lació és: defensors d'arreu, no podem deixar aquest tema amagat, ignorat, que ens atemoreixi i que no coneguem els seus límits. És un gran repte, que és la segona apel·lació a tota

la societat, és un gran repte democràtic. I són necessaris, com algú ha dit, ciutadans capaços digitalment per pensar i per criticar. No els hem d'informatitzar només per posar-los en una anella més de la cadena. Que siguin capaços de pensar i de criticar aquesta realitat. Per tant, una segona apel·lació a tota la ciutadania, especialment aquells (alguns parlaran avui i demà aquí) que van en primer línia de participació a desemmascarar on ens poden fer mal als drets humans. I la tercera apel·lació, ja l'he insinuat abans, legisladors, parlamentàries i parlamentaris, organismes internacionals, correu, és urgent reformar els tractats. Us ho diu un ciutadà de Catalunya, un ciutadà de Barcelona. Barcelona presumeix de ser un dels més importants *hubs* digitals d'Europa. M'agradaria que algun dia presumeixin, presumíssim, de ser un dels més importants *hubs* digitals en drets humans.

Moltes gràcies.

Pablo Martín

Sotssecretari del Ministeri de Ciència i Innovació

Bona tarda consellera, síndic, tinent d'alcalde, president de l'Institut Internacional de l'Ombudsman, autoritats, acadèmics, algunes cares conegudes. Senyores i senyors, en primer lloc, voldria disculpar l'assistència del ministre Pedro Duque. Estava previst que assistís avui a aquest acte, a aquesta inauguració, i que fos present en aquesta conferència tan interessant, però lamentablement qüestions d'última hora no li han permès ser aquí i em toca a mi representar-lo de la millor manera possible.

No és casualitat que tingui lloc a Barcelona, una ciutat molt compromesa amb el coneixement i la tecnologia, una ciutat coneguda internacionalment per l'excel·lent ciència que s'hi fa i les enormes capacitats innovadores que aglutina. N'és un exemple el centre nacional de supercomputació, la Barcelona Supercomputer Center, o l'estructura científica Sincrotró Alba, per esmentar-ne algunes.

Pel que fa a l'objecte d'aquesta trobada, ho comentàvem abans en el previ, la intel·ligència artificial és una cosa que en pocs anys ha passat de ser una àrea d'investigació i de computació a condicionar-nos tota la nostra vida, a ser present en la nostra vida d'una manera clara i per a sempre.

Sense haver preparat potser la societat adequadament, ja està modelant la realitat a nostre voltant, els nostres treballs, la nostra manera de relacionar-nos i de comunicar-nos, els nostres sistemes de salut, la manera d'ensenyar, la manera d'aprendre i molts altres aspectes que a mi se m'escapen ara.

En la història hi ha casos ben coneguts de tecnologies que han canviat les nostres societats. Puc esmentar-ne dues, i n'hi haurà moltes més sens dubte. El vapor, introduït inicialment per al transport, però ràpidament aplicat en molts altres sectors; o l'electrificació, que va automatitzar processos industrials i va canviar les nostres ciutats, els hàbits socials i els serveis. Aquests exemples van significar enormes canvis, van alterar la sort econòmica dels llocs que es van desenvolupar i després de tot el planeta. Almenys en les seves primeres etapes tots van generar guanyadors i perdedors. En molt poc temps es va generalitzar l'accés universal i immediat al coneixement, gràcies a la interconnexió de tots els ordinadors en inventar-se els protocols adequats que ara anomenem *Internet*. La quantitat de dades disponibles s'ha multiplicat extraordinàriament els últims pocs anys, incloent-hi enormes volums d'informació aparentment trivial, però que tractats amb les tècniques adequades i amb els accessos a capacitats de processos, també enormes, permeten treure conclusions inquietants, inexplicables. Potser qui té accés a dades i processos i conclusions es troba en una enorme posició d'avantatge.

Si les conclusions d'aquest procés massiu són encertades poden ajudar a guanyar més i més diners i desplaçar la resta del mercat. Si són esbiaixades, es poden produir tota mena d'injustícies sense que ningú sembli tenir-ne la culpa, ni ser-ne responsable.

Per tant, el debat avui dia gira al voltant de com s'ha de regular el desenvolupament i l'ús de la intel·ligència artificial, de determinar què és el que s'ha de regular i de quina manera es fa perquè el seu impacte sigui el benefici col·lectiu.

Durant els últims mesos han augmentat els fòrums que plantegen aquests dubtes. Alguns advoquen per un compromís d'autoregulació sense aprofundir en excés en principis ètics ni en mecanismes concrets, d'altres debaten sobre els marcs i els principis ètics que han de regir. Una activitat que ha generat multitud de manifestos que promouen una intel·ligència artificial ètica.

És molt necessari que continuem creant entorns de debat públic, que afavorim la comunicació sobre les implicacions socials i ètiques d'aquestes noves tècniques d'emmagatzematge i processos d'informació, fòrums amb perfils diferents i sensibilitats diferents que permetin facilitar acords col·lectius sobre els perills, beneficis i oportunitats que comporten.

Vaig acabant perquè suposo que el debat més interessant és el que ve a continuació, però voldria reafirmar que estic segur que en aquest congrés sobre intel·ligència artificial i drets humans els debats seran molt intensos i fructífers i que, després que finalitzin, tindrem una millor idea de com regular col·lectivament aquesta nova realitat i com gestionar entre tots l'interessant futur que ens espera.

Moltíssimes gràcies i sobretot moltíssima sort, que la necessitem.

Gràcies!

Àngels Chacón

Consellera d'Empresa i Coneixement

Good afternoon, buenas tardes, molt bona tarda, senyor Rafael Ribó, Mr. Peter Tyndall, senyor Pablo Marín, senyora Laia Bonet. Afegeixo també les salutacions al senyor Carles Grau, director de la Mobile; també al senyor David Ferrer, que avui ens acompanya, secretari de Polítiques Digitals, un departament en el qual també treballem tots aquests temes; la senyora Joana Barbany, directora general de la Societat Digital, i la senyora Elisabet Nebrera, directora general d'Afers Socials. També evidentment als col·laboradors i esponsoritzadors, sense els quals no es podria fer aquesta jornada, i evidentment als ponents.

El primer ho esmentava el senyor Ribó, som conscients que som incompetents sobre aquesta matèria i crec que és el primer pas pel qual hem de passar, per reaprendre.

Moltes gràcies per convidar-me a formar part d'aquesta inauguració, de la inauguració d'aquesta jornada sobre intel·ligència artificial i drets humans. I el mateix plaer que tenim que se'ns convidi com a autoritat és un obstacle perquè ens puguem quedar al que realment seria un plaer, que seria quedar-nos a tota la jornada i aprendre.

En primer lloc, deixeu-me agrair al síndic de greuges i a la resta d'ombudsmen i ombudswomen, perquè també ens calen dones, i aquí saludo especialment a la senyora Catherine De Bruecker, perquè també ens calen dones en aquest àmbit, la vostra necessària tasca en defensa de les persones i dels drets. I a aquesta preocupació pel benestar de tots els ciutadans també cal incloure les reflexions que es plantegen en actes com el d'avui.

El que es discuteix en aquestes jornades és com dibuixar entre tots (ciutadans, administracions, empreses) un futur on les noves tecnologies siguin les nostres aliades també per fer un món més just i més pròsper. Sabem que la tecnologia ha de ser una aliada, però ara ens cal saber també situar-la en el benefici col·lectiu, que més enllà de la seva estètica atractiva tingui una ètica al darrere.

Ens interessa i ens entusiasma la tecnologia, tot el que ens aporta i el que ens permet construir, però també ens hauria d'interessar i entusiasmar el cervell de la tecnologia, l'ètica i els valors sobre la qual es construeix, com fer-la oberta, accessible i beneficiosa a tothom, fugint del control i la concentració de poder en poques mans. I per fer això ens cal coneixement. El coneixement és l'element clau que marcarà també la competitivitat de territoris i empreses en els propers anys. Per això agraeixo especialment al síndic la voluntat de fer créixer i difondre el coneixement en un àmbit tan interessant com el de la intel·ligència artificial. I sobretot per l'interès pel que hi ha més enllà de la simple tecnologia, l'interès pel seu impacte en els drets humans i l'ètica en el seu funcionament.

La jornada d'avui se centra en la intel·ligència artificial i drets humans, i fa tot just dues setmanes presentàvem al Palau de la Generalitat l'estratègia en intel·ligència artificial de Catalunya. Aquell dia, al Palau, es va comentar que aquesta és una estratègia de país, és un pas més per al posicionament de Catalunya com a referència entre les societats del coneixement, que és on hem d'aspirar si com a país volem una societat més rica, innovadora i oberta al món.

La intel·ligència artificial, per tant, es troba al centre de les estratègies d'aquest Govern. I, un cop més, el coneixement ha de ser l'eix central des d'on fer créixer la nostra societat i la nostra economia. Però tota aquesta estratègia perdria el seu sentit si no parlés d'una reflexió prèvia, alinear-la amb la Carta catalana dels drets digitals, que són els drets del segle XXI. Entre d'altres, per exemple, l'accés universal a Internet, un model de governança obert, inclusiu, plural i amb diversitat d'actors, la llibertat d'expressió i d'informació i l'ètica en l'ús de la intel·ligència artificial i els algorismes.

Hem d'entendre i saber utilitzar les noves tecnologies com la intel·ligència artificial. Tenim eines cada cop més sofisticades. Però el llegir no ens ha de fer perdre l'escriure. L'objectiu ha de ser sempre dotar aquestes eines de l'ètica necessària perquè ens ajudin a dissenyar solucions per al bé comú i resoldre reptes socials de país, com la mobilitat, el model educatiu, la salut i l'envelliment actiu, el canvi climàtic o el creixement econòmic.

La intel·ligència artificial és una de les principals tecnologies que impacten en el dia a dia de tots els ciutadans i ciutadanes i que revolucionen de ple la competitivitat de les empreses. Per tant, és especialment important que el seu aprenentatge i constant evolució parteixin d'uns paràmetres on tots, tots, ens hi puguem trobar còmodes, i sobretot ens garanteixin els nostres drets. Ho explicava molt bé en Genís Roca, avui aquí present també, en un dels seus magnífics articles. Deia: No hem de debatre de què és capaç una intel·ligència artificial, sinó de com hauríem de pretendre viure els humans al món que estem creant. Qui farà les tasques? Basant-se en quines dades? Sota quines premisses? A qui rendirà comptes? Com es repartirà la riquesa? En qui recauran les responsabilitats? I confirmar si aquestes novetats tècniques realment revertiran en millores per a la subsistència i el confort de la majoria del grup i no només d'uns quants.

Per això, és important fer aquest esforç, per entendre el seu enorme potencial, però també per posar-hi els drets humans al centre. No ens podem permetre la resignació i deixar de qüestionar-nos com es genera aquesta intel·ligència artificial, sota quin paraigua i amb quin model d'aprenentatge. Estic segura que les aportacions que es facin en aquestes jornades seran d'allò més útils per seguir perfeccionant la seva aplicació en benefici de tota la societat. I per aconseguir que la intel·ligència artificial sigui un agent en favor de la societat és especialment rellevant que Catalunya sigui un subjecte proactiu en el seu desenvolupament i no un mer espectador de la seva evolució.

Avui Catalunya és un dels principals pols de recerca europeus, tal com ho demostren molts indicadors, i això és especialment cert en algunes àrees de coneixement, com la mateixa intel·ligència artificial. Aproximadament dues terceres parts de la producció científica en intel·ligència artificial que es fa a l'Estat prové de Catalunya. Són moltíssims els centres tecnològics i de recerca a Catalunya que treballen també en aquest àmbit. Actualment també, tal com ens indicava un estudi sobre intel·ligència artificial elaborat per ACCIÓ, tenim identificades 179 empreses a Catalunya en aquest àmbit i empreses que donen

feina gairebé a 9.000 persones. Parlem, per tant, d'un sector que creix i ho seguirà fent en els propers anys.

Hem de veure la tecnologia com una aliada i no com una amenaça, perquè només així podrem aconseguir apropar els seus beneficis al benestar col·lectiu.

Per acabar, voldria també posar en valor la importància de la col·laboració publicoprivada. També espais com aquests de reflexió i consens entre tots els espais de la societat contribueixen a enriquir les estratègies que ens enforteixen com a país, i especialment en debats tan transcendents com la defensa i la integració dels drets humans en un món copat de noves tecnologies, com el de la intel·ligència artificial. Cadascú, coneixedor dels seus deures i responsabilitats, ha d'aportar la seva part, saber sumar per seguir fent de Catalunya un país d'oportunitats.

Ara reflexionava i pensava com a consellera d'Empresa i Coneixement que hem de desenvolupar des del Departament diferents estratègies que ens facin promoure aquesta economia productiva, innovadora, sostenible i oberta al món? Sí. Entre tots hem de generar coneixement? Sí. Les noves tecnologies han d'evolucionar? Sí. Però sobretot no perdem mai la consciència. És aquesta consciència que hem de tenir tots. Espais com el d'avui ens ajuden a tenir-ne més. De qui som? De qui volem ser? Dels nostres drets i llibertats individuals i col·lectius, perquè sense aquesta consciència no serem mai una societat plena. Moltes gràcies!

Laia Bonet

Tercera tinenta d'alcaldia de l'Ajuntament de Barcelona

Honorable consellera Àngels Chacón; Mr. Peter Tyndall, president of the International Ombudsman Institute and Ombudsman of the Republic of Ireland; Sr. Pablo Martín, subsecretario del Ministerio de Ciencia e innovación, i Sr. Rafael Ribó, síndic de greuges.

Senyores i senyors, bona tarda i moltes gràcies per participar en aquestes jornades sobre intel·ligència artificial i drets humans.

En primer lloc, vull agrair al Síndic i a l'Institut Internacional de l'Ombudsman la invitació i, sobretot, l'organització d'aquestes jornades.

Gràcies també als ponents que avui ens acompanyeu i que contribuïeu, sens dubte, a aquests dos dies de debats importantíssims i urgents.

Fa poques setmanes en Daniel Innerarity, que avui participa també en aquesta sessió, es preguntava en un article: La intel·ligència artificial serà la nostra última invenció? És una pregunta que sintetitza molt bé el sentit d'aquestes jornades.

El debat sobre la intel·ligència artificial, i els efectes que té i tindrà sobre les nostres vides, és el debat més rellevant dels nostres temps. I ho és perquè ens remet a la pregunta essencial: Què és ser humà avui? Com apunta José María Lassalle, que també participa a la segona taula d'aquesta tarda, l'horitzó mateix de la nostra identitat està exposat al desafiament d'una nova alteritat. I aquesta nova alteritat ens obliga a repensar-nos, alhora que ens fa tornar als bàsics i, si em permeteu, també als clàssics.

Necessitem no oblidar que al llarg de la història només hem progressat quan hem posat la humanitat al centre de la reflexió i també al centre de l'acció. L'humanisme es va caracteritzar des dels seus inicis per posar la humanitat al centre de la reflexió filosòfica i política. Calia pensar els problemes de la humanitat sencera. D'aquesta preocupació en derivaren les diverses ideologies polítiques que ens han acompanyat els darrers tres segles. Totes es basaven en la idea de progrés, com una eina per resoldre els problemes de la humanitat.

Fa uns anys, però, que el progrés ja no és el que era, el futur no és el que pensàvem que seria i les distopies dels grans clàssics de la literatura, i també del cinema, s'assemblen massa a la realitat.

Per això ens cal un nou humanisme. Un nou humanisme tecnològic, si em permeten que el denomini d'aquesta manera, com ens agrada fer des de l'Ajuntament de Barcelona. Un humanisme tecnològic que torni a posar la humanitat al centre de la conversa en un món que ja és estructuralment tecnològic i està alterant els nostres marcs interpretatius.

La intel·ligència artificial ens està posant el mirall al davant amb les nostres pors, i també evidentment amb les nostres contradiccions, i per superar-les necessitem empoderar-nos, necessitem atribuir-nos la responsabilitat de controlar l'automatització del món.

Ens trobem en un moment de disrupció, de canvi profund, i hem d'evitar dos perills: el perill de la inacció i el perill de la por. Ambdós van molt lligats perquè la por sovint ens paralitza.

Ens trobem en un moment en què cal recordar aquelles paraules del president Roosevelt en la seva primera inauguració presidencial el 1933 quan deia: "Let me assert my firm believe that the only thing we have to fear is fear itself". L'única cosa a què hem de tenir por és justament a la por. Al terror injustificat i irracional que sovint ens paralitza. Una màxima que és vàlida per als Estats Units en la crisi dels anys 30, però que també ho és per al món d'avui amb la por als robots i a la intel·ligència artificial, o als contagis globals –com és el cas del coronavirus– que pot paraitzar-nos del tot.

Ens hem de conjurar davant d'aquest risc, hem d'evitar que la por –sovint irracional i injustificada– ens paralitzi, hem de convertir aquesta disrupció en una oportunitat.

Tenim el repte i l'oportunitat de posar la ciutadania i els drets digitals al centre del debat tecnològic. Ens trobem en una cruïlla i, en funció del camí que triem, acabarem construint un futur o un altre. Si triem el camí dels drets, hem de defensar un nou humanisme tecnològic, per pensar i per desenvolupar la tecnologia al servei de la ciutadania i dels interessos generals.

Parlar d'humanisme tecnològic, tal com nosaltres l'entendem des de l'Ajuntament de Barcelona, vol dir incrementar el nivell de responsabilitat, tant d'empreses com d'administracions, tant dels desenvolupadors com dels usuaris; augmentar la nostra capacitat de decisió per seguir mantenint el control sobre la intel·ligència artificial, i, finalment, regular els models algorítmics pensant en la seva utilitat per a la societat, introduint-hi criteris ètics i d'interès públic per evitar noves discriminacions.

Els algorismes treballen basant-se en dades, i el biaix en la recollida pot condicionar totes les decisions que prenem. Per tant, necessitem implementar una bona política de recollida de dades.

A Barcelona volem convertir l'humanisme tecnològic en un compromís de ciutat, un compromís amb la regulació de la intel·ligència artificial i l'ús massiu de les dades, i per això hem impulsat una aliança de ciutats pels drets digitals.

Aquestes jornades que s'inauguren avui ens permeten aprofundir en la reflexió sobre com garantir els drets digitals de la ciutadania, que han de ser considerats com una nova generació de drets humans. El fonament d'aquesta nova generació de drets humans digitals ha de ser la plena protecció de la dignitat humana. Hem d'evitar una era digital sense drets. De fet, per a alguns experts, la dictadura de l'algorisme és tan real que ja es podria parlar d'una situació d'emergència digital, tal com ho fem o ens hem acostumat a fer-ho en relació amb l'emergència climàtica.

Necessitem, doncs, respondre de forma proactiva a aquesta situació i l'única eina realment eficaç que tenim per regular la tecnologia, i vull dir-ho amb tota claredat,

és la política. El sistema polític que defensa els interessos generals a través de les lleis. És a això al que entenc que es referia el síndic de greuges en la seva intervenció quan parlava de la necessitat de regular, d'establir els marcs normatius per fer-ho possible. Només les administracions públiques, només els governs si voleu, a escala local, nacional i global, podem embridar el cavall desbocat de la tecnologia.

El control de les dades es juga a les grans ciutats, i per això també estem molt entestats des de la ciutat de Barcelona a intervenir, a participar i a promoure aquesta reflexió. I és per aquesta raó que és tan important com regulem el desplegament del 5G a les grans concentracions urbanes. A les ciutats hi ha milions de càmeres que vigilen el que passa als nostres carrers i ja tenim la tecnologia per incorporar-hi el reconeixement facial, amb tot el que això implica. Les ciutats europees hem de triar quin model d'espai públic volem.

De les vuit ciutats més vigilades del món, set són xineses. La sisena és Londres. Volem seguir aquest camí? O en volem transitar un altre? Si volem triar un altre camí, hem de treballar perquè la tecnologia sigui una eina al servei de les persones i dels seus drets, i no una eina al servei dels governs i de les grans corporacions. Les ciutats hem de ser la primera línia de defensa dels drets digitals.

La intel·ligència artificial ens ha de servir per donar resposta als problemes de la ciutadania i facilitar la democratització de la ciutat, no al contrari. Si sabem utilitzar-les correctament, la intel·ligència artificial i l'ús de les dades públiques per part de les administracions poden ser una oportunitat per oferir nous serveis a la ciutadania i, sobretot, per oferir-los d'una manera molt més proactiva i directa.

Ens cal definir un nou model de ciutat intel·ligent: una ciutat que es pensi a ella mateixa i doni sentit a la tecnologia que desenvolupa i que aplica; una ciutat humana que posi la humanitat al centre del desenvolupament tecnològic, i una ciutat democràtica, inclusiva i orientada al bé comú.

Per aconseguir que la tecnologia sigui una eina per fer efectiu el dret a la ciutat i garantir el que ens agrada denominar com el *dret al futur*, des de l'Ajuntament de Barcelona estem vinculant la transició digital amb els objectius de desenvolupament sostenible de l'Agenda 2030.

Estem convençuts que no assolirem els objectius de desenvolupament sostenible en els àmbits econòmic, ambiental i social si no som capaços d'utilitzar la tecnologia i la transformació digital per tenir una energia més neta, una mobilitat més eficient, unes polítiques socials més ben orientades als diferents col·lectius als quals serveixen, i uns serveis públics més eficients i més intel·ligents.

No ens calen només més serveis públics, sinó serveis públics millors: més eficaços, més eficients i més intel·ligents. I a Barcelona ho estem intentant. Hem posat en marxa diversos reptes perquè les empreses proposin solucions tecnològiques i innovadores per donar respostes als problemes concrets que existeixen a la ciutat i que cada vegada es troben, malauradament, més presents.

Pensem que el desplegament de la tecnologia 5G obre una oportunitat per incidir en el seu model de desplegament urbà, fent-lo verd i neutral i posant-lo al servei de la millora dels serveis públics municipals. Per això hem instaurat un premi 5G a la millor iniciativa tecnològica per donar resposta a un repte social.

A la primera edició, que vam atorgar ara fa poques setmanes, els premis van anar dirigits a dos projectes innovadors per donar assistència i suport a persones grans que viuen soles. El primer és un petit robot personal capaç de moure's i interactuar amb l'usuari per donar-li assistència i suport, i el segon és un dispositiu de videoconferència optimitzada per a un ús molt simple, per facilitar la realització d'activitats a distància per part dels col·lectius usuaris. Són projectes pilot que poden servir per desenvolupar nous serveis, més personalitzats i adaptats, gràcies a l'ús d'aquesta intel·ligència

artificial que tenim, que ens rodeja i que volem que serveixi a tots aquests objectius. Aquests projectes serveixen també perquè l'ecosistema tecnològic que existeix a la nostra ciutat col·labori amb l'Ajuntament a partir d'uns objectius compartits.

Barcelona ja és una ciutat de referència en innovació i tecnologia, però volem també que sigui ciutat de referència en l'humanisme tecnològic. Volem promoure la reflexió i, alhora, aportar solucions i mesures concretes per humanitzar la tecnologia. Estem treballant per convertir Barcelona en capital de la tecnoètica implicant-hi empreses, entitats socials i també universitats.

El proper mes de desembre, per exemple, la ciutat acollirà la Digital Future Society el fòrum internacional que abordarà els reptes de la transformació tecnològica i buscarà un compromís públic i privat per avançar en la bona governança del futur digital. Una iniciativa impulsada pel Ministeri d'Economia i també pel Mobile World Capital amb l'objectiu de millorar l'impacte de la tecnologia en la societat per construir un futur més just i més inclusiu en l'era digital en què vivim. Volem que aquest fòrum, que ja ha organitzat diversos debats a la nostra ciutat, tingui continuïtat i ens permeti també avançar en iniciatives concretes, de reflexió i acció, i d'aterratge sobre el terreny.

Per fer-ho possible comptem, n'estic convençuda, amb tots els que avui compartiu aquesta preocupació i per això assistiu en aquestes jornades. Comptem amb tots vosaltres, us animem a participar-hi també, i amb les vostres propostes i reflexions estic convençuda que podrem fer un debat molt ric. Estic convençuda també que tots compartim la idea que el futur no ens espera i ens calen respostes urgents. Moltes gràcies!

Peter Tyndall

President de l'IOI i ombudsman de la República d'Irlanda

Bona tarda. És un plaer estar de tornada a Barcelona. Vull començar agraint al síndic i president de IOI Europa per organitzar aquesta conferència avui. La seva oficina ha estat sovint a l'avantguarda del treball dels ombudsmen i ha identificat problemes que tindran un gran impacte en la manera com la institució de l'ombudsman es desenvolupi en el futur. Ha organitzat regularment conferències al voltant de temes emergents com aquest en moments importants.

Dir que vivim en un món canviant és una obvietat. Res no roman igual gaire temps. El treball de l'ombudsman no és immune a això. El món en què treballem actualment és molt diferent del que es va trobar el primer ombudsman a Suècia fa més de dos-cents anys. Tanmateix, els principis que sustenten el nostre treball segueixen sent els mateixos. Som independents, tenim lliure accés a la informació, som objectius. El nostre treball consisteix a equilibrar la diferència de poder entre l'Administració i l'individu. Som part integral dels estats democràtics. Oferim reparació a les persones que han patit i, al mateix temps, utilitzem les lliçons apreses en el nostre treball per millorar els serveis públics i evitar que es repeteixin els errors. La institució de l'ombudsman ha demostrat ser notablement adaptable a un món canviant. Sense perdre de vista les seves característiques fonamentals, ha assumit molts rols i responsabilitats i ha après a operar en un entorn molt canviant. La reeixida evolució del concepte d'*ombudsman* es reflecteix en la seva difusió global. L'IOI, per exemple, té gairebé dos-cents membres en més de cent països de tot el món.

M'agradaria considerar alguns dels reptes actuals a què s'enfronta la institució, en particular el que planteja la IA. En el passat, les oficines de l'ombudsman generalment rebien queixes per escrit. A les persones se'ls lliurava un formulari perquè l'emplenessin o acudissin a l'oficina en persona. Sovint, es trigava molt a investigar la queixa i a determinar-ne el resultat.

Avui dia, les persones fan la major part de les activitats en línia: compren, fan les operacions bancàries, reserven els viatges i l'allotjament, es comuniquen i, el més important, tracten amb els serveis públics. Les seves expectatives no són només obtenir resultats ràpids, sinó tenir uns nivells molt alts d'interactivitat. També esperen comentar les seves interaccions i

qualificar-les. La meua generació es relaciona amb el tema digital com un segon idioma. Tanmateix, moltes persones són natives digitals, han crescut amb telèfons intel·ligents i Internet, i no poden concebre el temps que tenen per davant. Esperen interactuar amb els serveis públics com ho fan amb Amazon o Facebook, comunicar-se com ho fan a WhatsApp. Per a molts, l'última vegada que van escriure una mica va ser en un examen. El bolígraf, per a aquestes persones, s'ha quedat obsolet.

Les oficines de l'ombudsman també estan acostumades a tractar amb organismes que tenien la informació en arxius en paper. Ara, tanmateix, cada vegada hi ha més informació en el núvol. No estem acostumats a oferir resultats instantanis. No esperem alts nivells d'interactivitat amb les persones interessades i no esperem que la nostra actuació es revisi en línia i sigui objecte de debat a les xarxes socials. Els nostres col·legues al Regne Unit han tingut alguna experiència amb grups de persones interessades que es van unir per criticar la seva actuació. Ara és molt més fàcil fer això en línia. I el to de les converses a les xarxes socials pot ser un to molt enfadat o fins i tot ofensiu. En línia la gent diu coses que dubta a dir-te a la cara. La majoria de les oficines d'ombudsman responen a aquestes noves circumstàncies: faciliten les queixes en línia i moltes ofereixen actualitzacions de l'estat de les queixes a les persones interessades.

El sistema i-People a Corea és un exemple establert des de fa molt temps amb un enfocament conjunt per a la gestió de queixes. El mateix sistema de reclamació el comparteixen els organismes públics i l'ombudsman. Això permet una transferència molt més ràpida de les queixes, que ara es resolen localment, i proporciona una rica font d'aprenentatge a partir de les dades de les queixes. Les persones interessades volen interactuar cada vegada més amb les nostres oficines, com ho fan en unes altres àrees de les seves vides, ja sigui per Facebook, Instagram o WhatsApp, per esmentar les plataformes més utilitzades en la meua part del món. També esperen respostes molt més àgils i expressen ràpidament el seu descontentament davant qualsevol retard o quan no acceptem les seves queixes. La necessitat de resultats ràpids reforçarà encara més la tendència a intentar resoldre les queixes mitjançant la interacció amb els proveïdors de serveis, en comptes d'investigar minuciosament cada reclamació i generar un informe molts mesos després. Això, al seu torn, presenta reptes sobre la forma en què les oficines de l'ombudsman responen a les tendències en relació amb el tractament de les dades de les queixes. Fa que sigui més difícil garantir que s'identifiquin i s'abordin els problemes sistèmics. No n'hi ha prou de resoldre les queixes. També hem d'assegurar-nos de millorar els serveis públics i evitar repetir errors. També hem de considerar la millor manera de millorar les habilitats del nostre personal per poder bregar amb les evidències que hi ha en els sistemes i no en el paper.

Esmentaré breument que no tothom està posat en l'univers digital, potser per l'edat, encara que hi ha molts *silver surfers* (gent gran internauta), o per mancances. Per tant, encara hem d'assegurar-nos que, mentre participem en aquest univers digital, seguim sent accessibles per a les persones que no fan transaccions comercials en línia. També hem d'investigar casos en què a les persones se'ls negui l'accés als serveis perquè els organismes públics només interactuïn en línia.

Ens hem adaptat amb diferents nivells d'èxit a l'entorn digital en què treballem actualment. La intel·ligència artificial ens plantejarà nous reptes. Recentment, hi ha hagut informes, per exemple, que mostren un gran avenç en la detecció del càncer mitjançant la intel·ligència artificial: les màquines ara poden llegir mamografies tan bé com els humans o les operacions es duen a terme utilitzant la tecnologia robòtica impulsada per la IA. Alguns podeu haver experimentat la intel·ligència artificial a l'hora tractar amb empreses per telèfon, a través d'una aplicació o pàgina web. La intel·ligència artificial està permetent que els cotxes es conduixin sols. Aviat serà endèmic i la gent opina que és probable que causi moltes pèrdues de llocs de treball.

Molts dels serveis públics de les nostres jurisdiccions inclouran la IA. Les decisions que ara prenen les persones les prendran els ordinadors. Com hem sentit, s'està desenvolupant una nova branca d'ètica que cobreix la IA i es tenen en compte dilemes inimaginables abans.

Si el cotxe que condueix una persona es troba amb un pare o una mare amb un nadó en un cotxet que creua una carretera i no es pot aturar, hauria d'atropellar-los i salvar l'ocupant del cotxe o xocar sabent que això posarà en perill la vida de l'ocupant? Són un tipus de qüestions ètiques que eren inimaginables abans de la IA.

Haurem de desenvolupar la nostra anàlisi de les queixes per determinar si s'ha produït una injustícia com a resultat de l'ús de la intel·ligència artificial. No es poden preveure totes les circumstàncies. De vegades, l'ús d'un criteri pot ser essencial per tenir en compte les circumstàncies particulars d'un individu. Pot la intel·ligència artificial fer això? La resposta probablement sigui que sí, si es va tenir en compte en l'etapa de disseny.

Haurem de decidir si un ordinador és susceptible de mala administració i, si fos d'aquesta manera, com es pot corregir? Com podem evitar que tornin a ocórrer les mateixes injustícies? El problema de la caixa negra és un dels que, sens dubte, presentarem en la nostra consideració en aquesta conferència. A mesura que l'aprenentatge automàtic es torni endèmic, s'assoliran resultats en què el procés de presa de decisions sigui opac. Estem acostumats a determinar si hi ha hagut mala administració examinant el procés de presa de decisions i veient si els passos fets estan justificats. Això serà potencialment molt més difícil en el futur. És fàcil preguntar-li a una persona com va arribar a prendre una decisió. Això no sempre serà el cas amb la IA.

També hem de considerar com l'ombudsman pot utilitzar la intel·ligència artificial. Podem utilitzar-la, per exemple, per determinar si les queixes són de la nostra jurisdicció o competència, si s'han presentat dins el termini, si hi ha evidència d'injustícia? Pot interactuar amb els nostres interessats en comptes que ho faci el personal? Pot desenvolupar un paper en la investigació de les queixes? Com a mínim, podria desenvolupar un paper en la identificació de problemes sistèmics tenint en compte les dades de les queixes.

És clar que l'entorn en què treballem continuarà canviant i portarà nous reptes, però també oportunitats. Tanmateix, confio que la institució de l'ombudsman seguirà adaptant-s'hi, però es mantindrà fidel als seus valors i principis fonamentals. Oferir accés a la justícia, millorar els serveis públics, promoure l'estat de dret i reivindicar els drets i les llibertats fonamentals sempre han estat els nostres principis rector, i seguirem guiant-nos-hi en qualsevol circumstància que es doni en el futur. Aquesta conferència és molt oportuna i ens ajudarà a comprendre l'evolució actual i futura de la IA i com les nostres oficines han d'adaptar-s'hi per aprofitar les oportunitats que ofereix i afrontar els indubtables reptes que presenta. Gràcies!

Intel·ligència artificial i drets humans

reptes, rols i eines de les institucions d'ombudsman

Dilluns, 2 de març de 2020

Realitats i límits de la intel·ligència artificial

Moderador: Andreas Pottakis, director de l'IOI i ombudsman de Grècia

Ulises Cortés, catedràtic i investigador titular de la UPC, Barcelona Supercomputing Center (BSC)

Montserrat Pardo Bayona, directora d'afers corporatius de Microsoft Espanya

Alfonso González, director de Government & Regulatory Affairs, IBN Spain

Oliver Smith, director d'estratègia i cap del Departament d'Ètica a Alpha Health

Carina Lopes, responsable de la Digital Future Society Think Tank (Mobile World Capital)

Andreas Pottakis

Director de l'IOI i ombudsman de Grècia

Estic encantat de ser aquí, i també de tornar a Barcelona. Ha passat un temps, des del 2017, també per a un acte organitzat pel síndic. Un esdeveniment que crec que va tenir molta rellevància per al que debatrem avui i demà, perquè part de la discussió del 2017 va tenir a veure amb la manipulació d'Internet i de diferents plataformes per promoure el discurs d'odi i la xenofòbia.

Esmento això perquè crec que, ja que faré un parell de comentaris abans de donar la paraula als nostres distingits membres de la taula, un dels principals problemes i una de les principals amenaces als quals ens hem d'enfrontar les institucions d'ombudsman en particular i les institucions de drets humans en general, són l'amenaça i els problemes que la intel·ligència artificial i l'ús o l'abús de la intel·ligència artificial poden tenir sobre els nostres drets cívics i la nostra consciència cívica.

En primer lloc, agraeixo al síndic la seva invitació per permetre'm presidir aquesta primera taula. Crec que comparteixo amb el síndic el mateix desconeixement sobre la matèria. Tinc la impressió que volia que presidís la taula perquè volia exposar la meva ignorància.

Així doncs, un cop exposada obertament la meva ignorància, el més assenyat serà contribuir al debat amb algunes preocupacions i algunes possibles preguntes, tot i que em temo que la majoria dels punts que volia plantejar ja han estat tractats per Peter Tyndall, el president de l'IOI. Ara, com a institució d'ombudsman, treballem amb drets humans i també lluitem contra la mala administració, però el nucli de la nostra feina és l'ésser humà. Ara crec que haurem de tractar amb el que anomenem *animal piratejable* i ja no amb l'ésser humà individual.

El tema de la primera taula, que els meus distingits amics ens ajudaran a abordar, té a veure amb la situació en què ens trobem actualment, i probablement cap a on ens dirigirem en un futur proper pel que fa a la intel·ligència artificial. Per tant, tinc molt per aprendre personalment. Pel que entenc, i ja em corregireu, per descomptat, encara estem en la fase d'aquesta anomenada *intel·ligència artificial estreta*, que apliquem, per exemple, en l'ús més comú per als usuaris de les plataformes d'Internet, etc. Els exemples típics són la personalització de contingut o els filtres bombolla, que tenen a veure amb la creació d'una mena d'imatge virtual de qui som o, almenys, segons els diferents algorismes, com pretenem ser. També entenc que el que en algun punt hem de tractar és l'anomenat punt de singularitat en què en algun moment arribarem a un nivell d'aplicació diferent pel que fa a la intel·ligència artificial. Espero que rieu perquè no soc lluny d'on som, però segurament em corregireu.

Per a nosaltres, els ombudsman, és de vital importància cercar i rebre l'assistència dels experts per entendre en primer lloc què és i què pot fer la intel·ligència artificial, és assegurar

que hi hagi una supervisió humana i un debat, i reflexionar sobre el tipus de supervisió humana en relació amb la intel·ligència artificial. Em sembla que els companys que m'han precedit ja ho han esmentat o insinuat, i només vull concloure el meu primer comentari, molt breu, reflexionant sobre un document molt recent que, de nou, és rellevant per a nosaltres, els ombudsmen. Em refereixo al Llibre blanc sobre intel·ligència artificial de la Comissió Europea, publicat fa només uns deu dies. I vull posar en relleu que una de les recomanacions o conclusions finals en què se centra la Comissió Europea en el seu Llibre blanc és la qüestió de la governança en relació amb la intel·ligència artificial. M'imagino que es refereix a una manera que hi hagi supervisió humana, una manera de reflexionar sobre el tipus de supervisió humana que tindrem. Serà de nou en forma de codi ètic? Serà potser en termes d'un conjunt d'instruments legals vinculants? A banda dels instruments legals existents que estem aplicant, i em refereixo, per exemple, a la legislació contra la discriminació, perquè crec que es va esmentar que la intel·ligència artificial pot provocar parcialitat i discriminació, en matèria de protecció del consumidor o del Reglament general de protecció de dades, el cèlebre reglament amb el qual la majoria de nosaltres tenim problemes.

Però la meua principal preocupació no són els nostres drets com a consumidors, sinó els nostres drets cívics, els nostres drets polítics pel que fa a intel·ligència artificial. Com a ombudsman, la meua gran por, i així conclouré, és que les actuals aplicacions d'intel·ligència artificial que utilitzem en el nostre dia a dia no només siguin una vulneració de la nostra llibertat d'expressió, sinó que també constitueixin un risc per a la nostra llibertat per formar-nos una opinió.

Doncs ara que ja dec haver avorrit la majoria del públic, potser toca escoltar alguna cosa molt més intel·ligent, i començarem amb Ulises Cortés, professor i investigador titular de la Universitat Politècnica de Catalunya i del Centre de Supercomputació de Barcelona.

Ulises Cortés

Catedràtic d'Intel·ligència Artificial de la Universitat Politècnica de Catalunya,
Barcelona Supercomputing Center (BSC)

Moltes gràcies per la invitació a aquest esdeveniment. Intentaré descriure, en quinze minuts, els fets essencials i els límits tecnològics de la intel·ligència artificial (IA). L'objectiu és explicar el meu punt de vista sobre la IA i els seus límits reals, però, primer, un descàrrec de responsabilitat: només soc un professor d'intel·ligència artificial. No soc un especialista en ètica. Puc fer errors importants a l'hora d'utilitzar alguns termes. Jo no soc un advocat. A més, no soc un futurista i, per tant, no us intento convèncer que tinc una visió clara sobre el futur. La meua idea és aclarir el que sabem fins ara sobre els usos de la IA i el que podem construir amb les eines i els ordinadors d'IA existents que tenim a mà.

L'impacte de la IA en la societat

Avui sou aquí perquè la IA té un impacte enorme en la societat. No hi ha dubte. Les grans empreses, com ara Google, Amazon, Facebook, Apple, Microsoft i Alibaba, contribueixen a la confusió, ja que fan molt soroll sobre l'ús de la IA per transformar la nostra forma de vida i guanyar molts diners amb això. Creen excel·lents campanyes de màrqueting en què ens diuen quin és l'abast de les capacitats de les seves tecnologies i com en són, d'útils. A més, el que sabem és que posen al mercat, a una velocitat vertiginosa, tecnologies que cal adaptar i corregir cada dia. Per tant, tenen a disposició noves versions d'aquesta tecnologia en tot moment, i ningú més que ells és capaç de comprendre el funcionament real d'aquests programes. En aquest moment, hi ha diferents punts de vista sobre el possible impacte de la IA des de les visions més extremistes i poc positives de la IA, com la d'Elon Musk i Stephen Hawking, que alerten sobre el mal ús de les tecnologies basades en IA. A l'altre extrem, hi ha persones que intenten regular-ne l'ús per evitar-ne un mal ús.

Tanmateix, potser no és només la intel·ligència artificial, sinó també la seva combinació amb la potència informàtica, que és cada vegada més gran, una gran quantitat de nous algorismes sense censura i la falta de regulació del mercat el que està creant la situació actual. És evident que la IA ens impacta, però hem de diferenciar el soroll del màrqueting de l'efecte real de la IA, i això no és una tasca fàcil. Hem de tenir clar el fet que la IA és només una part del problema, el més visible, però no és el panorama complet.

Per utilitzar la terminologia comuna, permeteu-me començar amb una definició pràctica d'intel·ligència artificial. N'hi ha moltes, agafo aquesta que va ser encunyada pel Grup Europeu d'Experts d'Alt Nivell en Intel·ligència Artificial:¹

Els sistemes d'intel·ligència artificial (IA) són sistemes de programari (i possiblement també de maquinari) dissenyats per humans que, atès el seu objectiu complex, actuen en la dimensió física o digital percebent el seu entorn a través de l'adquisició de dades, interpretant les dades estructurades o no estructurades recopilades, raonant sobre el coneixement, o el processament de la informació, derivat d'aquestes dades i decidint quin és la millor acció (és) a prendre per aconseguir l'objectiu fixat. Els sistemes d'intel·ligència artificial poden utilitzar regles simbòliques o aprendre un model numèric, i també poden adaptar el seu comportament analitzant com l'entorn es veu afectat per les seves accions anteriors.

La majoria de les meves referències seran als documents de la Unió Europea perquè Europa té la seva pròpia posició sobre els usos permesos de la IA. Europa vol una IA

¹ <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

fiable i centrada en l'ésser humà, i jo m'hi adhireixo, fins i tot tenint alguns dubtes. Els sistemes basats en IA són un tipus de programa que pot actuar tant a Internet, que és un espai virtual, com en el món físic. A més, el que és més interessant és que aquestes màquines no només acumulen informació (la gent sovint només parla sobre l'ús de dades personals, les dades en són només una part), sinó que també acumulen coneixement, i aquesta és la part emocionant de la IA.

Aquestes màquines intel·ligents utilitzen dades i coneixements; ambdós es fan servir per prendre decisions. Aquestes màquines prenen decisions per nosaltres cada dia, i ens representen en moltes situacions. Aquest ús està canviant la manera amb què vivim i interactuem amb el món. La majoria de les persones ara cedeixen aquesta part de les seves responsabilitats a una màquina perquè prengui petites decisions, per exemple, les que pren un optimitzador de ruta quan intentem anar a algun lloc. Aquestes parcel·les d'autonomia de la màquina són cada vegada més grans perquè s'agreguen per ampliar-ne el nombre i la freqüència. Com a resposta, la comunitat europea d'IA intenta convèncer les persones que necessiten educar-se en la tecnologia per poder comprendre què hi ha darrere d'aquests algorismes i del seu impacte.

En termes generals, hi ha dos tipus d'IA: una basada en l'ús de símbols per raonar sobre un problema i resoldre'l, que a més pot donar una explicació transparent sobre com utilitza el coneixement i les dades. L'altre tipus no utilitza símbols, funciona com una caixa negra per resoldre problemes. Aquests sistemes no poden explicar com van resoldre un problema. Aquests sistemes són coneguts com *deep learning* (aprenentatge profund o aprenentatge automàtic). El primer que heu de saber és que l'aprenentatge profund és només una petita part de l'aprenentatge automàtic i que l'aprenentatge automàtic és part de la IA. L'aprenentatge profund és molt popular avui dia a causa del màrqueting que rep i a alguns resultats espectaculars obtinguts. A més, hem d'evitar deixar-nos enganyar per la idea que la IA, en general, o l'aprenentatge profund, en particular, poden fer-ho tot per damunt del nivell dels humans.

Com de lluny o de prop estem d'aquesta terra promesa de la IA?

Com de lluny o de prop ens trobem d'aquesta terra promesa de la IA? Per esbrinar-ho, examinem alguns exemples d'aplicacions d'IA (vegeu la taula 1). Avui dia, una de les expectatives més comunes sobre els sistemes basats en IA és que aviat tindrem vehicles autònoms. Personalment, no puc entendre com algú pot voler tenir un automòbil autònom perquè vol que sigui lliure i que vagi a un altre lloc. No necessito una màquina perquè em porti a algun lloc. No necessito una màquina amb què comunicar-me cada dia sobre on soc o quina és la meua velocitat actual. No estem a prop de tenir cotxes amb una autonomia total, però hi ha una investigació molt potent per produir aquests cotxes. Mentrestant, com a resultat, avui disposem d'alguns mòduls de conducció útils, com ara per a persones que puguin tenir alguna discapacitat. Alguns cotxes poden tenir un suport per a estacionament o estacionar-se ells mateixos, etc.

Ara, vegem breument els usos militars de la IA. Els exèrcits de diversos països ja tenen armes semiautònomes. Tanmateix, estem molt lluny de veure armes autònomes reals. No obstant això, potser ho veiem en un futur no gaire llunyà, perquè, per descomptat, la major part de la tecnologia d'IA ha sorgit de la investigació militar i els exèrcits continuen invertint en aquest tipus d'investigació. El sentit comú convida a prohibir les armes capaces de prendre decisions autònomes per aconseguir els seus objectius o fins i tot per canviar d'objectiu. Si no podem prohibir-les, almenys hem de regular-les.

Sesame Street

NEAR	FAR
Assisted Driving modules	Autonomous Car
Semi-autonomous weapons	Autonomous weapons
Decision Support Systems	Fully Autonomous Decision Systems
Transportation route optimization algorithms for project planning optimization	Autonomous Building Systems
Clinical image recognition	Autonomous medical diagnostic systems based on clinical images
...	...



Taula 1: Terra promesa i IA

Permeteu-me presentar-vos alguns termes per classificar les armes. Una classificació aproximada conté aquests tipus: armes controlades, supervisades, automatitzades i autònomes. Aquests termes defineixen els diversos graus de supervisió humana sobre una arma. Per exemple, un dron que es controla amb control de ràdio en totes les accions que fa pertany a la primera classe. Aquests sistemes tenen supervisors, és a dir, es tracta d'una màquina que té un objectiu i l'executarà sota la supervisió d'una persona que sempre estarà a l'aguait. Les màquines automatitzades, com les mines terrestres –que no sé per què les continuem produint–, estan preparades per reaccionar cegament davant qualsevol element d'entrada. Després hi ha les futures armes semiautònomes, com els robots produïts per Boston Dynamics als Estats Units. Hi ha un debat continu aquí, o els agrada automatitzar les coses o volen ampliar les capacitats humanes utilitzant la IA. L'acumulació i la propagació incontrolada d'armes de totes les categories tenen un impacte humà devastador. Si anem cap a l'automatització del procés, perdem el control del dispositiu.

Quan diem que algú s'expandeix? Per exemple, una persona s'expandeix si utilitza lents per ampliar la seva capacitat visual. Ara podem ampliar les capacitats d'algú sense que aquesta persona tingui un problema o una discapacitat. Per exemple, utilitzant ulleres amb infrarojos per veure en la foscor i mesurar la distància envers l'objectiu. De la mateixa manera, podem parlar d'eines basades en IA que incrementen les capacitats humanes. D'entre aquestes eines podem distingir els usos correctes de la IA dels incorrectes.

Un exemple de doble ús són les eines d'autenticació biomètrica. És una aplicació beneficiosa i útil. Per exemple, quan sigui necessari restringir l'accés a zones crítiques, aquest sistema només permetrà l'entrada a persones autoritzades, no permetrà que hi entri ningú més. Una zona crítica pot ser el vostre telèfon que utilitza reconeixement facial per permetre obrir-lo. Aquest tipus d'eines es tornen alarmants si s'utilitzen

per a la identificació biomètrica permanent i remota. En aquest cas, els drets individuals es veuen compromesos. Especialment quan s'utilitzen sistemes per detectar continuament les persones que utilitzen la seva informació biomètrica. Els problemes de seguretat pública són complexos i no sempre són blancs o negres.

Un altre possible ús indegut és el capitalisme de vigilància o les tasques de vigilància, en què la IA s'utilitza per recopilar dades sobre nosaltres, les dades es recopilen en un altre lloc i després s'integren per identificar, classificar o perfilar les persones. Això ja ocorre cada dia, moltes vegades al dia en qualsevol dispositiu que estigui connectat a Internet. Cadascú de nosaltres, quan no tanquem el servei de localització dels telèfons mòbils, donem informació no només sobre nosaltres mateixos, sinó també sobre totes les persones que ens envolten. Hem d'estar socialment convençuts que no és bo donar aquesta informació sense la nostra conformitat explícita.

Entre altres usos incorrectes de la IA, hi ha les notícies falses i els ciberatacs potenciat per la IA. Són part del costat fosc de la IA. Una de les coses que vull recordar és que la IA no és perjudicial per si mateixa. No és la primera vegada que la investigació o els resultats científics s'utilitzen de manera poc ètica. Tenim exemples el 1776 quan Jensen va fer el primer experiment de vacunació. Permeteu-me enumerar alguns exemples com ara el Tuskegee Syphilis Study-USPHS, els estudis del Departament d'Energia dels EUA sobre la pluja radioactiva, etc.

Ètica i intel·ligència artificial

Per donar resposta als usos poc ètics de la IA, el 2017, la comunitat d'IA va proposar els Principis d'Asilomar. Així mateix, el 2017, la comunitat europea d'IA va proposar la Declaració de Barcelona sobre el bon ús de la IA, liderada pel professor Luc Steels i el professor López de Mántaras. Les guies ètiques proposades pel Grup d'Experts d'Alt Nivell són el posicionament polític europeu.

Avui dia, tots creen els principis que s'ajusten als seus interessos. Tenim molts principis, però no es respecten gens. Fins i tot les grans empreses contribueixen a aquesta confusió. Hi ha un article a la revista Nature, publicat a principis de 2020, en què els autors informen de més de noranta-quatre declaracions diferents d'associacions, països i empreses. Noranta-quatre! Van des dels set als vint-i-set principis d'IA, i només quatre són comuns per a tots. A més, per si no n'hi havia prou, també l'Església està implicada en aquest moviment. Microsoft i IBM cerquen tenir una aliança amb l'Església per a una IA. Espero que ara tingueu una idea més clara de per què la IA és al centre de moltes discussions socials, polítiques i econòmiques.

Permeteu-me parlar breument sobre el que fem a Europa. Com he dit, tenim un grup d'experts que ha elaborat pautes ètiques per a la IA. Aquestes directrius estableixen tres requisits perquè la IA sigui admissible a Europa. Primer, ha de ser una IA legal. Això significa que compleixi totes les lleis. En la seva majoria totes les que emanen del Consell d'Europa i de l'RGPD. En segon lloc, aquesta IA ha d'estar alineada èticament i, en tercer lloc, ha de ser sòlida. Aquesta última característica és important perquè les empreses s'afanyen a posar les seves tecnologies al mercat abans que es provin a fons. Hem de repensar quin és el significat exacte de sòlid en aquest context.

En les directrius HLEG-AI, hi ha quatre imperatius ètics: el principi d'autonomia, com he dit abans, que requereix tenir un ésser humà a bord en tot moment. El principi de no-infracció: l'acció de la IA, ja sigui programa o maquinari, no ha de danyar els humans. Principi de justícia, ha de ser just. Hem de definir el concepte d'equitat, però, bàsicament, una de les coses que hem de tenir són màquines lliures de biaixos, cosa que significa que els creadors de programa han de

provar les seves dades i els seus programes abans que surtin al mercat. A més, el principi d'explicabilitat: necessitem màquines capaces d'explicar per què prenen determinades decisions. Si vaig a un metge i em diagnostica càncer, vull saber quina és l'evidència d'això i les raons que hi ha darrere. Si no estic convençut, cercaré una segona opinió. Si no m'agrada el diagnòstic que he obtingut fet per Google, puc anar a Yahoo per obtenir un nou diagnòstic. Reemplaçar els humans en aquestes tasques no és una bona idea. A Europa necessitem un enfocament ètic centrat en l'ésser humà. Necessitem respecte per la dignitat humana, llibertat per als individus, respecte per la democràcia i la justícia, i veurem exemples de com podem influir en això, igualtat, no-discriminació i solidaritat, i els drets civils han de situar-se en primer lloc.

Diluvi de dades o economia de dades

Conclusions

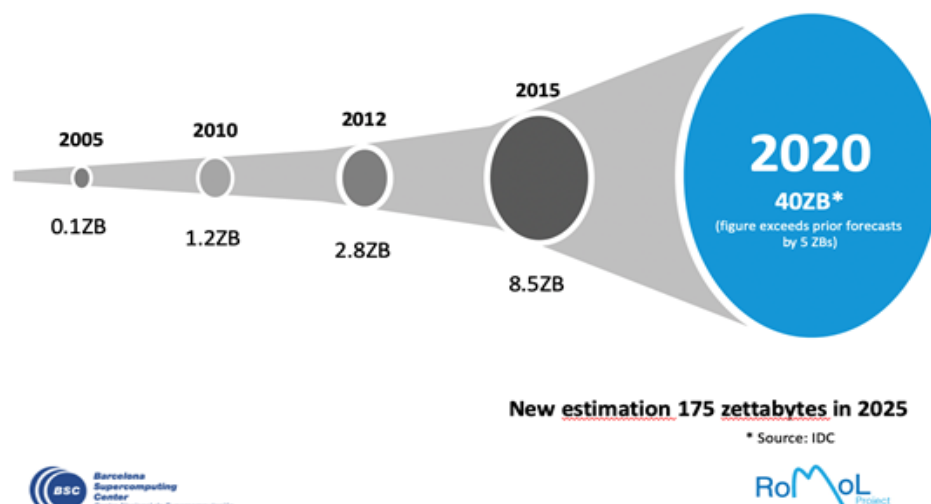
Ara descriuré l'opinió de la UE sobre les dades i la IA. Hi ha un llibre blanc sobre IA que se centra en la nova economia de dades. L'objectiu és clarament regular la nova economia basada en dades:

El creixement econòmic sostenible actual i futur i el benestar social d'Europa es basen cada vegada més en el valor creat per les dades. La IA és una de les aplicacions més importants de l'economia de dades. Avui dia, la majoria de les dades estan relacionades amb els consumidors i s'emmagatzemen i es processen en una infraestructura central basada en el núvol. Per contra, gran part de les dades en un futur, dades que seran molt més abundants, provindran de la indústria, de les empreses i del sector públic, i s'emmagatzemaran en una varietat de sistemes, especialment en dispositius informàtics que funcionin en la perifèria de la xarxa.

Tots sabem que les dades són un d'aquests motors que mouen tota la nostra economia, i Europa reconeix que estem molt endarrerits en la competència amb els EUA, la Xina i les gegantines plataformes comercials. Tanmateix, tenim una destacada indústria de la microelectrònica a Europa. És possible utilitzar aquesta tecnologia per portar la computació al límit i utilitzar les dades que es produeixen. La pregunta és com utilitzar-les de manera justa?

Sabent que les dades són la primera matèria més valuosa, parlem de l'allau de dades en què vivim immersos, parlem de big data. No diria que m'agradi el terme big data. Prefereixo parlar de dades correctes. Per fer-vos una idea de la quantitat d'informació de què parlem, consulteu la taula 2, us donarà una visió panoràmica del que parlem. Per exemple, i si la meva memòria no em falla, la intel·ligència artificial va néixer el 1956. Aquest mateix any, IBM va produir el primer disc d'emmagatzematge de dades amb cinc megabytes. Llavors cada megabyte costava 10.000 dòlars. Actualment, tenim a les nostres càmeres molts GB. Cada dia, Internet produeix de 10 a 18 zettabytes i creix amb l'ús progressiu de la Internet de les coses (IoT) (interconnexió digital d'objectes quotidians amb Internet). És una quantitat sorprenent de dades. La Unió Europea afirma que aquest any (2020) produïrem 40 zettabytes. A més, al mateix Llibre blanc, hi ha previsions per als propers cinc anys. En cinc anys hi haurà 175 zettabytes l'any. Com a societat, produïm tones de dades, consulteu la taula 3.

The Data Deluge



Taula 2: El diluvi de dades.

Qui controla les meves dades?

Un es pot preguntar, on són totes aquestes dades? Què farem amb aquestes dades? Qui assumeix la responsabilitat de les dades? Individualment, la pregunta és qui controla les meves dades?

A més, en aquest punt ja ha quedat clar que són les grans empreses les que controlen les nostres dades. Això és cert ara i seguirà sent vàlid per un temps indeterminat. A més, hi ha algunes empreses que diuen que no volen tenir totes les dades. No les volen, perquè ja saben quina és la dada rellevant per conservar.

Podem preguntar-nos, confieu en els gegants d'Internet per protegir les seves dades personals? Quin és el benefici de l'ús de les nostres dades? Ricardo Baeza-Yates va escriure que cada any les grans empreses produeixen 180.000 dòlars per persona. El Llibre blanc parla de la diferenciació entre serveis i productes, i tota la legislació europea s'ocupa dels productes, però no dels serveis. En català sona millor: "Els serveis de Google més enllà de la mort."²

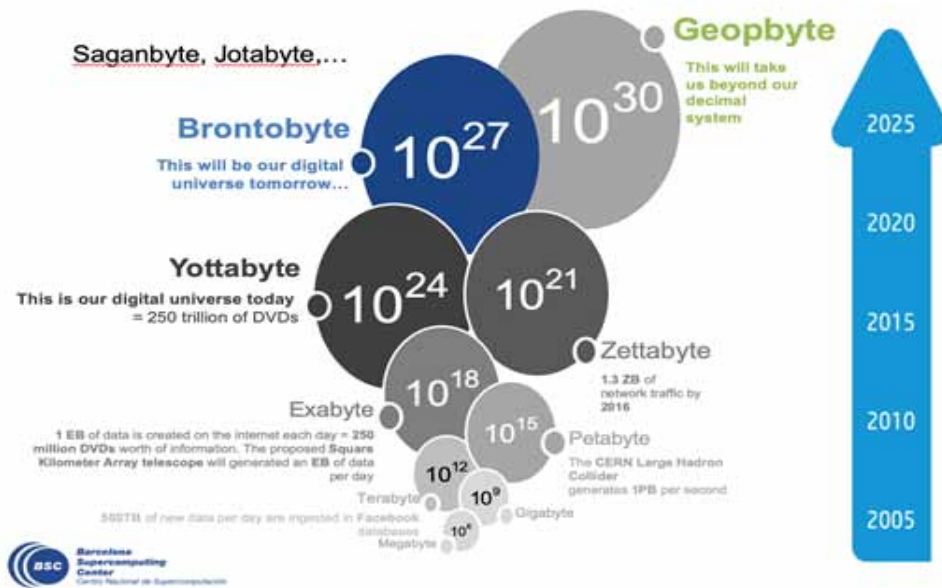
A més, si llegiu, perquè cap de nosaltres llegim tots els contractes d'aplicacions, això és el que diu el contracte de Google sobre les seves dades: Aquesta llicència continua fins i tot si deixa d'utilitzar els nostres serveis. Per exemple, quan és mort. Tenen totes les dades. Ens classifiquen. Intenten predir la manera amb què ens comportem. Utilitzen filtres de bombolles per posar el nostre comportament en una pista específica, i si no poden dissenyar el nostre perfil, llavors trien induir-nos a canviar hàbits.

Confieu que el vostre govern protegirà les vostres dades personals dels gegants d'Internet? La meua resposta és no; ells no estan fent això. Confieu que el vostre govern protegirà els vostres drets, els vostres drets humans i els vostres drets civils dels gegants d'Internet? La meua resposta continua sent no, i si no, podeu veure el que va succeir amb els casos de Cambridge Analytica o Compass en el passat. Llavors, és clar que hi ha una falta de regulació. Tenim tots els mitjans, tenim tota la investigació, som forts a Europa per produir molts resultats científics, però no tenim els diners per pagar els investigadors. Les grans empreses paguen molts diners per aconseguir els seus objectius; porten totes les ments brillants per fer el que han

²"Google services beyond death"

de fer. Llavors, el major repte és evitar convertir-se en idiotes inútils i rebutjables simplement connectats a un telèfon. El director del Centre de Supercomputació de Barcelona, el professor Mateo Valero, sempre diu que en els teus mòbils tens deu mil vegades la potència que va portar l'home a la lluna. I, el que és més, estic segur que gairebé cap de vosaltres ha programat mai una línia de codi en els vostres telèfons mòbils. Llavors, per què teniu tota aquesta potència? La necessiteu? O qui necessita aquesta potència informàtica en tots els mòbils? Per fer què?

Data volumes driving AI



Taula 3: Terra promesa i IA.

Una conclusió: És clar que la IA està influint en les nostres societats. Som a Europa, intentem, com a europeus, influir en la manera en què la IA està marcant el ritme de la societat. Podem aconseguir aquest objectiu dissenyant el programa d'IA d'una manera que estigui alineada amb un conjunt comú de valors ètics, culturals, legals i socioeconòmics europeus que haurien d'estar en la nostra altura. És obvi, però val la pena recordar que conèixer l'ètica no és ser ètic. Llavors, no importa si tenim totes aquestes declaracions i regulacions; hem d'ensenyar, educar els nostres fills a casa amb valors ètics. A l'escola poden complementar-ho, però és a casa on aprenen tot això.

La intel·ligència artificial ha de mostrar responsabilitat i transparència. Vull que en un futur pròxim les regles del govern per als nous productes basats en IA tinguin un cicle de vida similar al d'un químic per a farmàcia. Aquests medicaments es desenvolupen i tenen alguns anys d'assaig clínic, assaig social, i no només han d'estar regits per les empreses i el govern, sinó també per les organitzacions no governamentals. Si no, en perdrem el control. Els sistemes d'IA són artefactes que construïm i dissenyem. Siguem explícits en la manera amb què la IA ha de donar forma al nostre futur. Les màquines no són ètiques, és necessari dissenyar-les perquè estiguin èticament alineades, i quan són completament autònomes, poden necessitar un altre tipus d'ètica. Llavors, siguem-ne conscients, eduquem críticament i tècnicament la societat. Només les societats informades poden reclamar els seus drets i reclamar que es respectin.

Montserrat Pardo Bayona

Directora d'afers corporatius de Microsoft Espanya

Hola, bona tarda i gràcies al síndic de greuges per la invitació.

Avui us vull donar una mica la nostra visió, de Microsoft. La nostra missió és donar poder a totes les persones perquè puguin aconseguir més coses. I per què una missió? Jo al principi, jo us ho dic, jo era bastant anti-sistema americà i vaig voler entrar a Microsoft per entendre realment com es treballava en una companyia americana, i una de les coses que he après és que si no saps bé cap on vas, no trobes vents favorables. Llavors, saber cap on anem, saber realment el que volem a Microsoft cadascuna de les persones, perquè, al final, a Microsoft, l'únic que som és persones. El que volem és donar poder. Ja sigui a les persones, les organitzacions, les institucions públiques. I com ho fem? Doncs ara que parlem del canvi de la nova era, que realment la intel·ligència artificial el que està fent és pivotar i liderar aquest canvi d'era.

Realment el que nosaltres creiem és que la intel·ligència artificial no ve a substituir la intel·ligència humana, sinó que ve a amplificar-la, absolutament, i el que es dona és una sèrie de circumstàncies, que realment tenim una quantitat de dades, una capacitat de computació enorme i uns algoritmes... i Jordi Ribes, que és el nostre vicepresident en intel·ligència artificial, també originari d'aquí, de Manresa, sempre diu: "Ja estem arribant als resultats amb els algoritmes que estan trigant realment dies, coses que pensàvem que trigarien anys". Llavors estem en un moment realment molt important.

I què veiem? Que els països estan agafant el lideratge. Hi ha la Xina, hi ha França ... I on som nosaltres? També som aquí. Ja a Davos, l'any passat, el president Pedro Sánchez es va reunir i es va accelerar tot el que és l'estratègia de la intel·ligència artificial. I aquí, a Catalunya, la setmana passada, també presentàvem, conjuntament amb la Generalitat, amb l'Ajuntament de Barcelona, amb el Supercomputing Center, presentàvem el que és realment una cosa innovadora: el centre innovació en dades i intel·ligència artificial, que l'últim objectiu que persegueix és realment constituir Catalunya com una regió excel·lent. Que sigui realment un model al món que es pugui replicar i que mostri que realment aquí hi ha un ecosistema súper potent, de talent, de persones, d'empreses, d'institucions que realment volen augmentar la productivitat i la competitivitat d'aquesta regió.

Llavors hi ha reptes i hi ha responsabilitats. I aquí ... tot aquest seguit d'incerteses, de pors, què passarà? Perdrem llocs de treball? Tots aquests biaixos, pel que fa a gènere, pel que fa a races, que realment són temes de drets fonamentals, de drets humans. Què està passant? Això de la intel·ligència artificial en realitat el que mostra és el món en què vivim, i ho està transparentant, i la intel·ligència artificial el que fa és aprendre. Per això és tan important que, ara mateix, les màquines s'estiguin assemblant cada vegada més als homes. L'humanisme, l'ètica, els drets humans passen a tenir el major protagonisme en la història. Perquè, perquè això funcioni, haurem de treballar conjuntament i amplificar aquest enginy humà, tenint en ment aquests drets humans i aquesta ètica.

I què fem a Microsoft? Doncs intentar comunicar millor, perquè avui s'ha dit: "Jo només sé que no sé res". S'ha començat avui la sessió i és veritat, però si hi ha alguna cosa a Microsoft que hem après és aprendre a aprendre, i a aprendre que no ho sabem tot, sinó que volem ser els que en sabem i aprenem més ràpid. Però realment estem comunicant, intentant comunicar a la societat que efectivament això va d'oportunitats, d'eines, però que també hi ha unes certes amenaces, hi ha uns reptes. A mi m'agrada parlar de reptes. I us animo a llegir aquest llibre, perquè realment l'han escrit Brad Smith i Carol Ann, del nostre equip, i ells se senten amb tots els governs del món, escolten, intentem construir, i això és universal i atemporal. Al final del que parlem aquí, sigui d'intel·ligència artificial, en el passat parlàvem d'altres tecnologies,

és el bon ús o el mal ús que li donem a aquestes tecnologies i que per això al final si posem en el centre les persones, els drets humans, anirem en la bona direcció.

I, efectivament, no totes les intel·ligències són iguals. Les intel·ligències artificials, em refereixo, perquè no és el que la tecnologia pot fer, sinó el que ha de fer, i ens hem trobat moltes vegades, i en el nostre cas, a Microsoft, hem dit que no a certes institucions, perquè no hem considerat èticament que el projecte s'hagués de fer. Per què? Direu, no, Microsoft en intel·ligència artificial tindreu molts enginyers, moltes enginyeres ... Tenim en aquest equip d'intel·ligència artificial, en tota la companyia, molts antropòlegs, també, molts filòsofs, psicòlegs. De nou la multidisciplinarietat serà fonamental més que mai perquè això realment tingui èxit. Perquè la intel·ligència artificial i la innovació, per a què? Si no és per fer un món millor, que econòmicament sigui més competitiu, que socialment tingui i assegurí que tothom hi és inclòs i no es queda ningú enrere. Si no fem un món més sostenible, no ho estarem fent bé.

I què vam fer? I vam ser pioners, el 2016, fa quatre anys, i vam llançar aquests principis ètics i realment són atemporals, perquè avui veureu que parlem d'aquests principis ètics atemporals, perquè amb aquesta intenció es van fer: privacitat per disseny; seguretat, tots com a drets fonamentals; l'equitat, com a dret fonamental; la inclusió, com un altre dret fonamental; la transparència i la responsabilitat. I això no són paraules, són fets, perquè si no passem de les paraules als fets, al final de què et serveix ser una persona, si no tens valors? De què et serveix ser una institució, si no tens una cultura que ho reflecteixi?

Llavors, no entraré en cadascun d'ells, compartiré aquesta presentació, però que sapigueu de veritat que el tema de privacitat el considerem per disseny com a dret fonamental, i vam ser la primera companyia que, més enllà d'Europa, el tema de GDPR l'expandim globalment *by design* a tots els nostres països i és fonamental assegurar aquesta seguretat. I sabeu que cada vegada més hi ha atacs, ciberatacs, i és fonamental, ja no només en l'àmbit institucional, les estructures crítiques, sinó també les empreses, les grans i les petites, que són atacades dia a dia i fins i tot nosaltres com a ciutadans, i en el nostre paper, i que algú pot ser polític i alhora tenir la seva vida personal, evitar qualsevol tipus de vulnerabilitat en aquest sentit, no? La transparència, perquè si realment, com a companyia, no som transparents i transparentem totes aquestes decisions en què ens basem, no podríem guanyar aquest *trust*, aquesta confiança de tots. I per descomptat la responsabilitat, el que es diu en anglès *accountability*. I que tenim en el nostre ADN. I l'equitat. Perquè es parla molt de biaixos.

Nosaltres fem una gran tasca per tenir desenvolupadors i desenvolupadores, precisament per evitar tots aquests biaixos, que poden venir de gènere, però també, no només de gènere, també hi ha temes d'accessibilitat. Els nostres millors desenvolupadors, els que ens han fet millors, han estat persones invidents a la nostra companyia, que ens han fet la tecnologia molt més accessible. I tornem al tema i al debat dels drets humans. Per primera vegada, a Espanya, hem contractat una enginyera jove. Direu: Que bé! Enginyera, jove, la millor de la seva promoció... sordcega. Us asseguro que aquesta persona, la Sandra, ens ha fet molt millors. Pel que fa a les millores en la nostra tecnologia, però sobretot en la manera en què treballem avui en dia a Microsoft a Espanya. I després, la nostra responsable en diversitat i accessibilitat és sorda, en corporació. Llavors toquem tots aquests eixos i la diversitat i l'equitat les entenem en el sentit ampli. La discapacitat, veureu que és recurrent en el nostre cas, perquè realment la intel·ligència artificial ens està permetent avenços increïbles. Avui dia és possible per a les persones invidents fer el Camí de Santiago, cosa impensable abans, no? I ho poden fer amb l'ajuda de les tecnologies.

I només estem començant, efectivament. I això no ho estem fent els desenvolupadors i les desenvolupadores de Microsoft. Això ho fa l'empatia. L'empatia de persones que saben com trobar una solució d'intel·ligència artificial, i que escolten metges o

que escolten aquests pares d'aquest nen amb autisme. I realment la conjunció, la interacció, de nou, del món educatiu, per exemple, amb la intel·ligència artificial, és el que farà la màgia que es produeix en aquests casos.

Per què la diversitat? I per què la defensem tant a Microsoft? Perquè hem entès que no hi ha diversitat, no hi ha innovació, sense diversitat, i que si la nostra força laboral no és un reflex de la societat ens estarem equivocant. Per això la defensem tant. I no voldria estar aquí i no parlar d'aquesta necessitat de tenir més dones en el nostre sector, perquè realment si volem una intel·ligència artificial que reflecteixi la societat, hem de aconseguir-ho. I que aquestes dones de fer servir la tecnologia passin realment a desenvolupar-la, a treballar-hi i a desenvolupar-la.

Aquí, i a Barcelona especialment, i gràcies a tota la tasca que estan fent les diferents institucions i el Mobile, també, i a través també de projectes com Barcelona Digital Talent o Digital Future Society, us vull llançar aquesta *calltoaction*, serà fonamental, aquí, a Barcelona, a Catalunya, a Espanya, a tot Europa i el món. En aquesta cursa per realment fer una societat millor i que ningú es quedi enrere, que tots estiguem preparats. I quan dic tots, començant per nosaltres, però a tots els nivells i totes les audiències. Serà un *reskilling* que es diu *upskilling* i *skilling* en tots els sentits. Però jo aniria més enllà. Pel que ens hem de formar és a la *learnability*, en la capacitat d'aprendre de forma contínua i jo us diria fins i tot més, *skilltoskill*, *to be readytoskill*, no? Que estiguem sempre preparats, capacitats per capacitar-nos constantment, perquè no sabem realment el que demanarà el mercat, perquè això va molt ràpid.

Però sí que sabem que desapareixeran llocs de treball i que hauréu d'estar preparats, perquè n'apareixeran d'altres. I al seu dia ja en parlàvem, d'això, no? I no cal tenir por, perquè totes les revolucions i els canvis d'era han generat haver d'aprendre i haver de canviar els models, però realment seran aquestes societats les que adoptin abans aquests canvis, les que seran més avançades. I per això us deia abans, us mostrava la Xina, us mostrava Macron, però nosaltres també som aquí. I aquí, a Barcelona, hi ha un ecosistema que és increïble el que està succeint aquí. Hem de aprofitar-ho, simplement unes pinzellades. Clarament a LinkedIn ho veiem, perquè hi ha intel·ligència artificial aquí, que realment aquestes ja són les capacitats més demandades i ho veiem perquè efectivament hi ha totes aquestes dades que, per al benefici de tots, es poden utilitzar precisament per formar les persones en el que per primera vegada demanda el mercat. Perquè moltes vegades hem format aquestes persones en coses que no es demanaven i, aquí, simplement, a tall d'exemple, veieu les diferents audiències i tots ens tenim, des dels cecs fins realment les persones i els col·lectius més vulnerables, perquè de veritat que el nostre futur depèn d'això. I aquesta és la nostra ambiciosa aspiració. Arribar a deu milions en els propers cinc anys de persones i que ningú es quedi enrere en aquest *skilling*.

I això d'AI for good (IA per a bé), de nou, m'agradaria que us quedeu amb aquest missatge: intel·ligència artificial, per a què? Jo em pararia d'aquest món si realment les tecnologies no ens han d'ajudar a fer un món millor. I un món millor en tres eixos. Us posaria, per posar un focus avui, a banda del de sanitat, que és molt important, també, però per a un món millor planeta, per a realment l'accessibilitat i per al tema humanitari. I la intel·ligència artificial realment i hem llançat cinquanta milions per a institucions, ONG, *start-ups* que realment estiguin treballant en projectes en aquests quatre eixos, en intel·ligència artificial per al planeta. Ella és una persona, una companya nostra, que ens ha ajudat molt a generar aquest tipus de programes que, realment, en els propers cinc anys, també volem dedicar a projectes d'accessibilitat, quaranta milions destinats a institucions de manera que desenvolupin. Volem aquesta intel·ligència artificial per al bé comú.

I no voldria deixar aquest final sense aquesta part, perquè realment hi ha aquí una responsabilitat molt molt gran, en tots aquests casos. Cada vegada tenim més desastres naturals amb el tema del canvi climàtic. Cada vegada tenim més infants que estan en posicions molt vulnerables i temes de refugiats. I la intel·ligència artificial ens pot ajudar.

No ho podem fer des de Microsoft. No som tan gegants com sembla. A Catalunya, per donar-vos d'una perspectiva, som setanta persones. Setanta persones que realment el que generen és un ecosistema de pimes, que són les que realment innoven mitjançant la nostra tecnologia, que són quaranta mil persones aquí, a Catalunya, i dues-centes mil en el cas d'Espanya. Aquests són els que ens fan grans i aquests són els gegants. Aquestes pimes tecnològiques, aquestes institucions que desenvolupen a partir d'aquesta tecnologia base.

I el que s'ha dit, això no ho guanyarem si no fem coses com les que s'estan fent aquí al voltant de Barcelona, on estem unint esforços des de les institucions, inclòs també en molts casos el mateix Govern central juntament amb la Generalitat, juntament amb els centres d'innovació, juntament amb les empreses i la societat civil, que realment estem transformant i tenint en ment aquests valors ètics, aquests principis, aquests drets humans, que si no els tenim i no els posem al centre, ens equivocarem. Així que us convido a crear aquest futur i la millor manera de predir el futur és crear-lo conjuntament. Així que pintem-lo tots junts i us deixo aquí les meves dades per si desitgeu continuar la conversa. Moltes gràcies!

Alfonso González Herrero

Director de Government & Regulatory Affairs, IBM Spain

Intel·ligència artificial i drets humans. Permeteu-me primer fer-vos una breu descripció del que és IBM. És possible que tingueu en ment una imatge d'IBM que no és gaire actual, ja que l'empresa està en constant evolució. Probablement, si haguéssim fet aquesta conferència fa cinc o deu anys, IBM hauria estat molt diferent del que és avui en relació amb la intel·ligència artificial. IBM es va fundar fa 109 anys, i vull mostrar això en una primera diapositiva perquè no som una empresa fundada fa deu o vint anys. No som una empresa fundada fa setze anys (com algunes altres...). Tenim una història molt més llarga, vam cometre els nostres errors, en vam aprendre i sabem bé el que es necessita per fer les coses correctament.

Durant vint-i-set anys seguits hem liderat el rànquing de patents als EUA. Continuem sent el número u en el rànquing d'Interbrand en termes de serveis empresarials. Som, essencialment, una empresa orientada a la venda de productes i serveis entre empreses (*business-to-business company*). No som en el mercat de consum, tot i que molts dels nostres clients sí que hi són; això significa que venem a unes altres empreses. I avui liderem en diverses àrees en què no érem presents fa alguns anys, com ara la informàtica en el núvol (*cloud computing*), la intel·ligència artificial, la internet de les coses (*IoT*), l'analítica empresarial. També fa més de cent anys que som a Europa. Tenim oficines i centres a tot Europa. Tenim un centre de dades aquí, a Barcelona. Tenim la seu europea a Madrid. Des de Madrid dirigim no només les operacions espanyoles, sinó també les portugueses, gregues i israelianes. Per tant, sabem quins són els valors europeus. Som europeus i fa més de cent anys que som a Europa, i aquest és un altre missatge clau que us vull transmetre. Com deia, som a Espanya, no fa cent anys, però gairebé, tant a Madrid com a Barcelona fa al voltant de noranta-quatre anys, i tenim centres de responsabilitat internacional a Madrid i a Barcelona i oficines a Sevilla, València, Bilbao, La Corunya, etc.

Així doncs, es tracta d'IBM, i com us deia abans, el nostre centre avui, el nostre enfocament, en termes de negocis, es troba en el *big data* i en l'anàlisi, en la informàtica en el núvol i, per descomptat, en la intel·ligència artificial, que és el que debatem aquí avui. Ara parlem d'intel·ligència artificial, però moltes vegades ens referim a conceptes que, i probablement el professor Cortés d'això en sap molt més que jo, no són el mateix. Em refereixo a tots els termes, els diferents conceptes que utilitzem per a la intel·ligència artificial quan parlem d'elements diferents. I aquí (en aquesta diapositiva) teniu alguns d'aquests conceptes: aprenentatge automàtic, robòtica.

A IBM preferim parlar sobre *els últims noms que apareixen a la diapositiva: intel·ligència augmentada i computació cognitiva*. Per què? Perquè aquests termes posen l'èmfasi en la intel·ligència artificial com una tecnologia que és aquí per ajudar els humans, en contraposició al terme artificial, que sembla que és una mica oposat a *natural*, com si hi hagués una intel·ligència natural i una intel·ligència artificial, i aquesta semàntica és probablement part del problema quan ens acostem a la intel·ligència artificial. Creiem que pensem, que triem una cosa que s'oposa a la nostra intel·ligència natural.

Com ha dit abans el professor Cortés, el concepte intel·ligència artificial es va utilitzar per primera vegada el 1956 durant una conferència, crec que va ser el senyor de la samarreta vermella de la diapositiva, John McCarthy, si no recordo malament, qui va crear aquest terme. Hi havia dos membres d'IBM en aquell moment en aquesta conferència. Dues persones que van treballar per a IBM. El senyor de l'esquerra, Trenchard Habiti, és un d'ells. Aquesta és una foto presa fa uns quants anys per commemorar un curs d'estiu a Dartmouth College el 1956. Per tant, IBM ha estat implicada en la intel·ligència artificial des del principi. Des de l'inici a Dartmouth College. Nosaltres sabem de què parlem. Recordareu alguns dels passos més emblemàtics de la intel·ligència artificial, com el 1997, quan el sistema IBM Deep Blue va derrotar Kaspàrov. I aquest va ser un moment molt icònic, perquè hi ha la creença que els escacs és un dels jocs en què els humans pensem que juguem millor que les màquines. El 2011, a Jeopardy, un programa molt famós als Estats Units, que s'ha transmès durant més de trenta anys, també una màquina d'IBM, Watson, va derrotar els dos concursants més d'èxit en aquest moment. Avui estem implicats en el que en diem Project Debater, que implica ensenyar les màquines a debatre com a éssers humans sobre qüestions complexes.

Actualment, què fem en el negoci? Us donaré alguns exemples dels últims tres mesos: Tenim més de 20.000 projectes sobre intel·ligència artificial al món en aquests moments, a IBM. Per exemple, tenim un assistent d'Orange a Espanya que ajuda els clients a obtenir respostes a través de WhatsApp. Un equip de Barcelona, fa amb prou feines un parell de mesos, va guanyar un concurs internacional per a IBM, anomenat Call for Code, que ajudarà els bombers a estar més segurs quan apaguin un incendi. Ajudem Cerealto Siro Foods a posar nous productes en el mercat, analitzant les tendències del consumidor i posant en el mercat productes que tinguin en compte aquestes tendències. I, per descomptat, estem en el ram de la salut amb el nostre sistema Watson Health intentant ajudar a aconseguir un millor diagnòstic per als metges. Ajudar els metges a fer la seva feina, no reemplaçar-los.

La intel·ligència artificial ja s'utilitza en la nostra vida diària. La utilitzem quan paguem amb la nostra targeta de crèdit; quan utilitzem el sistema al nostre automòbil per conduir del punt A al punt B; en els nostres correus electrònics, els filtres de correu brossa utilitzen intel·ligència artificial; quan utilitzem el nostre telèfon, SIRI per exemple, o quan fem una cerca a Google. Llavors, veiem que ja és aquí. No parlem del futur. La intel·ligència artificial ja és aquí i la utilitzem cada dia. Ara diré una paraula que ja s'ha esmentat aquí: *confiança*. Aquesta és la clau per a l'adopció de la intel·ligència artificial. Si no confiem en els sistemes, aquesta tecnologia no evolucionarà. És a dir, confiem en el nostre cotxe perquè sabem que el cotxe frena quan trepitgem el pedal, i entrem en un ascensor perquè sabem que, si premem el número cinc, l'ascensor s'atura al cinquè pis, o anem a un caixer automàtic i traiem diners del caixer automàtic perquè sabem que, si demanem cent euros, la màquina ens en donarà cent i no cinquanta. Llavors, la confiança és al darrere de tota aquesta tecnologia.

Algunes de les preguntes que ens fem quan parlem d'intel·ligència artificial: primer, és justa? Vull dir, hi ha algun biaix? És fàcil d'entendre? Entenem com funciona aquest sistema? Algú l'ha manipulat? Vull assegurar-me que aquest és un sistema sòlid i que ningú ha fet res d'il·legal o que m'hagi posat en perill. I és responsable? Em refereixo que, si alguna cosa surt malament, hi ha algú darrere que en sigui responsable?

Per tant, sí, també tenim els nostres principis. No estic segur si publicats el 2016 o abans, no discutiré sobre això, tenim els nostres principis, i aquests principis giren al voltant del que sempre ha estat clau per a IBM, que és la confiança i la transparència. Aquests són bàsicament els tres principis (a la diapositiva): primer, assegurar-se que la intel·ligència artificial sigui aquí per ajudar els humans, per augmentar la intel·ligència humana, no per reemplaçar-nos, i que hi ha alguns conceptes al darrere, com ara supervisió humana, etc. Les dades i els coneixements pertanyen al seu creador. El nostre model de negoci és força diferent del d'algunes de les plataformes que esmentava abans i posava en pantalla el professor Cortés. No guanyem diners amb les dades dels nostres clients. Oferim una tecnologia, els nostres clients es basen en aquesta tecnologia, però ens assurem que conservin els seus drets sobre les seves dades. I aquest és un model de negoci molt diferent dels de les plataformes en el mercat d'empresa a consumidor.

Ara, el tercer principi: les noves tecnologies que inclouen, per descomptat, els sistemes d'intel·ligència artificial han de ser transparents i explicables. Vull dir, no se'ns pot demanar com a humans que confiem cegament en una caixa negra. Llavors, necessitem transparència, d'acord? Per tant, demanem regulació? Sí, demanem regulació. No qualsevol tipus de regulació. Demanem una normativa de precisió. Creiem que no totes les empreses s'estan comportant de la mateixa manera, que hi ha algunes empreses que han creat un problema i aquest problema s'ha d'abordar. Però no tot l'ecosistema digital necessita o ha de patir aquesta regulació. Per tant, demanem responsabilitat, demanem transparència, i demanem justícia i seguretat. En altres termes, una regulació de precisió. La mateixa tecnologia utilitzada per diferents persones en diferents contextos s'ha de regular de manera diferent. Aquesta és bàsicament la idea que hi ha darrere. Hem estat col·laborant, llavors, no es tracta d'autoregulació, és coregulació.

Hem estat col·laborant en les Directrius ètiques per a una intel·ligència artificial fiable, publicades l'abril de l'any passat per la Comissió Europea. Vam donar la benvinguda als principis de l'OCDE per al desenvolupament i l'ús de la intel·ligència artificial, per exemple. Hi ha una consulta oberta sobre el Llibre blanc de la intel·ligència artificial, de la Comissió Europea, una consulta que estarà oberta fins a finals de maig, mitjan maig, i en què contribuirem amb els nostres comentaris. I penso que si estem aquí debatent sobre com els ombudsmen poden contribuir a aquest debat, aquí tenen una oportunitat en què poden participar amb les seves idees.

Com podeu veure en la diapositiva següent, estem implicats en unes altres iniciatives. Som membres del Future of Life Institute. També hi estem treballant juntament amb unes altres empreses, com Microsoft, etc. Som fundadors del Partnership on Artificial Intelligence, perquè mostrar lideratge en una indústria no és només desenvolupar les tecnologies més avançades. Vull dir, el lideratge també està intentant fer que la indústria es comporti de manera correcta i és per això que estem intentant implicar-nos en totes aquestes iniciatives. I, per descomptat, la Iniciativa vaticana, que abans esmentava el professor Cortés.

Ara fem això, però també prediquem amb l'exemple. Ja tenim productes en el mercat que ajuden unes altres empreses a actuar de manera justa i a assegurar-se que els productes que llancen al mercat, els productes d'intel·ligència artificial, no tenen biaixos i són transparents. En tenim un parell d'exemples, com ara Artificial Intelligence Fairness 360 i Watson OpenScale. El de l'esquerra, més pensat per als desenvolupadors i científics de dades, i el de la dreta, Watson OpenScale, per a usuaris de negocis. Tenia un vídeo preparat. No sé si és possible posar-lo. El de la dreta, Watson OpenScale. És un vídeo de dos o tres minuts, no estic segur de si està preparat per ser emès. El podeu veure a YouTube i podeu veure en termes pràctics com un producte ja pot ajudar les empreses a provar les solucions d'intel·ligència artificial que comercialitzen, per assegurar-se que la seva transparència ajudi les empreses a saber si hi ha cap biaix i a oferir transparència en cas que un regulador preguntí per què s'ha pres una determinada decisió.

I acabaré amb una sola frase que el nostre actual CEO, Ginni Rometty, va dir l'any passat a Davos i que reflecteix el que la nostra empresa pensa sobre la intel·ligència artificial i és molt clara: "Les empreses que no poden explicar les decisions preses per la intel·ligència artificial en els seus productes han de sortir del mercat, no s'ha de permetre que estiguin al mercat". Aquesta és la posició d'IBM i aquest és el final de la meva intervenció. Moltes gràcies!

Oliver Smith

Director d'estratègia i cap del Departament d'Ètica a Alpha Health

Gràcies, Andreas. Gràcies al Síndic i a l'IOI per convidar-me avui. Vull fer un agraïment també als traductors, que crec que tenen la feina més difícil i que són la prova viva que la intel·ligència artificial encara no ha arribat del tot.

El gran missatge que us vull fer arribar avui és que l'ètica en la intel·ligència artificial és molt difícil. Tot i que això probablement ja ho sabíeu, no és nou. Però no és difícil per l'aspecte tecnològic, tot i que aquest és un aspecte complicat, i per descomptat cal aprendre'l, sinó que és difícil per l'aspecte humà. Perquè són els valors humans els que pensem i implementem. Són els humans els que creen una nova tecnologia d'usuari. I crec que vosaltres, com a "ombudspeople" (no estic del tot segur si aquesta paraula existeix, "ombudspeople"), homes i dones, enteneu la gent, i això és el que esteu acostumats a tractar. Per tant, crec que podeu confiar més en el que esteu fent, ja que esteu intentant superar els reptes de la intel·ligència artificial. Per tant, intentaré parlar molt més sobre una mena de reptes pràctics als quals ens hem enfrontat en incorporar intel·ligència artificial als nostres productes, per entendre després com poden ser ètics o no.

Primer us donaré una mica de context sobre Alpha i sobre el que hi fem. Després us comentaré el que estem fent en matèria d'ètica i finalment us parlaré d'alguns d'aquests reptes. Alpha Health forma part de Telefónica. Per als que no sou d'Espanya, Telefónica és una gran empresa de telecomunicacions espanyola que també té presència al Regne Unit i a Alemanya, i a molts països de l'Amèrica del Sud. Alpha es va crear el 2016 per desenvolupar innovació a molt llarg termini, per intentar abordar alguns dels grans reptes als quals s'enfronta el món i per trobar-hi solucions. I tenim la seu aquí, a Barcelona. Per als que viviu aquí, aquest és el nostre edifici. Els dies que fa molt vent, com avui, es balanceja una mica, és la primera vegada que em passa. És divertit.

A Alpha Health, la nostra missió és donar a mil milions de persones cinc anys addicionals de salut, fent que adoptin hàbits més saludables. I per fer-ho, volem crear un assistent de salut personal per a totes les vostres necessitats de salut que realment us entengui a vosaltres, a la vostra salut, al vostre comportament i a allò que l'impulsa, i que després us aconselli i us ajudi a construir la via més eficaç perquè pugueu ser més feliços i estar més sans. Això, doncs, és el que intentem fer, i per donar-vos una idea de com ho fem, actualment tenim quatre productes o prototips complets a la nostra cartera. I tots quatre tenen intel·ligència artificial integrada.

Comencem amb aquest producte de l'esquerra, Evermind, que està pensat per a l'autocura. És per a persones que tenen estrès i per ajudar-les a combatre aquest estrès. Es tracta d'intentar prevenir problemes de salut en primer lloc. Però, per descomptat, no només ens podem centrar en la prevenció, de manera que també tenim un parell de productes, Perspectives i Mindset, que s'ocupen del tractament. Aquests dos utilitzen una tècnica anomenada *teràpia cognitivoconductual*, que es demostra molt, molt bé en la bibliografia sanitària, i funcionen a través del telèfon mòbil. I finalment, a la dreta, també voldríem predir si la gent estarà malalta. Per tant, tenim un producte anomenat Foresight, que estem creant en col·laboració amb el Servei Nacional de Salut del Regne Unit, i que utilitza en primera instància dades

electròniques de la història clínica, és a dir, la història clínica del pacient, per formar un algorisme de manera que pugui predir les crisis. Perquè, per descomptat, a l'individu li agradaria no tenir la crisi. A l'hospital també li agradaria que no tingués la crisi, perquè així es pot proporcionar una millor atenció i evitar totes aquelles despeses, dolor i càrrega a l'individu, al sistema i als ciutadans. I, per descomptat, pensem que les versions més potents d'aquestes eines seran les que utilitzin intel·ligència artificial, de manera que es puguin personalitzar.

L'avantatge de la tecnologia és que pot funcionar a distància, la qual cosa és ideal en l'àmbit de la salut, en particular per a la salut mental, en la qual ens centrem molt, perquè hi ha moltíssimes necessitats no cobertes. Però si bé la tecnologia pot utilitzar-se a distància, un dels seus grans problemes és que això no és gaire personal i, per descomptat, no hi ha res millor que ser davant del metge, cara a cara, i sentir-se cuidat, i rebre un assessorament personalitzat i adequat a cada persona, i això és el que volem que faci la nostra tecnologia. Amb Evermind, per exemple, la intel·ligència artificial intenta entendre quin és l'estat emocional de la persona en aquell moment. Està estressada? Llavors què li podem donar? Quina intervenció podríem fer? Quins consells li podríem donar que li anessin bé ara mateix? De la mateixa manera, amb Perspectives i Mindset, a causa del que acabo d'esmentar, estem intentant predir si algú està a punt de tenir una crisi.

Per tot això, creiem que no funcionaran tret que siguem ètics, tret que es pugui confiar en nosaltres, i hi ha tres raons per això. En primer lloc, la credibilitat. Quan vam parlar amb els metges del servei nacional de salut amb els quals treballem, van dir: "Sabem que podeu fer tot tipus de coses enginyoses amb la intel·ligència artificial, però no les utilitzarem tret que ens expliqueu per què el producte ha fet una determinada predicció. Per què ha dit que hauria de parlar amb aquests pacients, perquè en dues setmanes tindran una crisi. Perquè si parlo amb aquests pacients, no puc parlar amb aquests altres. Per tant, heu d'explicar-me per què passa això."

La segona raó està relacionada amb l'eficàcia. Per exemple, Evermind ajuda les persones a gestionar l'estrès. Un dels procediments és que les persones escriuen com se senten, i aquest és un producte que oferim a les empreses per ajudar el seu personal a combatre l'estrès. Per tant, si una persona, quan escriu un missatge sobre el seu cap, no confia que l'aplicació no enviarà aquestes dades al cap, evidentment no utilitzarà aquesta aplicació. No l'utilitzarà i no serà tan eficaç. Per tant, hem de crear aquesta confiança perquè la faci servir.

I, finalment, per descomptat, ja era d'esperar que diria això, és d'esperar que qualsevol que treballi en intel·ligència artificial digui això: Accés a les dades. Evidentment, si no generem confiança, no tindrem accés a les dades que necessitem per formar els algorismes i, per tant, no podem aconseguir la personalització que volem perquè siguin realment efectius.

L'ètica, doncs, és fonamental per a tot el que fem. Pel que fa a com l'apliquem, també tenim un marc ètic. Hem analitzat molts dels marcs que existeixen, tant els de la Unió Europea o l'OCDE com els d'algunes altres empreses, perquè Microsoft i IBM també fan una gran feina. I també hem analitzat la Declaració universal dels drets humans i la Carta europea dels drets humans. I ho hem agafat tot plegat i hem elaborat cinc principis en el nostre marc, que estan relacionats amb la felicitat i la salut per a tothom, és a dir, sense biaix; estan relacionats amb el control i l'autonomia de les persones usuàries, amb la comprensió i l'explicabilitat, amb la seguretat i la privadesa i, finalment, amb la governança.

En lloc d'entrar a analitzar cada un d'aquests principis en detall, el que vull és aprofundir en alguns exemples que estem posant en pràctica i després, com he dit abans, parlaré d'alguns dels reptes als quals ens hem enfrontat en el procés. Així doncs, aquests són tres exemples que m'he atrevit a etiquetar com a "èxits inicials".

Odio els apartats de termes i condicions i odio llegir-los, haver de pensar a llegir-los i després sentir-me culpable de no haver-los llegit perquè acabo de prémer "sí" i he tirat pel dret, perquè només vull arribar al que m'interessa. Per tant, sabem que els apartats de termes i condicions són terribles; les polítiques de privadesa són terribles. I hem de fer-ho millor, i això forma part de la nostra feina, intentar crear nous dissenys, noves formes de fer arribar aquesta informació a la gent. La imatge que mostro aquí és per a una eina per al son que tenim a Evermind. Volíem transmetre a la gent molt més clarament quines dades recopilàvem. Per tant, vam dir: "Farem una cosa visual, hi posarem una foto d'una persona al llit en què es vegi quines dades està recopilant el telèfon". I aquest és un exemple d'entre els que estem intentant posar en marxa, en què intentem utilitzar un llenguatge molt senzill, mitjançant il·lustracions sempre que podem. També combinem el procés de consentiment i el procés de termes i condicions. No cal que s'exposi tot d'entrada. L'experiència que tenim la majoria de nosaltres és que obrim l'aplicació i de seguida ens demana que ho aprovem i consentim absolutament tot, i després ja no tornem a veure aquesta informació mai més, mentre que les empreses líder intenten anar més enllà.

El segon exemple del qual vull parlar és la intel·ligència artificial explicable, i alguns companys de la taula ja n'han parlat. He mencionat que per al Servei Nacional de Salut del Regne Unit això era molt important, i l'Ulises tenia tota la raó quan ha dit que l'aprenentatge profund no és l'única forma d'aprenentatge automàtic, per tant no us centreu només en això. Però, per descomptat, el repte que hauran d'afrontar les persones, al qual ens vam enfrontar amb el personal mèdic del Servei Nacional de Salut, és el següent: D'acord, això descarta l'aprenentatge profund. Ja ens estava bé en aquest cas, podíem utilitzar altres formes d'aprenentatge automàtic. Però no volíem descartar l'aprenentatge profund per sempre, perquè és una tècnica molt potent. Per alguna raó la gent en parla tant.

I crec que vosaltres, com a defensors i defensores, teniu aquí un repte amb el qual heu de lluitar, i és que podeu tenir el que s'anomena *explicacions locals d'una xarxa neuronal*, d'una xarxa neuronal d'aprenentatge profund. I és fantàstic si sou l'usuari final i us pregunteu: "Per què ha fet això? Per què el sistema m'ha recomanat això en aquest cas concret?" Tot i això, probablement voldreu un model més global que entengui i pugui explicar la totalitat del funcionament d'un algorisme. I això encara no existeix. Hi ha sistemes que poden donar l'explicació local, però es poden enganyar. Hi ha alguns articles recentment a *IA, Ethics and Society*, de Nova York, que demostren que es pot fer creure a la gent que un algorisme esbiaixat en realitat no n'és, d'esbiaixat. Per tant, una de les coses que fem és dir: podeu crear un model global per a xarxes neuronals d'aprenentatge profund? I ho podem fer fins a cert punt, per a algunes xarxes neuronals. Encara no podem fer-ho per a xarxes neuronals enrevessades, però hi estem treballant. Però crec que aquest tipus de treball és molt important per a vosaltres, com a defensors i defensores.

L'exemple final són les auditories externes. I, de nou, ja s'ha parlat d'això, aquesta noció de responsabilitat i responsabilitat externa és importantíssima. I, mentre hi reflexionàvem, vam reconèixer que no ens podíem autoavaluar. Per tant, vam encarregar a una empresa local, *Eticas Research and Consultancy*, que treballa estretament amb la Universitat Pompeu Fabra, també aquí a Barcelona. I auditen la nostra feina, i una de les auditories que van fer recentment, l'any passat, de fet, va ser sobre un dels nostres prototips. El van analitzar per comprovar si era esbiaixat, i ens van respondre: "No, no n'és, d'esbiaixat, està bé". I això és molt important, però, per descomptat, hem d'anar més enllà i publicarem aquestes auditories aviat, però no podem, com he dit, avaluar la nostra pròpia feina.

Em penso que del que realment voleu sentir parlar és de les coses que no han anat tan bé. I també n'hi ha tres que vull mencionar. La primera és: els principis

poden entrar en conflicte. Resulta que en realitat els principis ètics poden entrar en conflicte. I tornaré a l'exemple que acabo d'esmentar. Aquesta auditoria sobre el biaix. Fantàstic, Èticas i la UPF ens diuen: "No és esbiaixat". Però, de fet, el que van dir inicialment, sobre aquest prototip que vam construir, va ser: "En realitat, no sabem si és esbiaixat, perquè heu fet una feina tan bona en no recollir cap informació personal, que no podem demostrar-vos si és esbiaixat o no". I, és clar, tan aviat com ens van dir allò vam pensar: "Això que dieu és increïblement obvi, com és que no hi vam pensar des del principi?" No ho vam fer perquè ens vam centrar molt en la privadesa i a assegurar-nos que realment la complíem, com a principi, però després, pel que fa al biaix, això volia dir que no podíem fer-hi res directament.

Afortunadament, en aquest cas, Èticas i la UPF van poder examinar algunes proves indirectes, els comentaris que s'havien fet a Google Playstore, ens van entrevistar i van examinar tota la nostra documentació, i van dir: "Creiem que en aquest cas, aquesta aplicació, aquest algorisme, no estan esbiaixats contra grups particulars. Per tant, podeu estar tranquils, però heu de pensar per endavant en els vostres diferents principis". I també van dir: "No podeu culpar el Reglament general de protecció de dades (RGPD)". Perquè, per descomptat, sempre que parlo amb els nostres advocats em diuen: "Hem d'assegurar-nos de no recopilar cap dada personal". Però l'RGPD també parla de la importància de la precisió i la importància de l'equitat. Per tant, no podem dir només: "Bé, com que seguim l'RGPD, no hem estat capaços de donar cap prova directa". Per tant, en realitat hem d'acceptar que els principis entren en conflicte i cal tenir-ho en compte des d'un primer moment.

El segon repte al qual ens hem enfrontat recentment és del tipus "tot o res" de la confiança i l'ètica. Podem desenvolupar els nostres principis, podem desenvolupar les diferents parts del pla que hem d'implementar i, per descomptat, us podeu imaginar un bonic gràfic de Gantt que digui: d'acord, aquesta setmana privadesa, la setmana que ve biaix, i la setmana següent farem alguna cosa sobre explicabilitat. Però no es pot intentar fer una cosa perfectament, després passar a la següent i després a la següent, perquè us trobareu que, quan parleu amb persones externes, us diran: "Oh, espera, no has fet això". Per posar-vos un exemple més concret, quan treballàvem amb el Servei Nacional de Salut del Regne Unit ens vam centrar en la privadesa, ens vam centrar en l'explicabilitat per al personal mèdic i infermer de l'hospital, però no ho vam fer tan bé amb la transparència. En realitat, no en parlàvem gaire fora de l'hospital local o fora de la comunitat local. Podríem haver fet una feina millor si haguéssim parlat amb gent a escala més nacional. I d'això ens n'hem adonat recentment, quan van aparèixer als mitjans anglesos moltes històries que en parlaven i es preguntaven què fèiem. I, és clar, no parlar d'una cosa d'entrada fa que la gent pensi que amagues alguna cosa. Així doncs, estàvem posant-nos al dia i dient "No, no, està bé. No amaguem res. Això és el que fem, aquests són amb els qui parlem, parlem amb el regulador, no farem res fins que no hagin aprovat la fase següent". Però vam entendre que s'ha d'anar avançant en tots els aspectes alhora.

I, finalment, "perfecte contra prou bo". I crec que aquest és un repte que només turmenta la feina en intel·ligència artificial i ètica. He parlat abans dels termes i les condicions i de la política de privadesa. Tenim un conjunt de polítiques millor que el que teníem, però no són com haurien de ser. Tinc un debat permanent amb un dels nostres advocats en matèria de privadesa, i Alpha compta amb el suport de Telefónica, de manera que tenim la sort de poder utilitzar els seus advocats, i discutíem fins a quin punt podem simplificar el llenguatge dels apartats de termes i condicions i polítiques de privadesa. El seu argument és: "Bé, és clar, hi ha un llenguatge tècnic que cal que hi sigui forçosament". I el meu argument és: "Però crec que hem de ser capaços d'explicar-ho a la gent d'una edat lectora d'onze anys. I segur que això ha de ser possible perquè, creu-me, si m'ho pots explicar a mi, que

probablement tinc una edat lectora d'onze anys pel que fa als contractes legals, llavors hem de ser capaços de fer-ho per a tothom". I encara no ho hem resolt. Però si hagués esperat fins a tenir-ho solucionat del tot i haguéssim redactat un contracte legal comprensible, llavors ens hauríem encallat en la versió més tradicional.

Així doncs, hauríem de moure'ns ara i tenir una cosa imperfecta o hauríem d'esperar? I no hi ha una resposta correcta a això. I hi ha una segona part de "perfecte versus no prou bo". De vegades no hi ha una resposta correcta, de manera que quan la gent parla de biaix i algorismes imparcials, és un àmbit que cal debatre perquè, possiblement, per a un sistema de classificació, per a un sistema recomanat, es necessita un biaix perquè en certa manera d'això es tracta. Es pot voler distingir unes persones d'unes altres, un grup d'un altre. En realitat, cal aprofundir una mica més, cal entendre què vol dir realment determinada organització o empresa quan parla de biaix. A quins grups es refereix? I quins prejudicis preveia que hi podria haver com a conseqüència que hi hagués un biaix o no? Per tant, crec que hem d'anar més enllà d'aquestes nocions de perfecció i viure al món real.

I vull acabar tornant al que he dit abans, al començament, sobre que l'ètica en intel·ligència artificial és difícil. Una de les coses que més em molesten, un dels aspectes que m'irriten més en el diàleg actual sobre ètica, és que la gent té aquesta... hi ha dues coses. Tornem una mica enrere. La primera és que quan vaig a conferències acadèmiques i quan parlo amb els meus estimats companys d'informàtica, sovint amb aquest debat sobre ètica en la intel·ligència artificial simplement estan descobrint l'ètica i estan descobrint que és bastant difícil, i el que realment volen és tenir un algorisme per l'ètica. "Dona'ns els càlculs". Per tant, sovint es trobaran amb un enfocament utilitarista o conseqüencialista, i això és fantàstic! Significa que hi ha un resultat final per al qual puc optimitzar. M'encanta, i sé com funciona, i té tot el sentit. I llavors algú diu: "Aha! Però heu llegit Kant? I heu pensat en la deontologia?". Si s'ha pensat en un enfocament més basat en regles. Ah d'acord! Per tant, això significa que... no, tot és bastant complicat. Per tant, hi ha una part del debat que crec que és... el que tot just comencen a entendre els informàtics, que és que no hi ha una resposta perfecta en ètica. No l'hem tingut durant milers d'anys, i no la trobarem de sobte només perquè hem introduït la intel·ligència artificial.

I la segona part frustrant del debat en aquest moment és una cosa de la qual crec que les empreses són especialment culpables, i és aquesta idea que simplement es pot... afegir ètica, oi? Tenim aquest producte excel·lent i només afegirem un grapatet d'ètica i llavors serà perfecte. I això no és així. Cada empresa, cada organització té un conjunt de valors que està implementant i té una ètica que utilitza. Per tant, crec que la vostra pregunta és, defensors i defensores, ho sabeu? I hi heu pensat? Moltes gràcies pel vostre temps.

Carina Lopes

Responsable de la Digital Future Society Think Tank (Mobile World Capital)

Gràcies. Després d'aquestes meravelloses presentacions, m'agradaria portar-vos de tornada al nucli de la nostra sessió, que és la gent. Som aquí per parlar sobre els ciutadans i els seus drets, i què succeeix quan els sistemes d'intel·ligència artificial afecten el seu dia a dia. M'agradaria que féssim un exercici col·lectiu, perquè de vegades és realment un repte parlar sobre intel·ligència artificial i ètica i valors sense comprendre realment les conseqüències que tenen en la vida de les persones. Perquè quan algú s'acosta a un altre amb un problema, es tracta d'una cosa personal, una cosa que realment impacta en la seva vida diària.

Avui vull que pensem en el Salvador i us el presento. És aquí fent unes copes. Celebra que s'acaba de jubilar, és un home de seixanta-cinc anys. Ha estat taxista durant

tota la vida i pren unes copes per celebrar el començament de la seva nova vida. I està molt emocionat per això després de tants anys conduint un cotxe durant hores i hores cada dia. Després de gaudir d'una copa de vi amb els amics, torna a casa seva. De fet, viu en un habitatge social. Viu sol. No té família, la majoria de les seves connexions són amics locals, alguns són veïns seus. I decideix revisar el seu ordinador i veu un correu electrònic de fa sis setmanes que en realitat li diu que té fins a final de mes per actualitzar els serveis socials perquè pugui començar a rebre les seves noves prestacions de salut i de jubilació. I per a això té sis setmanes. I després recorda que també va rebre una carta que deia que el sistema ja estava funcionant, que tots els serveis socials s'estaven tornant digitals i també que havia de transferir-ho tot a una identificació digital. Una identificació digital, una cosa molt simple que està passant en tants països i que presenta tants reptes: com es mou tota la població d'un sistema analògic a un sistema digital, i a banda d'això entendre quines són les conseqüències d'aquesta introducció. Estic segura que teniu alguns d'aquests casos a les vostres taules de treball esperant una resolució.

Llavors, el Salvador truca a la línia d'ajuda i rep aquest servei on-line. Parlem de bots que algunes de les empreses han estat dissenyant, i que són molt útils a vegades, sobretot si s'entén què passa i que realment no s'està parlant amb un humà, sinó que s'està parlant amb un robot. Però, ja sabeu, passa pel sistema, prem els botons clau, intenta arribar a una cara humana i comprèn que en realitat tots els ajuts que tenia s'han eliminat perquè no va fer tota aquesta nova tasca administrativa a temps. Llavors, ja ha estat assenyalat. *Assenyalat*, què significa això? Només significa que ha estat marcat en el sistema per no haver complert les noves regles i les regles digitals.

Llavors, què li passa al Salvador? Quan se'ns va ocórrer aquesta història, vam pensar inicialment que aquest era un cas de 2030, això és molt hipotètic. Però després comences a llegir les notícies, comences a veure el que està succeint a Europa amb molts dels sistemes de decisió automàtica i comprens que en realitat no és hipotètic. I el que va passar amb el Salvador és que de sobte va perdre les seves prestacions de salut i el seu fons de jubilació. Sí que es va queixar i sí que va interposar recurs, perquè va percebre que l'estaven tractant injustament, però això va trigar mesos i mesos a resoldre's, perquè com es deté aquesta màquina? Com aconseguir que el sistema compregui que en realitat l'està tractant de manera injusta? Llavors, gasta els estalvis, perd l'habitatge, perquè no té cèdula digital, i tampoc no pot accedir als serveis de salut i té una malaltia cardíaca crònica, per la qual cosa acaba vivint al carrer. És com travessar per una espiral de males decisions, passant per alt correus electrònics o cartes. I com es pot culpar una persona que això hagi passat? Ens pot passar a qualsevol. Com a conseqüència, el Salvador acaba dormint al carrer i se sent víctima de la burocràcia i de l'automatització.

I no vull dir que la intel·ligència artificial no pugui presentar beneficis per a tothom, i estic segura que hi són, però també vull crear consciència sobre com la presa de decisions automatitzada pot incrementar i augmentar la desigualtat en els serveis públics digitals. Per tant, abans d'aprofundir en això, només vull dir: Realment volem un futur digital en què subcontractem les decisions difícils a les màquines? I em faig aquesta pregunta, perquè quan es tracta dels serveis públics i de salut digital, aquestes són les raons més comunes que trobareu, i això apareix una vegada i una altra en els informes sobre la transformació digital dels serveis públics: reducció de costos, reducció de la quantitat de persones que es necessiten per tramitar aquests casos, detecció i reducció del frau.

Llavors, bàsicament, comencem a tractar el problema des de la perspectiva equivocada. Posem la mirada en els col·lectius vulnerables, parlem de les persones que més necessiten aquests serveis, i assumim que l'ús que fan d'aquests serveis no és el correcte, mentre ignorem l'impacte real que pot tenir una mala decisió en la vida d'aquestes persones, perquè moltes vegades aquests serveis públics són la xarxa de

seguretat d'aquests individus, i quan aquesta xarxa de seguretat no hi és, és molt difícil sortir d'aquesta situació.

En el sector digital sempre parlem de la bretxa digital entre diferents sectors i col·lectius de la societat. Però m'agradaria parlar-vos sobre la bretxa de coneixement, perquè crec que és aquí on realment tenim el repte, per a tots els qui treballem en empreses o treballem en el sector privat o per als ciutadans. Com sentim aquesta bretxa de coneixement a l'hora d'entendre el que fan aquests sistemes? Com ho fem quan som assenyalats? I la majoria de nosaltres en alguna etapa de les nostres vides serem assenyalats per un d'aquests sistemes. Pot ser un banc, pot ser un servei públic. Ja ho sabeu, hi ha moltes formes i molts moments en les nostres vides en què constantment som avaluats i avaluats per riscos. Llavors, la meua pregunta és: Quan som assenyalats, què succeeix? I com evitem aquestes situacions?

Per tant, com avui som on som, vull entrar amb més detall en tres mostres de sistemes automatitzats de presa de decisions que s'estan implementant a Europa. Perquè crec que és aquí quan comencem a comprendre realment els reptes que tenim sobre la taula i a comprendre el que realment significa aquesta bretxa de coneixement. I el repte que tenim, per exemple, com a servei públic i com a proveïdor d'un servei públic. El primer exemple que aporto avui és Gladsaxe. Aquest és un sistema que s'està implementant a Dinamarca. Tot va començar molt bé, amb molt bones intencions. Parlem d'un sistema en què tres petites ciutats fora de Copenhaguen van demanar dur a terme un projecte amb què van crear un sistema d'avaluació de riscos, un sistema basat en punts que anticipa el risc a què s'enfronten els infants. La idea és que els serveis socials puguin identificar els infants abans que realment tinguin unes necessitats que requereixin haver de ser atesos pels serveis socials. Llavors, com funciona aquest sistema basat en punts? Bàsicament és, dit d'una manera molt cruel, com un sistema de la vergonya. S'obtenen punts pel fet de no comportar-se com la societat pensa que t'has de comportar. Per exemple, et divorcies i obtens mil punts. Si perds una cita amb el pediatre per al teu fill ja tens uns altres mil punts. Tens un problema de salut mental? Tens tres mil punts.

Per tant, parlem d'un sistema que intenta identificar els infants que en un futur pròxim podrien necessitar l'ajuda dels serveis socials a Dinamarca. Llavors, veiem que, en realitat, es poden seguir de prop aquests infants. Però el que certament també fem és etiquetar i controlar la vida d'aquestes famílies; famílies que de moltes maneres ja es troben en una situació difícil. Penseu en la mare soltera que té cura dels seus fills, però que alhora també ha d'estar en una o dues feines i que la penalitzen constantment pel fet de no encaixar en aquest sistema basat en punts. I estic segura que quan va començar tot es va fer amb bones intencions. Però és només quan comencem a observar l'impacte que té això en la vida de les persones, quan realment comprem com les bones intencions es poden tornar en contra, que veiem que poden anar en realitat malament molt ràpidament. Aquest exemple està relacionat amb la salut, i és interessant.

RiskEr és un sistema que s'està desenvolupant a Itàlia i, bàsicament, és un procediment estadístic que combina variables demogràfiques i de salut per comprendre i predir el risc d'hospitalització de determinades persones. És un sistema molt bo, molt interessant. Es tracta de fer que el vostre sistema de salut sigui més eficient. S'està provant a vint-i-cinc centres de salut a Itàlia. El repte comença quan llegim la lletra petita. Es recopilen cinc-cents elements, variables, sobre el context social i de salut d'aquestes persones. Llavors, parlem de cinc-cents variables. Com a servidor públic, això ja és molt difícil de controlar. Com es pot parlar de biaix, sent justos, quan s'estan combinant cinc-cents indicadors?

Però el repte aquí és que en realitat aquest sistema es va desenvolupar en col·laboració amb un hospital dels EUA, d'on fa prou feines unes setmanes va arribar la notícia que el sistema intencionalment no té biaixos en termes de raça. Per tant, no hi havia

indicadors de raça en aquest algoritme. Però el que feia el sistema als EUA era avaluar el tipus d'atenció que hauria de tenir cada persona en funció de les seves condicions cròniques i, mitjançant una investigació a la universitat, es va acabar veient que, en realitat, s'havia indicat que les persones negres necessitaven menys atenció, menys serveis dels hospitals.

Però, ja sabeu, el sistema en realitat va ser dissenyat per no tenir aquest biaix. Llavors, com va succeir això? Perquè històricament les persones negres gasten menys, van menys al metge o gasten menys diners en metges, o també ha passat perquè es tracta del sistema estatunidenc i tenen una abast de l'assegurança més baix. Llavors, històricament, utilitzen els serveis de salut pública en una taxa més baixa que les persones blanques. Per tant, novament parlem d'un sistema que va ser dissenyat per no incloure dades de raça, indicadors de raça, però per les condicions socioeconòmiques s'estava tractant injustament aquest col·lectiu. Aquests són els reptes que presenten aquests sistemes. Com adonar-nos-en, com assenyalar que aquests sistemes han de tractar aquests col·lectius de manera justa abans que realment ja estiguin en acció? Perquè el nivell de complexitat és molt, molt alt.

L'últim cas que vull presentar, i que estic segura que la majoria de vosaltres coneixereu, és el sistema que acaba de ser frenat pel Tribunal dels Països Baixos, i que és un sistema d'indicació de risc. Es tracta d'un sistema de serveis de benestar que ha estat assenyalat per la cort holandesa com una violació dels drets humans. I per què van arribar a aquesta conclusió? Perquè aquest sistema s'estava implementant principalment en veïnats de baixos ingressos i recopilava dades a tot el govern. Era un sistema per detectar el frau. La possibilitat que un ciutadà cometí frau en termes de dret laboral. Parlem d'un tema molt específic, probablement un percentatge molt baix de la població realment comet frau en termes de treball. Però després el que ens trobem aquí és que s'estaven recopilant dades i les dades recopilades pel govern de sobte es van ajuntar, es van barrejar, i com a resultat hi havia un col·lectiu específic de la població sota vigilància contínua. Per tant, aquest és un gran repte, comprendre el que passa, perquè no es tracta només del que fan les dades. Sovint, comences a utilitzar conjunts de dades que no van ser dissenyats o no van ser creats per a aquest tema específic. En conseqüència, és molt difícil assegurar realment la qualitat d'aquestes dades.

Llavors, crec que també és molt important tenir en compte la periodificació, i avui estic molt contenta que s'hagi organitzat aquest esdeveniment a Barcelona per diferents motius. Fa un temps que parlem a Barcelona de tecnologia i tecnologia centrada en l'ésser humà, però també hi ha diferents moments en l'àmbit internacional que crec que situen el tema de la intel·ligència artificial i els drets humans al centre de l'agenda, i realment espero que tots nosaltres, quan tornem a les nostres institucions i ciutats, realment puguem posar aquest tema a l'agenda de tothom.

Aquest any també és un any d'oportunitats perquè, el mes d'octubre, Philip Alston, que també va visitar Espanya molt recentment, va publicar un informe exactament sobre aquest tema en el sector públic, i com n'és, de rendible, i per detectar el frau en els serveis socials es va veure que castigaven constantment les mateixes persones, que eren les que més necessitaven aquests serveis. És com castigar dues vegades aquests col·lectius. No només estan en situació de pobresa, sinó que els tornem a castigar perquè pensem que necessiten estar sota control. La setmana passada vam tenir accés a la publicació de l'Estratègia digital de la UE, que estableix moltes de les agendes en relació amb el que seran els propers quatre o cinc anys. I durant aquests dies la tecnologia ha de funcionar per a la gent. La meua pregunta és: Què significa això en el nostre treball diari? Com posem la tecnologia al servei de les persones i ens assurem que el que desenvolupem no és innovació per innovació, sinó tecnologia al servei de les persones, però també assegurant que no sempre sigui en benefici dels mateixos? És el nostre deure assegurar-nos que quan parlem de tecnologia al servei de les persones, en realitat sigui per a tothom independentment del seu context, de la seva condició social o condició migratòria.

Us explicaré una mica més sobre la nostra tasca a la Digital Future Society, perquè aquest any celebrem els vint-i-cinc anys de la Declaració de Pequín. Aquesta declaració parla sobre els drets de les dones i, igual que en la Declaració de drets humans, aquesta agenda no parla realment de tecnologia. Però el que ens trobem en la nostra investigació és que aquests mateixos sistemes de què parlem avui també discriminen de manera exponencial les dones. Perquè històricament la condició socioeconòmica de la dona sempre ha estat en una posició més feble. Tenim coneixement sobre això, per exemple, pel que fa a les pensions, pel que fa a l'accés a la tecnologia, al coneixement, a l'educació. Llavors, realment hem d'entendre quines són les tecnologies que desenvolupem i quins són l'objectiu i els afectats per aquest tipus de disseny de tecnologies.

Llavors, el que vull de vosaltres, i que m'agradaria que us emportéssiu a casa, a més de les xerrades d'avui, és que realment penséssiu en intel·ligència artificial més enllà de la tecnologia i de la recopilació de dades. Com assegurem la qualitat d'aquestes dades? Com ens assegurem de comprendre aquests biaixos quan no són tan visibles? Només els entenem una vegada s'estan implementant. Però també penso en com afecta això certs col·lectius: dones, infants, migrants, sempre acaben sent les mateixes persones. Potser és el moment d'implicar-los en aquests debats, d'escoltar la seva experiència, de com és sentir-se continuament alertats pels mateixos sistemes que haurien de donar-los suport.

Hi ha diferents coses que podem fer des de les nostres institucions. Crec que una és la regulació, i és molt fàcil parlar de regulació, de vegades és fins i tot més difícil implementar-la. També la fiscalització, la rellevància dels òrgans de fiscalització, i també tenint un procés específic d'auditoria. Com a ciutadana vull saber on puc anar quan sento que un sistema digital m'està tractant injustament i que algú té el poder d'evitar ràpidament que aquest sistema faci això.

Vam estar parlant que als Països Baixos els activistes socials han estat treballant en aquests temes durant gairebé una dècada. Quantes vides ha destruït el sistema que avaluava aquest poble durant una dècada? Quantes vides han de passar per això per saber el bàsic d'una xarxa de seguretat abans que puguem realment prendre mesures? Aquesta és la pregunta que sempre em faig: Quant ens costa repensar realment com dissenyem aquests sistemes i també les dades històriques? I un repte, crec que el repte més gros que tenim entre mans és com desenvolupar la capacitat pública dins el sector públic, perquè la bretxa de coneixements és molt important i àmplia. Però també parlem de tot el sector públic, parlem de contractació. Com es fa la contractació per garantir que aquestes situacions no es produeixin i una vegada que aquests sistemes entren en funcionament, quines són les eines que té el sector públic per aturar aquests sistemes? Quins són els aspectes legals que realment donen seguretat als ciutadans quan els governs canvien sobtadament i les dades que proporcionem per a un ús específic s'estan reutilitzant per a un altre objectiu? I, finalment, com capacitem els ciutadans perquè compreguin realment què està passant i com els sistemes esdevenen traçables?

També cal recordar que, per a cada transformació digital, per a cada transició digital, hauríem de tenir un anàleg, una cara humana, perquè al final del dia, i està molt bé dir que la intel·ligència artificial està dissenyada per humans i que inclou prejudicis humans, però això no és suficient. Necessitem assegurar-nos de tenir sempre una cara humana a l'altre costat i algú amb poder per aturar una decisió injusta.

Només per posar-vos una mica en context, això és part de la feina que hem estat fent a Digital Future Society, i el nostre objectiu és realment establir l'agenda per als propers cinc anys per als diferents actors dels ecosistemes: autoritats públiques i més enllà, la indústria, etc. I l'any passat vam començar a publicar un informe, que crec que està exposat a l'entrada, amb alguns exemples de sistemes automatitzats de presa de decisions en el sector públic i els reptes que les tecnologies emergents ofereixen al

sector públic i també les oportunitats. I aquest any, de fet, farem més feina sobre això, però observant l'impacte que aquestes tecnologies també tenen en les qüestions de gènere. Llavors, us convido a saber més coses sobre la nostra tasca. Ens podeu trobar a digitalfuturesociety.com. També sou benvinguts a venir i parlar amb nosaltres. També tenim Carlos Grau, el president de la nostra fundació, aquí avui amb nosaltres. Així que espero haver-vos donat alguna cosa en què pensar i emportar-vos a casa. Gràcies!

Intel·ligència artificial i drets humans

reptes, rols i eines de les institucions d'ombudsman

Dilluns, 2 de març de 2020

De la societat analògica a la digital

Moderador: Genís Roca, expert en Internet

José María Lassalle, director del Fòrum d'Humanisme Tecnològic d'ESADE i professor de Filosofia del Dret d'ICADE

Daniel Innerarity, catedràtic de Filosofia Política, investigador Ikerbasque en la Universitat del País Basc i professor de l'Institut Europeu de Florència

Renata Ávila, directora executiva de la Fundació Ciudadano Inteligente

David Cabo, director de la Fundació Civio

Genís Roca

Expert en internet

És un plaer coordinar aquesta taula, és un luxe. Els quatre ponents són castellanoparlants. Tenim Bilbao, Santander, Guatemala i Saragossa, encara que la vida us ha portat a viure qui sap on.

És un panell interessant. Classificar i etiquetar la gent és un vici molt lleig, però podríem dir que hi ha dos filòsofs i dos activistes, encara que insisteixo que les etiquetes aquestes són una mica injustes.

Veníem de la sessió anterior, hem pogut veure acadèmia, món d'empresa i món de la societat discutint sobre què és intel·ligència artificial (IA). Podríem convenir que sí, que estem d'acord que la IA està molt bé, que salvarà el món, i que necessita ètica i que tot està molt bé, però la veritat és que sembla difícil que això pugui avançar pel bon camí, si no fixem algunes normes d'algun tipus. Qui ens convoca avui és un defensor del poble que el que demana és un marc legal per poder fixar les regles. Estic segur que moltes de les empreses i dels actors que intervindran en tot això agrairien que hi hagués regles perquè quedés clar el que està dins i el que està fora de la llei i no tenir aquests debats ètics. Que al final el debat ètic és l'absència de norma, però també, i Ulises crec que ha estat força clar, et fies que el Govern pugui garantir que es compleixin les normes, no?

Estem davant d'un tipus de tecnologia en què els mecanismes de control, de com s'està fent tot això, són força difícils. David ens podrà explicar com costa posar sota control el que passa darrere d'un algoritme. Les regles que estan prenent decisions que poden afectar la nostra vida no són clares, i és molt menys clar com supervisar i auditar aquestes regles, amb la qual cosa per molt que estiguem desitjant un marc legal, podríem tenir el problema que estem posant en joc llibertats. I, encara que hi hagi una norma, la dificultat de complir una norma és patent. Això ens podria portar a mecanismes en què si introduïm autocontrol, si introduïm vigilància ciutadana, si introduïm el marc normatiu (que sí, fem les tres coses, introduïm marc normatiu, introduïm autocontrol i introduïm vigilància ciutadana), però on estan els rols? I fins a quin punt això és redefinir la democràcia o les regles del joc de la democràcia? I fins a quin punt els marcs mentals que teníem fins ara de què és norma i de què és compliment de la norma seran suficients per viure en aquest món?

Algú va dir una vegada que la intel·ligència artificial en aquests moments és molt artificial i poc intel·ligent, però sí que no hi ha cap dubte que progressa adequadament, i que el que avui és embrionari tindrà una força i unes conseqüències contundents en el nostre dia a dia.

Volem discutir amb vosaltres quatre aquestes coses. Teniu un temps inicial per exposar el vostre plantejament. Cadascun de vosaltres porta una motxilla i el seu punt de vista,

i per això sou aquí, perquè ens interessa saber com es complementa tot això i, en la mesura que sigui possible, veure com us creueu, si hi ha una mica de debat i si el públic té o ens regala alguna pregunta o fa alguna consideració al que pugueu dir ara.

José María Lassalle

Director del Fòrum d'Humanisme Tecnològic d'ESADE i professor de Filosofia del Dret d'ICADE

Bé, jo també em reivindico com una persona que està en l'activisme, i crec que des de la política, encara que jo ja l'he abandonat, continua sent possible l'activisme.

Jo en el temps que vaig ser secretari d'estat d'Agenda Digital, i lamento citar-me, vam posar en marxa la comissió de drets digitals, que ha donat peu, així que va arribar, al que ha estat la primera declaració continguda en la Llei de protecció de dades de l'any 2019 i el compromís, almenys en l'exposició de motius, d'afrontar una constitució digital per al nostre país.

La comissió, a més, la van formar experts en drets digitals, però també activistes i gent del sector privat, i la va coordinar un exministre socialista, Tomás de la Quadra, amb la qual cosa garantim, crec, una relativa independència i autonomia en la gestió.

Després vam posar en marxa una cosa que ara ja està en l'imaginari dels debats públics, que és la comissió de la bretxa de gènere digital. Quan la vam posar en marxa el gener de 2018, ens miraven molts com dient ¿de què parlem? I la veritat és que és una de les coses de les quals em sento més orgullós, perquè va ser difícil tirar-ho endavant, i el govern que va venir després ho va assumir i ha continuat la tasca.

Vam voler posar en marxa la comissió d'IA, però malauradament no va poder ser pel canvi de govern i ara s'ha reprès amb una secretaria d'estat d'agenda, vaja, d'IA, i vam posar en marxa l'embrió del que després ha estat la Digital Forum Society, amb la reflexió que plantejem sobre el Davos Digital i l'aportació pressupostària que va permetre arrancar aquesta meravellosa iniciativa, que ha generat el que per a mi era una dada fonamental, que era, al costat dels drets, al costat de la recerca d'equitat, que afrontéssim també un espai de reflexió crítica que fes viable el que jo crec que seria el debat més profund que hauríem de plantejar en aquests moments.

Que és que no parlem d'ètica, parlem de regulació i parlem de regulació per una senzilla raó, perquè quan es va produir l'altra gran revolució tecnològica, que va ser la revolució industrial, es va alliberar una potencialitat de canvi que va transformar radicalment les societats europees. Va enterrar les societats de l'antic règim, va modificar el model econòmic, va canviar les mentalitats, però va generar durant molts anys un dèficit democràtic, unes situacions d'injustícia i desigualtat extraordinàries que la civilització democràtica i les societats europees van trigar més de 100 anys a corregir.

És a dir, no podem entendre l'estat social i democràtic de dret com l'entenem avui, si no és el producte de veure que és la conseqüència de tot un intent de pal·liar el dèficit d'equitat que va generar la revolució industrial. Una revolució que va girar al voltant de dos vectors fonamentals: el vector de la velocitat i el vector de la capacitat, que van transformar radicalment el món i van canviar les mentalitats i van donar peu a la revolució de 1848, a la revolució soviètica, és a dir, a tot un seguit de conscienciacions socials, de lluites que han generat un *feed-back* que no podem llençar ara per la borda, quan estem afrontant un altre moment tan important o molt més important que el que va suposar la revolució industrial, que va ser un canvi extraordinari, perquè era un canvi associat al progrés, per tant associat a un fet, i és que l'home, com deia Ortega, és un ésser amb un dèficit epistemològic innat que necessita generar, a través de fonts d'informació, instruments amb què cobrir la seva imperfecció i així sorgeix el poder de la tècnica des de l'antiguitat fins avui. Des de l'*homo habilis* fins avui.

I això és la generació d'una sèrie de capes de coneixement d'institucionalitat que ens han portat fins aquí, i a les quals no podem renunciar perquè la tècnica és un instrument de progrés, és un instrument de canvi i és un instrument d'avenç a tots els nivells, però el que ens ha succeït des de probablement la primera dècada de segle

XX a avui és que quan s'ha produït un esclat que ha estat disruptiu de l'acumulació en la capacitat de generació de velocitat i de continguts que podien circular a través les xarxes que han permès amb la revolució tecnològica més plàstica, que ha estat la introducció de la telefonia mòbil, ha generat una eclosió de desigualtat, una situació de manca de regulació democràtica, de dèficit crític que la democràcia assumís realment una reflexió sobre el que estava passant que quan han transcorregut deu anys en aquesta acceleració del temps que hem viscut, hem descobert que el nostre món ha canviat.

L'ésser humà experimenta una revolució antològica que està alterant les claus sobre la seva manera de viure's a si mateix, és a dir, l'ésser humà per primera vegada està fent que els cossos es retirin de la realitat, que la nostra identitat comença a veure com el cos, que ha estat el substrat sobre el qual s'ha construït la nostra identitat, comença a replegar-se en la seva capacitat de gestionar la informació a través de la qual construïm la nostra pròpia vida, quan articulem una nova experiència d'identitat, a través de la identitat digital, quan ens vivim amb els altres a través de la telefonia mòbil, quan vivim atrapats en pantalles, ens adonem que ens hem perdut pel camí, que ens hem perdut ja no només l'ètica, ens estem perdent l'ésser humà i ens estem perdent les dimensions jurídiques que han protegit l'ésser humà tal com l'hem entès fins ara.

Però no en tenim prou només de trasplantar la mentalitat jurídica de la modernitat a la nostra postmodernitat tecnològica, ja no en tenim prou amb això, en tenim prou d'imaginar jurídicament com ens gestionarem el segle XXI sota les noves experiències a què ens haurem d'enfrontar i que estan vinculades a una nova sensibilitat, a una nova manera de relacionar-nos, a una nova experiència antropològica de l'ésser humà que necessita ser pensada jurídicament. No n'hi ha prou només a trasplantar els drets fonamentals, concebuts en la societat analògica, a l'àmbit digital.

El dret té per davant l'esforç d'imaginar jurídicament una nova generació de drets fonamentals, és a dir, de la mateixa manera que hi va haver una primera generació de drets vinculada a la llibertat, una segona generació de drets vinculada a l'experiència democràtica, una tercera generació de drets vinculada a com es pal·liaven les marginalitats o les externalitats negatives que generava la societat democràtica i el mercat, hem d'afrontar una nova generació vinculada a les externalitats negatives que està generant la revolució digital.

I per això no podem pensar en clau analògica, ja no ens serveix la distinció entre privacitat i publicitat. És que hi ha la intimitat, és que no només som usuaris i consumidors de continguts, és que tenim un dèficit com a persones quan treballem i operem com a tals en les xarxes, i ja no dic res quan som ciutadans. Com a ciutadans hem perdut la nostra identitat, perquè és que no hi ha un marc regulador. Hi ha un marc regulador que defineixi com confeccionar un algoritme? Hi ha un marc regulador que ens digui de qui és la propietat de les dades? Quins drets tenim a l'hora de gestionar les nostres dades? Per primera vegada el capitalisme afronta una teoria del mercat al voltant de l'economia de plataformes i no té una teoria de la propietat.

Les *commodities* de la nova economia de plataformes, que són les dades, no tenen propietaris. Sí, tenen un propietari: aquell que introdueix l'algoritme i llavors obté les externalitats que hi afegeixen valor i li permeten monetitzar el seu model de negoci.

Aquest és el nou disseny que estem construint, i en aquest disseny no hi ha hagut reflexió democràtica. Aquest és el gran dèficit que tenim en aquests moments. I per això no només ens calen drets, ens fa modificar i fer evolucionar les institucions.

A mi m'obsessiona, sobretot, un tema, és a dir, quan construïm una societat democràtica, que és el producte de generacions, de revolucions i de canvis socials molt profunds, es va afrontar, perquè la modernitat va comprendre que volia drets perquè tenia una idea de progrés i, per tant, una acció cap a la qual orientar els

esforços humans de canvi. El canvi social es va identificar com una realitat possible a través de la llei i, per tant, es va pensar des d'una modernitat com orientar l'esforç de l'acció humana, individual i col·lectiva.

Ara necessitem fonamentalment, i d'això no se'n parla, una llibertat per comprendre, és a dir, necessitem no llibertat per actuar, que també, necessitem llibertat per comprendre què està passant, és a dir: Quin és el sentit que volem donar a les màquines? Quin sentit volem donar a tot el que ens passa?

Hi havia un sentit clar de la tecnologia vinculada al progrés i al canvi social. Ara veiem que la tecnologia està alterant els eixos gravitacionals de l'experiència de l'humà, no solament en un terreny col·lectiu, sinó també en un terreny individual. I necessitem capacitat per pensar-ho.

La gran llibertat a hores d'ara és recuperar el temps per poder pensar, entre altres coses, que la nostra corporeïtat està en desús, per exemple. I hem construït tota la nostra emocionalitat, una part molt important de les nostres categories jurídiques, per exemple la lluita contra la tortura, a partir de l'experiència de saber el que era un cos que patia.

I això que ha estat acompanyant la construcció del dret i la teoria política fins avui s'està qüestionant, se'ns està anant de les mans. I no tenim temps per pensar-ho, perquè, entre altres coses, estem atrapats per una acceleració quotidiana on estem imbuïts, a més, per una obsessió de rendibilitzar al màxim els esforços mentals, com si estiguéssim traslladant el compte de resultats de les grans corporacions tecnològiques que busquen monetitzar exhaustivament totes les inversions que fan, també en ètica ara, però les porten al terreny del que és humà i al terreny de la nostra quotidianitat i no tenim temps per pensar.

Jo crec que aquest és el gran repte revolucionari que hauríem de ser capaços de reclamar a les societats democràtiques, temps per pensar el que ens està passant i per donar sentit a tot el que fem.

Daniel Innerarity

Catedràtic de Filosofia Política, investigador Ikerbasque en la Universitat del País Basc i professor a l'Institut Europeu de Florència

Soc un filòsof; em sap greu, ningú no és perfecte. I els filòsofs, quan ho fem bé, que no és sempre, el que aconseguim és mostrar una especial habilitat per detectar problemes, fins i tot allà on la gent no els veu, fins i tot allà on es diria que no n'hi ha. I això té a veure amb el fet que tenim una espècie d'addicció, com la gent que té addicció a les drogues o a l'alcohol, nosaltres tenim una addicció als problemes i també estem acostumats que ens preguntin què fa un noi com tu en un lloc com aquest? Sempre ens fan aquest tipus de preguntes i anem per la vida amb el nostre detector de problemes. Allà on hi ha una disrupció, un canvi, un desajust i tot es posa cap per avall, nosaltres ens freguem les mans perquè aquí hi ha material per al pensament.

Aquest moment en què ens trobem, que té una dimensió històrica, tràgica per a molta gent, és un moment que per a un filòsof que té alguna cosa de paràsit, és un moment fascinant.

Jo crec que a la intel·ligència artificial li falta democràcia, és a dir, ens falta la consideració que aquest canvi que hem de fer afectarà la nostra vida democràtica fins a uns extrems que potser no són tolerables, però que al mateix temps constitueix un instrument que ben pensat i ben articulat pot fer-nos grans serveis als nostres sistemes polítics.

Tinc la impressió que falta reflexió sobre aquesta dimensió política de la intel·ligència artificial i sobren codis ètics. L'any passat jo n'havia recollit ja 84. Ja no sé per on va la xifra.

Hem fet com a societat, i a Europa especialment, grans avenços en el que podíem anomenar la reflexió des del dret privat respecte d'aquests temes. Hem fet el GDPR, el sistema de protecció de la propietat i la privacitat, que és un gran avenç des de la perspectiva de l'habeas corpus, però hem fet molt poc de la perspectiva del bé comú.

La democràcia és decisió popular, és autogovern de la gent, però el primer i el més important és que es tracta d'una decisió. Què passarà quan cada vegada que hi hagi més assumptes automatitzats? El que realment em fascina des del punt de vista conceptual és com conciliem democràcia, (*demos*) i (*cratos*), govern del poble, decisió del poble, amb cada vegada més sistemes automatitzats.

El camp ideològic, diguem-ne, es divideix referent a això en dos grans tipus d'anàlisi; quan no en tenim ni idea, el camp es divideix entre els optimistes i els pessimistes. Els optimistes consideren que l'automatització és fantàstica perquè la vella política ideologitzada que teníem fins ara serà substituïda per sistemes que decidiran amb moltíssima més precisió. Els pessimistes creuen veure-hi alguna cosa terrible, la distopia que deixarem de decidir, haurem de tancar els parlaments i uns sistemes intel·ligents decidiran per nosaltres. I llavors entra aquesta imatge que ho embruteix tot de robots que ens substitueixen. Optimistes i pessimistes en el fons pensen que la política és una cosa que es pot superar, que pot ser substituïda per una altra instància de l'estil que sigui, i es diferencien únicament en considerar-ho una bona o mala notícia.

El meu plantejament és que la política no es pot substituir absolutament per cap altra cosa, que fins i tot en les tecnologies que es presenten com a mancades d'ideologia hi ha sempre una política.

Què passa quan tenim una esfera pública automatitzada? Què li passa a la democràcia? Com podem pensar la seva compatibilitat amb la democràcia? Quin sentit pot tenir la política democràtica en un món en el qual cada vegada hi ha més coses automatitzades? Un món cada vegada més "deshumanitzat", no en el sentit ètic, sinó un món en què no hi ha humans en cada vegada més sistemes de decisió. Ara no hi ha ningú als ascensors ajudant-te a pujar, hi ha camps de conreu robotitzats, satèl·lits geoestacionaris, portes automatitzades, és a dir, una gran quantitat de processos automàtics, fins al punt que algú ha dit que podríem arribar a un món que funcionaria sense nosaltres. O com deia Minski, en què tindríem molta sort si les màquines ens adoptessin com a animal de companyia.

I al mateix temps, es diu, això és ple de caixes negres, la societat de les caixes negres, aquesta societat on hi ha dispositius de decisió la lògica dels quals no comprenem. Aquí la solució màgica ha estat la transparència, després ja una mica matisada amb la idea d'explicabilitat, com ens explicaven en la taula anterior. Això és molt important i al mateix temps és molt complicat perquè, en primer lloc, hi ha diversos tipus d'intransparència, d'opacitat. Hi ha una opacitat que té a veure simplement amb el fet que per protegir la propietat per secrets d'estat o patents o pel que sigui, cal establir un sistema per fer opaques certes coses. N'hi ha una altra que té a veure amb la complexitat de la màquina. Encara que ens expliquin un algoritme, no ho entendrem. Però hi ha una altra tercera opacitat, que és encara molt més complexa, que és l'opacitat que ha de tenir amb els aparells de l'època de la postprogramació en què estem, època en la qual les màquines obeeixen a una lògica que no està en continuïtat ni tan sols amb la voluntat de qui l'ha dissenyat.

Com serà la democràcia en aquesta època en la qual per primera vegada en la història de la humanitat som capaços de dissenyar màquines que no entenem?

Perquè això és el que significa la intel·ligència artificial. Com traduïm això en termes d'explicabilitat? Hi ha un problema amb l'explicabilitat de què podem parlar, per descomptat, però crec que l'article 22 del GDPR és molt poc realista (que no podem estar condicionats per processos automàtics que no ens resultin d'alguna manera comprensibles), i a més té una visió privada de l'explicabilitat, és a dir, pensa que l'explicabilitat és una cosa que es pot fer individualment, que es pot gestionar com el que signa una clàusula d'acceptació. L'exemple més eloqüent és quan abans de la crisi econòmica tu demanaves un crèdit, anaves al banc i firmaves un acord que no llegia ningú, que no llegíem ningú. Després Europa va endurir les condicions perquè el banc hagués de garantir que s'havia entès. Quan posem aquesta exigència d'intel·ligibilitat individual apliquem esquemes d'un món simple que els subjectes podien comprendre a un món complex que només pot ser comprès col·lectivament.

No es tracta d'obligar la gent a llegir finalment les condicions dels crèdits, que això seria la solució privada, sinó de crear les condicions perquè la gent assumeixi crèdits de manera responsable i, per tant, no hi hagi una incitació excessiva a assumir riscos excessius, que hi hagi unes regulacions anticíclics, és a dir, l'important és el context, el marc regulador.

I acabo amb un parell de coses sobre quin futur pot tenir la democràcia en aquest context.

Primer, crec que estem davant d'un risc que cal pensar bé, que és el que podem anomenar el risc del populisme antidemocràtic. El que està passant a la Xina, no això del virus, sinó el fet que cada vegada sigui més temptador el recurs a màquines exactes davant de la política ideologitzada.

La IA té molts avantatges, per exemple, al voltant de la possibilitat de fer unes polítiques públiques molt més ajustades a la realitat i després, per descomptat, el mesurament de l'impacte d'aquestes polítiques públiques en la realitat. Es podria fer una política algorítmicament *tunejada*, si se'm permet l'expressió, aprofitant precisament aquesta granularitat dels mecanismes de la IA.

Ara bé, quin és el problema d'això? Per què els algoritmes i les dades tenen biaixos? Perquè hi ha gent malvada o perquè els han dissenyat malament? No, perquè la vida com a tal té biaixos i si la IA està pensada sobre la hipòtesi que el nostre comportament futur serà una continuació del nostre comportament passat, no hi ha manera de donar curs polític a experiències d'innovació, canvi o ruptura. Com trenques amb unes dades que són masculines, perquè el nostre comportament i estructura social, que els algoritmes reproduïen, també ho són?

Renata Ávila

Directora executiva de la Fundació Ciudadano Inteligente

Bé, una cosa que no s'ha esmentat en la bio, i crec que és rellevant, és que també soc part del Global Retrospective Board de Digital Future Society, de la qual Carina també forma part, i crec que és important esmentar-ho perquè per a mi és una de les meves activitats que més m'agrada, poder contribuir en diferents directors per filtrar idees que s'estan veient des de les bases i que s'estan veient des de la gent, i també poder definir noves formes de pensament que crec que és no només tasca dels filòsofs a les facultats, sinó que part de l'activisme que practico també és ser un canal que connecti els seus centres de pensament amb la realitat del carrer. I a això em volia referir avui, la meua xerrada està dedicada, i que aquí en veureu més, però jo sé que aquestes idees es recolliran i es transmetran precisament a aquests ombudsmen, més que al públic en general, perquè el que vull deixar aquí és un missatge importantíssim en què tan creatius i creatives han de ser presents en aquest nou rol, perquè tenim un problema.

El plantejament que porto aquí és com podem dissenyar un futur diferent i com podem dissenyar o codiseñar un futur diferent i quin serà el paper de les defensories en això?

Com heu vist en el vídeo, els problemes que veiem que està portant la tecnologia són problemes que veiem en la societat cada dia –discriminació, manca d'inclusió, manca d'oportunitats i d'estar equipades amb les eines que necessitem–, i el problema que tenim aquí també és que els efectes de la tecnologia, sempre se'ns diu que necessitem marcs ètics, necessitem noves regulacions, necessitem una regulació específica per a tecnologia, i no és així.

Jo crec que ja des d'avui, des del moment en què aquestes tecnologies comencen a discriminar, a afectar dignitat humana, a produir efectes en la societat, aquí entra el paper de les defensories, perquè tenim un marc universal de drets bastant bo, força sòlid i bastant avançat que ens arriba per cobrir aquests problemes. A més, tenim lleis específiques que cobreixen molts d'aquests problemes que veiem.

El problema que tenim és de visió. No som prou visionaris amb les regles que ja tenim, a què em refereixo ..., si jo llegeixo els mandats de les defensories de la majoria dels països, és bastant ampli i sé, perquè no entenc què passa en el procés, que els ombudsmen són molt tímids a entrar en els temes digitals. Ho veuen com una cosa tan allunyada, tan sofisticada, tan complicada, que no s'han atrevit a exercir de ple la seva autoritat en aquest segle i en aquest temps.

Creo que la figura de la defensoria de les dones i dels infants, per exemple, i dels grups especialment protegits, si es limiten i s'esperen que s'aprovi un marc regulador d'IA, per exemple, o s'esperen que tinguem un tractat internacional en aquests temes, arribarem molt tard, perquè ara estem en el risc real de cristal·litzar el patriarcat 2.0, perpetuar i entrenar els nostres sistemes amb models d'una societat que no és justa, que és la societat que tenim avui. És clar que no és una societat justa, però podríem revertir aquests biaixos, podríem dur a terme accions preventives i accions proactives per corregir aquests biaixos i no traspasar-los a allò digital.

El problema és que no hem corregit els biaixos de la nostra societat, i la modernització, la transformació digital que estem projectant, i que la podem llegir en totes les comunicacions de la UE en tots els plans de negoci de la companyies, és delegar a les màquines, delegar als algorismes amb dades de pobra qualitat, amb dades que reflecteixen la societat que tenim avui i no necessàriament la societat que volem demà.

I aquí hi ha uns exemples molt simples, però aquests exemples molt simples imagineu-los perpetuats en sistemes d'educació arreu de la UE, per exemple. Cada vegada més deleguem funcions amb l'educació dels infants a sistemes automatitzats que ja estan esbiaixats.

Tenim un problema gros perquè tampoc no corregim a qui està treballant en aquests sistemes i a qui els elabora. Imagineu-vos que si només el 0,8 de les persones de color amb habilitats tècniques a Google són dones, on es veuran reflectits els problemes com ara la identificació facial i altres? És a dir, són molts problemes, un sobre un altre, però quina resposta esperem nosaltres de les defensories?

Qui regula aquestes màquines? Perquè hi ha molts espais que regulen les màquines, però hi ha un espai específic i un mandat molt preciós de les defensories d'acord amb el qual els ombudsmen són els nostres advocats, són els advocats de la gent, són els que han d'anar i advocar per evitar que aquests problemes s'estenguin d'una manera desmesurada. Han de tenir un paper central en el futur i el meu plantejament, que porto avui, és un humil plantejament: que siguin les nostres *superwomen* i *supermen*, que siguin els defensors aquests pioners del futur qui posin les bases per a aquest futur digital que esperem, i pot ser que mai s'hagin assegut davant un ordinador i que la

seva major interacció amb un sistema sofisticat hagi estat parlar-li a la Siri, a l'Alexa, a la cuina, però cal prendre el mandat, llegir-nos el mandat i veure tot el poder que tenen a les seves mans i que tan limitadament exerceixen. És una de les coses que he après, una de les lliçons més importants de Julian Assange.

Tenim molt d'espai, tenim moltes possibilitats i sempre ens quedem en la línia. Crec que la tasca avui, les defensories en l'era digital, és interpretar d'una manera àmplia i actual aquests mandats que tenen i poden fer tres coses si no saben com funcionen les tecnologies. És una cosa que han fet els relators especials de les Nacions Unides. Cal fer les preguntes incòmodes a les companyies que estan desenvolupant aquestes tecnologies i cal recolzar-se en l'ecosistema d'universitats, de sector públic i altres per entendre-les millor.

Un primer pas podria ser, per exemple, que les defensories d'Europa es coordinin i truquin a les portes de les companyies per tractar un tema, per exemple, discriminació racial, i facin aquestes preguntes incòmodes a les companyies, preguntant-los què estan fent per mitigar-la, què estan fent per corregir-la en la nova generació de tecnologies que tenim.

La segona tasca que jo els deixaria a les defensories seria la de vigilar que les bones intencions de servir, i de servir ràpid, i l'eficiència, que s'investiguin intervencions tecnològiques que afectin especialment els més vulnerables i, ho connecto amb la xerrada que va fer Carina, que no es transformin en estructures d'opressió per als més pobres.

Crec que això va en el centre del mandat de les defensories, i que el tema sigui tecnològic no hauria de ser excusa per allunyar-se'n, per distanciar-se'n. Crec que al contrari, crec que després d'aquesta jornada, i després de tot el que hem sentit avui, s'hauria d'incloure en una declaració i llista de tasques immediata de les defensories fer un inventari de tots aquests sistemes que hi ha a Europa i veure si compleixen els estàndards mínims i si aspiren a elevar i millorar els estàndards al màxim de drets humans.

El tercer punt seria obrir una finestreta real per als ciutadans perquè puguin exposar com i quin tipus d'efectes negatius o positius tenen aquestes tecnologies en les seves vides. Crec que no hi ha gaires instàncies. Moltes vegades l'únic recurs que té la ciutadania són les unitats d'atenció a consumidor, i no és suficient, perquè moltes vegades parlem de violació de drets fonamentals, moltes vegades parlem de coses que no són quantificables i de relacions o transaccions que no són de mercat, que són asocials.

Llavors aquestes serien com tres qüestions: el paper dinàmic de consulta i aclariment, vigilància i verificació, però també experimentació.

Com les mateixes defensories poden prendre un paper actiu, per exemple, a l'hora de cocrear o participar i opinar de codiseñar quan, per exemple, una mancomunitat o un estat vol desenvolupar aquests sistemes.

I, bé, la crida a l'acció, perquè a mi, a més d'aquesta feliç tasca, m'agradaria activar els ombudsmen, és el primer assumir el rol que tenen, tenen una responsabilitat i tenen un deure de rendir-nos comptes a la ciutadania d'aquesta tasca.

El segon jo crec que un paper fonamental que poden tenir les defensories és el d'inclusió. Moltes vegades en el seu mandat tot el tema de no-discriminació, d'equitat de gènere, possibilitats per a totes les persones entra en el seu àmbit. Aquí podria entrar un altre paquet d'obligacions relacionades amb allò digital i amb la IA, en què podrien tenir un paper en també al costat positiu de la tecnologia i identificar com, ben implementades amb un marc de drets, es poden tornar mecanismes del futur que incrementin i que facin efectius els drets econòmics, socials i culturals.

I el tercer, cooperació. Crec que la coordinació podria ser una font riquíssima d'experiència i de millora d'aquests sistemes, però també sistemes de *blacklist* perquè si hi ha un sistema dolent, per exemple, el que es va declarar a Holanda, si aprenc com es van protegir els drets, tot i que bastant tard en aquests sistemes, puc evitar traspasar-lo a altres parts d'Europa.

I, bé, hem de fer incidència estratègica i coordinació global per poder fer un futur millor. Formació de capacitats en espais clau, defensories (és un espai clau) i pilots tecnològics amb la intervenció de les defensories i amb la intervenció molt més àmplia d'experts en temes de drets humans.

David Cabo

Director de la Fundació Civio

Moltes gràcies per la introducció tan fantàstica. Sí, soc el codirector de la Fundació Ciutadana Civio. La nostra missió és, bàsicament, augmentar la transparència i la rendició de comptes de les administracions públiques espanyoles des de fa vuit anys. Fem sol·licituds d'informació, periodisme de dades i, si fa falta, litiguem. Llitem contra el Govern i de vegades guanyem, com fa poc al Tribunal Suprem un cas contra el Tribunal de Comptes.

El primer que vull dir és que soc enginyer informàtic i soc plenament conscient del potencial i de totes les històries positives que s'expliquen sobre la tecnologia. Jo vinc del costat tecnològic i quan vaig cofundar Civio ho vaig fer perquè creia que la tecnologia podia servir per augmentar la transparència en les administracions. I continuo creient que hi ha eines, com ara anticorrupció, que poden millorar el públic, que és molt important, perquè el públic és de tots nosaltres, podem fer que el públic funcioni millor. Però, a la vegada, també vaig passar una mica de tenir una visió una mica naïf de tot això, una visió molt utòpica que la informació ens faria lliures, a tenir una visió una mica més crítica i adonar-me'n fins a quin punt no només la tecnologia no és neutra, com de vegades es diu, sinó que és utilitzada de forma intencionada per aquells que tenen el poder, que són els que tenen els recursos per crear la tecnologia, dissenyar la tecnologia i aplicar-la d'una o una altra manera.

És molt important que la societat civil o els ombudsmen estiguin enfront, una mica, de l'Administració per lluitar contra els excessos del poder, com ara els intents de discriminar o de mantenir la discriminació que hi pugui haver a la societat.

Llavors, des del punt de vista de la intel·ligència artificial i de la tecnologia, hi ha un concepte que sempre s'utilitza que és el del *digital welfare state*, estat del benestar digital, que, per exemple, el relator de les Nacions Unides de pobresa extrema, que s'ha esmentat en una xerrada anterior, ja el menciona. Quan sentim la paraula *digital* podem pensar en serveis més accessibles, més pròxims al ciutadà, que tot sigui més fàcil, que l'Administració estigui al servei del ciutadà, però moltes vegades el discurs preponderant, com deia Carina en una xerrada anterior, és una mica lluitar contra el frau de qui demana una subvenció, no?

Aquest discurs de l'eficiència i del control de qui demana una ajuda per a l'atur fa que, com denunciava el relator de les Nacions Unides, hi hagi zones cada vegada més creixents dins l'Administració que són opaques, que no estan regides pels drets humans, que és una mica el que ha passat amb el cas d'Holanda, on el Govern holandès crea aquest megasistema de control, i crea moltíssimes fonts de dades, i l'aplica només per a una ajuda en concret, que és la que utilitza la gent pobre o determinades barris pobres.

Llavors, quan es parla del biaix de l'algoritme de vegades no només és el biaix de l'algoritme en si, sinó el biaix a l'hora d'aplicar aquesta tecnologia. El fet que tu creïs aquesta eina de control per a una cosa que afecta els ciutadans pobres i que el despleguis en ciutats o barris pobres.

Dins aquest canvi que s'està produint de la transformació a l'estat del benestar digital, a nosaltres ens preocupen moltes coses, però diguem-ne tres per ser molt breu.

Un és el tema de la bretxa digital: hi ha gent que no té accés a Internet o, si en té, s'enfronta a processos extremadament complexos, com explicaré després amb el tema del bo social. Complexos tant des del punt de vista conceptual, de notificacions, múltiples passos, formularis com també tecnològics. Si heu intentat utilitzar certificats digitals, us heu enfrontat amb les versions de Java, els navegadors... Nosaltres fa anys que ens queixem que és extremadament difícil fer una sol·licitud d'informació per barreres tecnològiques que hi ha, i tot això són coses que es van sumant, i que jo soc capaç d'esquivar perquè m'he passat anys lluitant amb aquests sistemes, però hi ha moltíssima gent, i és especialment la gent vulnerable, la gent que necessita aquestes ajudes, que no pot enfrontar-se a aquestes barreres.

Una altra cosa que ens preocupa és el rentat de cara del biaix, és a dir, la sistematització del biaix. El cas típic que se sol posar és el de la policia predictiva, però aplicats a uns altres casos que hem vist aquí. És a dir, tu entrenes un sistema amb dades històriques que mostren un biaix, un racisme, per exemple, de la policia, que es pot adreçar especialment contra gent pobre o gent negra, tu entrenes un sistema i utilitzes aquesta sortida per dir: Això és neutral perquè no ho dic jo, ho diu la tecnologia. Hi ha una frase que a mi m'agrada molt, que és "machine learning is like money laundering for bias", que l'aprenentatge automàtic és com el blanqueig de diners per al biaix, perquè et permet dir que ja no ets tu el que ets racista, que hi ha un sistema objectiu, i els números no menteixen, que diu que hi ha més crim en aquest barri.

Això fa que s'elimini el responsable últim de les decisions polítiques, que és un tema fonamental per a la qualitat democràtica. És a dir, veiem en molts països que en implantar-se un sistema, per exemple, que regula ajudes públiques, a la pràctica, el que s'està fent és reduir aquestes ajudes. Però no hi ha una decisió política oberta que pugui ser discutida o que pugui ser debatuda, sinó que el que es fa és dir: No, aquesta és una solució tècnica que gestionarà les ajudes. Però, a la pràctica, té una sèrie de criteris que no es discuteixen que el que fan és que a molta gent se li rebutgin aquestes ajudes.

I, finalment, una altra cosa que ens preocupa és que s'estiguin creant noves àrees d'opacitat, es creen noves excuses. Una d'aquestes excuses és que això és molt complicat, i que aquesta és una discussió que pertany als tècnics o que pertany als enginyers, i aquí és una mica el cas concret en què nosaltres treballem, que explicaré molt breument.

Una de les coses que fem, o que fa Eva Belmonte, periodista i codirectora de Civio, és llegir-se el BOE cada matí i veure si hi ha notícies rellevants per als ciutadans, perquè parteix que la nostra missió és explicar les decisions de les administracions. El 2017 vam veure que es canviava el sistema del bo social, el sistema que dona ajudes a gent vulnerable que no pot pagar la factura de l'electricitat. Comencem a denunciar que s'estaven quedant milions de persones fora, que no demanaven el nou bo i que es quedarien sense les ajudes quan caduqués el sistema anterior. Ens vam posar en contacte amb la Comissió Nacional dels Mercats i de la Competència (CNMC) per crear una calculadora en línia per facilitar tot el procés. És un procés extremadament complex. Perquè us feu una idea de la complexitat, els advocats i els responsables de la CNMC, gent que són els reguladors del sector elèctric, tenien moltíssims dubtes sobre quin era el procés a seguir i vam estar setmanes lluitant tots i fent preguntes al Ministeri, per entendre exactament els casos que no eren clars, on si eres família nombrosa i a la vegada eres víctima del terrorisme, etc. És un sistema que no poden entendre ni els advocats de la CNMC.

Arran d'això i de treure la calculadora, la van utilitzar centenars de milers de persones. Ens van arribar centenars de trucades de gent que tenia problemes, hi havia una desinformació brutal. Hi havia gent que també es queixava que la nostra calculadora deia que sí, però que després l'elèctrica que gestionava l'ajuda havia dit que no, que no sabia què fer per defensar-se, hi havia una indefensió enorme.

I, arran de tot això, decidim fer una sol·licitud d'informació per intentar monitorar el sistema o auditar el sistema externament, que creiem que és una cosa fonamental. I, molt breument, demanem l'especificació tècnica de l'aplicació, els casos d'ús i el codi font. El Ministeri va dir que no. Apel·lem, vam anar al Consell de Transparència, aconseguim l'especificació tècnica i els casos d'ús, i demostrem que hi havia errors en l'aplicació. El Ministeri ho va reconèixer i va dir que ho arreglaria, la qual cosa és part de la nostra missió. Ens alegra que s'hagi arreglat això, però continuem lluitant i hem anat a judici i estem en judici ara mateix per aconseguir el codi font. Perquè, igual que hi ha hagut errors en la transposició en l'especificació tècnica, pot haver-hi errors a l'hora de fer el codi.

I aquest és un cas que tècnicament és molt senzill, és a dir, no parlem de *machine learning*, no parlem de xarxes neuronals, és un algoritme que ve d'una llei i, tot i així, és extremadament complex d'auditar. Tenim moltíssima por de com seria això amb un sistema més complex. Ara estem en la primera instància. El Ministeri, per un costat, ha dit que és propietat intel·lectual, que no ens vol donar el codi font. Responem que l'ha desenvolupat el Govern, per la qual cosa ha de ser públic. I, a més, diuen que hi ha un tema de seguretat, que si ens expliquen com funciona això i ens donen el codi, qualsevol persona podrà connectar-se a l'Agència Tributària i treure les dades del seu veí i fer-li xantatge, literalment. Lluitem per aconseguir que els mecanismes de control que tenim ara, encara que siguin molt ineficients, es continuïn aplicant a les noves tecnologies i als nous sistemes.

I, per tancar, bàsicament, jo de vegades utilitzo l'analogia que la IA és una mica com el Mag d'Oz. Si heu vist la pel·lícula, sabreu que tot el món anava a Oz i esperava que li resolgués tots els problemes, que li donés un cor, que li donés un cervell... I després es descobreix, bé, això és una mica *spoiler*, que el Mag d'Oz en realitat era un senyor que estava darrere de la cortina. La IA és una mica això. La nostra missió és obrir aquesta cortina, i el que descobrirem és que darrere de tot això no hi ha un sistema màgic, sinó que hi ha humans o persones que han desenvolupat aquests sistemes; persones que, a més, com deia Renata, són blancs i són homes, amb la qual cosa això introdueix una sèrie de biaixos.

I el meu objectiu és intentar lluitar contra el fatalisme i la sensació aquesta de fetitxització de la tecnologia i la idea que la tecnologia és fantàstica i que arribarà i que es desplegarà per tot arreu i que ja està. Hem de vigilar cadascun dels sistemes tecnològics que es despleguen. Esperem que els ombudsman assumeixin aquest rol i que lluitin per vigilar la implantació de cadascun d'aquests sistemes, i necessitem que tots els sistemes que regulen drets públics puguin explicar les seves decisions, puguin ser auditats externament i tinguin un responsable últim, una mica en la línia del que deia Carina, que hi hagi una persona, que hi hagi una cara que et pugui ajudar, que et pugui defensar en cas que el sistema digui que no. No podem estar a la mercè de la tecnologia, encara que de vegades sembli una mica el progrés i que sigui tot fantàstic.

I aquest seria una mica el meu missatge. Gràcies.

Intel·ligència artificial i drets humans

reptes, rols i eines de les institucions d'ombudsman

Dimarts, 3 de març de 2020

Drets, reptes: Aproximació sectorial

Moderadora: Esther Paniagua, periodista independent especialitzada en tecnologia

Ángel Gómez de Ágreda, coronel d'Aviació, responsable de l'Àrea d'Anàlisi Geopolítica, DICOES

Josep M. Argimon, director gerent de l'Institut Català de la Salut

Liliana Arroyo, experta en transformació digital; Institut d'Innovació Social (ESADE)

Itziar de Lecuona, doctora en Dret i Professora Agregada, Facultat de Medicina i Ciències de la Salut; subdirectora de l'Observatori de Bioètica i Dret-Càtedra UNESCO de Bioètica de la Universitat de Barcelona

JM Sánchez Bursón, consultor

Esther Paniagua

Periodista independent especialitzada en tecnologia

Moltes gràcies per ser aquí de nou. No vull deixar d'agrair, encara que ja es digués ahir, però mai és sobrer, a Rafael Ribó i a tot l'equip del Defensor del Poble de Barcelona que hagi organitzat aquestes jornades sobre un tema tan important, reconeixent que no som experts, però que necessitem saber com abordar això perquè és fonamental, ho està sent ja, i serà crucial per als drets dels ciutadans.

Ahir es van donar exemples, inicialment en la primera taula que va haver-hi, de com la intel·ligència artificial pot ajudar els drets humans a tancar bretxes, etc. Però al final el que competeix al Defensor del Poble, el que és realment important, és veure el contrari, com pot fer que no es vulnerin aquests drets, no? Els riscos d'aquest desequilibri entre Administració i ciutadans que pot produir-se arran de l'ús d'aquestes tecnologies i de com s'està fent en diversos àmbits. D'això també se'n va parlar una mica ahir.

Es va parlar de discriminació. La Carina va parlar de com es pot automatitzar la desigualtat, com es pot automatitzar la discriminació. També la Renata va donar alguns exemples d'això. I d'aquest tema en parlarem més avui, aterrant en diferents àrees i veient com afecta això directament els ciutadans. Primer per tenir una panoràmica general de com està afectant això la gent i després també per intentar veure com abordar-ho, com es concreta això en solucions o què pot fer el Defensor del Poble.

La Renata també va fer una crida a l'acció que em sembla, independentment de com es faci, necessària. Jo no em vull allargar gaire perquè ahir vam veure que va haver-hi poc temps per al debat i m'agradaria que hi hagués temps per al debat amb la taula, però sobretot cal que tingueu l'oportunitat de preguntar el que us inquieta, d'exposar els dubtes que tingueu, perquè al final som aquí per a això, som aquí per a vosaltres. Us animo, doncs, que no us talleu i que escriviu, si us plau, aquestes preguntes. Us demanaria que fossin o en castellà o en anglès, li fareu un favor a la moderadora. I bé, després d'unes breus exposicions, farem algunes preguntes per a la taula i passarem a les preguntes de l'audiència.

No vull allargar-me més, parlarem d'intel·ligència artificial i de ciberseguretat, i de com la ciberseguretat i les polítiques relatives a la ciberseguretat poden afectar, poden tenir un impacte directe en la vida de les persones. I per a això tenim aquí Ángel Gómez de Ágreda, que és coronel d'aviació, responsable de l'àrea d'anàlisi geopolítica de la Secretaria de Política de Defensa del Ministeri de Defensa i, a més, autor d'un llibre excel·lent que no se'ns explicarà, perquè no hem vingut aquí a parlar del teu llibre, Ángel. *Mundo Orwell*, es diu el llibre, oi?

Parlarem també de l'impacte de l'ús de la intel·ligència artificial en investigació, tractament de dades, bases de dades i investigació en salut, concretament en una àrea molt, molt sensible, amb Josep Maria Argimón, director gerent a l'Institut Català de la salut. Parlarem, seguirem parlant, perquè ja vam començar a donar-nos pinzellades ahir de temes d'impacte social, discriminació de gènere, etc., i per això tenim aquí la Liliana Arroyo, que és experta en transformació digital de l'Institut d'Innovació Social d'ESADE, que a més ha fet diversos informes sobre aquests temes molt interessants i ens en podrà parlar.

Parlarem de bioètica amb Itziar de Lecuona, que és subdirectora de l'Observatori de Bioètica i Dret, de la Càtedra de la UNESCO de la Universitat de Barcelona.

I parlarem finalment, però no perquè sigui menys important, perquè de fet és fonamental això, de l'impacte de l'ús de la intel·ligència artificial en la infància i l'adolescència. Per a això tenim aquí el consultor José María Sánchez Bursón. Benvinguts tots i començarem per ordre, que així facilitarem la vida als tècnics amb les presentacions.

Ángel, què pots explicar-nos sobre el tema de la ciberseguretat?

Ángel Gómez de Ágreda

Coronel d'Aviació, responsable de l'Àrea d'Anàlisi Geopolítica, DICOES

Bé, voldria començar per agrair al síndic de greuges, el senyor Ribó, i a l'IIO la invitació a participar en el seminari i l'acollida aquí a Barcelona. Moltes gràcies.

El primer dir que soc aquí a títol particular, més com a autor del llibre i com a acadèmic que com a coronel d'aviació, però evidentment el paper de coronel d'aviació –la meua formació és com a pilot i com a paracaigudista–, doncs et dona una visió més específica respecte al tema de seguretat.

Tinc algunes bones notícies i sobretot alguns reptes pel que fa a la ciberseguretat i a la seguretat en general en el tractament de les dades de la intel·ligència artificial.

Som a Barcelona. Som en una de les ciutats més conscients, més innovadores pel que fa a intel·ligència artificial, però hem de pensar que vivim en el que alguns anomenen un *món físic*: un món físic per un costat i digital per l'altre. Ahir bàsicament fem el tractament del món digital de la intel·ligència artificial, de la informàtica en general com una eina. Jo fa temps que defenso que es tracta d'un ecosistema. Vivim en un món físic i vivim en un món lògic de forma simultània; és a dir, no es tracta d'una eina, es tracta d'un lloc o d'un no-lloc en el qual vivim, i això té connotacions que són realment importants.

Les dades, per exemple. Ahir parlàvem de la importància de preservar les dades perquè són nostres. Les dades no són nostres, les dades són nosaltres. No és el mateix. Estem fets de carn i os, però també estem fets d'uns i zeros en el món digital, i el nostre jo digital ara mateix és gairebé tan important com el nostre jo físic. I no anem regalant nassos, orelles, braços per aquí a la vora, doncs tampoc no hauríem d'anar regalant les nostres dades, els nostres uns i els nostres zeros. No és que les dades siguin nostres, no es tracta d'una propietat, no es tracta d'una eina, es tracta d'una identitat que estem defensant, i aquest és un concepte que em sembla important.

Vivim en les dues realitats. El 94% dels espanyols utilitzem Google habitualment, tot el sistema de Google. Vivim a Barcelona, vivim a Catalunya, vivim a Espanya i vivim a Google. Tinc una amiga, Pilar Trucios, que diu sempre: On amagaries un mort a Internet?...En la segona pàgina de Google perquè quan cerquem una mica ens quedem normalment amb el que Google ens diu en la primera pàgina. La primera pàgina és diferent per a cadascun de nosaltres, i en funció de si estem connectats amb el nostre usuari o estem en el nostre ordinador.

Després l'univers que veiem depèn de Google, depèn de l'algoritme de Google i del que sap Google sobre nosaltres. El problema per als d'una generació que ja hem passat dels 50 és que continuem pensant en Internet moltes vegades com la Internet dels anys 90, una internet en què ningú sabia que eres un gos, perquè doncs no existien encara les plataformes, que ara mateix és el vehicle que utilitzem per connectar-nos a Internet, i, per tant, accedíem directament als continguts, més o menys directament als continguts, quan ara hi ha una acumulació de dades sobre nosaltres que permet determinar no només que som un gos, sinó a més la raça de gos amb qui vivim, on som, d'on venim i on anem. Les tres preguntes fonamentals, no?: Qui som, d'on venim i on anem.

Qui som ens ho diu Google en la primera pàgina de les cerques, d'on venim i on anem ens ho diu Google Maps cada vegada que estem geolocalitzats, que és sempre absolutament. Tendeixo a dir en les meves conferències que quan deshabilitem la funció de localització del nostre telèfon, li estem dient al telèfon que nosaltres no volem saber on som, no que el telèfon deixi de saber on és.

Hi ha un experiment molt interessant que podeu fer. Escriviu a Google simplement "codes sensor" i el model del vostre telèfon, i us donarà un número que truca en el qual surt la pàgina en què hi ha tots els sensors que té actius el telèfon en un moment donat i veureu setanta i tants/vuitanta paràmetres diferents que el telèfon sap sobre tu en cada moment:

ubicació, acceleració, en els tres eixos, pressió atmosfèrica, temperatura, longitud i latitud. Un munt de paràmetres que nosaltres no coneixem, però que el nostre telèfon sí que coneix de nosaltres. Després les dades que s'estan acumulant aquí són tremendament importants.

Aquesta imatge em suggereix dues coses i crec que les dues són importants. D'alguna manera vivim en un món mitjançat. Fins fa uns anys teníem una experiència de primera mà. Coneixíem el nostre barri, coneixíem si de cas la nostra ciutat, un entorn més o menys pròxim i l'experiència que teníem de les coses era molt directa. L'experiència que tenim ara el 90% dels casos està mitjançada per una pantalla, una pantalla a través de la qual les nostres dades són transparents per a tot el món. Bé, no per a tot el món, sinó per als que hi ha darrere, per a les companyies que hi ha darrere, que utilitzen aquestes dades per construir la realitat que després ens tomen en forma d'una visió concreta del món alineat amb els seus interessos, moltes vegades comercials. Què suposa això? Que perdem llibertat i perdem seguretat. Perdem seguretat en el moment que donem les nostres dades i llibertat en el moment que no veiem la realitat tal com és, sinó que construïm la nostra veritat en funció d'unes percepcions que ens arriben a través de la pantalla.

L'altra cosa que em suggereix la imatge és respecte a la intel·ligència artificial. La intel·ligència artificial el que fa és tornar-nos la nostra imatge magnificada i optimitzada en funció dels paràmetres, dels algorismes, i això és el que ens espanta, que ens fa veure'ns tal com som sense cap tipus de filtre i portats potser fins a l'extrem.

Al final, tot el tema de seguretat, tot el tema de l'autonomia, de les màquines o dels algorismes, amb el que té a veure és amb la dignitat humana, i únicament amb la dignitat humana. És amb el sentiment de superioritat que té l'ésser humà respecte a la resta de les criatures o a la resta de les màquines, o a la resta del nostre entorn, i que ens fa veure'ns com uns éssers únics, i ens espanta la forma en què ens veuen les màquines, la forma en què ens veiem nosaltres.

No entraré en detalls dels diferents tipus de ciberatacs que es poden fer, entre altres coses, perquè és irrellevant. Al final, gairebé mai hi ha darrere el que ens creiem que hi ha. Ara mateix els ajuntaments tenen un paper fonamental en la seguretat i en la protecció de les nostres dades. De fet, veiem que els últims atacs de *ransomware*, dels que t'encripten l'ordinador o les dades o els arxius i et deixen sense poder-hi accedir, s'estan acarnissant sobretot amb els ajuntaments, amb els hospitals i amb les asseguradores; és a dir, precisament amb les institucions, amb les entitats que tenen moltíssimes dades sobre nosaltres, i que les tenen ordenades de manera que són perfectament utilitzables ràpidament.

Algú de veritat es creu que atacaran el teu ordinador per cobrar-te 300 euros de rescat? O és més lucratiu, mentre tu estàs preocupat de recuperar el teu ordinador, robar-te les dades i després poder vendre-les per un costat i utilitzar-les per un altre?

Des de 2007 l'atac a Estònia, de denegació de servei, que va deixar Estònia desconnectada durant dues setmanes i mitja, un atac que provenia de Rússia, però que no hi ha forma de demostrar que fos un atac institucional. Des de 2007 som conscients de la nostra dependència d'Internet.

Ahir, comentàvem què passaria si s'acabés Internet, si algú eliminés Internet de sobte.

Bàsicament, la conclusió m'imagino que serà que ja no estem preparats per viure sense Internet i que tornarem gairebé al neolític. Estem exposats cada dia, en cada moment. La seguretat de les nostres dades està exposada. Bé, va sortir no fa gaire, devíeu veure la notícia de la gent que va pel metro o pels aeroports amb un lector de targetes de crèdit i que van llegint les dades de les teves targetes. Bé, m'imagino que a hores d'ara tots porteu les vostres targetes posades dins una gàbia de Faraday perquè no es pugui llegir la targeta mentre la teniu a la butxaca fins que la traieu, no?

Al que em porta això una mica és el món que hem dissenyat, el món digital que hem dissenyat tots, estem connectats amb tots. Després, l'estat, les institucions d'alguna manera poden proporcionar-nos un cert nivell de seguretat, però el que és clar és que la connexió directa que tenim cadascun amb tots els altres implica que tots retenim una responsabilitat sobre la nostra pròpia seguretat, que no podem delegar i que no ens pot proporcionar cap entitat, cap institució externa.

Els atacs de *ransomware* que esmentava fa un moment estan creixent de forma exponencial, com ho fa tot normalment en el ciberespai, i estan creixent precisament perquè hi ha gent que s'ha adonat del valor de les nostres dades, i sobretot perquè hi ha molta gent que no ens hem adonat del valor de les nostres dades.

Clar, un capítol fonamental serà probablement la protecció de les infraestructures crítiques, la necessitat que tenim d'assegurar l'accés a les nostres infraestructures crítiques. Hi ha un parell de pel·lícules molt reveladores en aquest sentit que expliquen que és el que passaria en cas d'un atac.

Fa un parell d'anys va haver-hi un atac, precisament a Ucraïna, que va inutilitzar la infraestructura elèctrica durant unes hores un 23 de desembre, que a Ucraïna és una data molt poc convenient per quedar-te sense electricitat i sense gas a casa, fins i tot va arribar a provocar en un moment donat algunes morts per congelació als habitatges sense electricitat. Les nostres finances ara mateix és un altre dels camps que cal protegir, però urgentment. Pràcticament gairebé tots els fons estan utilitzant el *fast-trading* i els algorismes financers per guanyar en eficiència a l'hora de fer les transaccions borsàries i aquests sistemes són tremendament fàcils de *hackejar*.

De fet, fa uns anys l'exèrcit electrònic sirí simplement *hackejant* el compte de Twitter d'Europa Press, una agència de notícies, va induir els algorismes a pensar que hi havia hagut un atemptat a la Casa Blanca i, per tant, va haver-hi un crac financer, un petit crac financer que va durar cinc minuts. Dos minuts va trigar a baixar la borsa de Nova York 142 punts i més dos minuts va trigar a tornar pujar, una vegada que van recuperar el control sobre el compte. Per al grup aquest que s'havia posat a curt inicialment i havia capitalitzat en la baixada de la borsa i després havia tornat a capitalitzar a l'hora de pujar va suposar una operació financera molt lucrativa i molt fàcil de fer. Això mateix ho havia fet Ossama Bin Laden abans de l'11-S, apostant que la borsa de Nova York, si els avions no es desviaven dues illes de cases, baixaria. Si es desviaven més de dues illes de cases desapareixeria directament perquè estava just al costat de les torres bessones. I, efectivament, sembla que va estar finançant Al Qaeda durant un temps amb l'operació financera que va fer d'especulació a la borsa de Nova York.

Un altre dels camps en què s'està utilitzant molt, en què la seguretat s'ha d'enfortir per la permeabilitat i per la ubiqüitat de les nostres dades, és el tema del blanqueig de diners. S'ha creat d'alguna manera una sèrie de paradisos fiscals en l'àmbit mundial i una de les característiques fonamentals dels atacs en el ciberespai és que tenen a veure també amb la intel·ligència artificial, és que ja no existeix la simetria que hi havia abans d'empreses que ataquen empreses, particulars que ataquen particulars, estats que ataquen estats. No, ara mateix veiem un nombre creixent cada any d'estats que ataquen empreses.

Corea del Nord es finança en molt bona part robant bancs estrangers. Evidentment, per a una empresa, el fet que t'ataqui un estat, amb tot el poder que té darrere, és particularment greu i rellevant, amb la qual cosa hem de posar molta atenció amb aquest tipus d'atacs, sobretot perquè són atacs que moltes vegades no venen signats per l'Estat, sinó que es deixen una mica a l'aire.

Per anar acabant, bé a Espanya estem força bé en qüestió de ciberseguretat. Aquesta és una classificació que va canviant amb molta freqüència, entre altres coses, perquè això és com la borsa, que rendibilitats passades no garanteixen rendibilitats futures. Avui estàs molt bé, t'adorms sobre els llozers durant dos dies i passes a estar en el lloc 200 sense

adonar-te'n en un moment. Vaig agafar-ne una que ens afavoria, que ens anava molt bé i ens donava un cert optimisme, que falta que ens fa.

Després d'aquestes jornades segurament necessitarem un toc d'optimisme, i estem bé, entre altres coses, perquè estem fent els deures des de l'any 2013, que es crea l'Estratègia nacional de ciberseguretat. Es crea també el Comandament Conjunt de Ciberdefensa, on vaig estar destinat tres anys, i una sèrie de comitès que permeten des de la Moncloa la coordinació de tots els actors, tant en l'àmbit d'empreses, privat, com en l'àmbit estatal i entre els diferents nivells de l'Administració i en l'àmbit de defensa. Creant vincles, creant formes d'intercanviar informació a cadascun dels nivells estratègic, operacional, tàctic. I em sembla molt important el nivell tècnic també perquè es pugui intercanviar la informació del dia a dia accionable, que se sol dir.

Això també té les seves aplicacions positives. Podem generar una economia sostenible, podem generar (i en això no hi ha gaire a innovar sobre el que ja està fent Barcelona), podem generar iniciatives que permetin estimular una millora en la conducta urbana dels ciutadans, podem crear duplicats de les nostres ciutats per fer els experiments amb gasosa, en comptes de fer-los directament sobre la ciutat real; fer-ho sobre no els duplicats que hi ha, encara que puguem haver-ho fet generant o utilitzant les dades que hem tret del ciberespai.

Al final hi ha dos aspectes, tornant al principi, que es poden *hackejar*. Es pot *hackejar* la xarxa, es pot *hackejar* la part tecnològica i es poden *hackejar* també les persones, i veiem que l'ús, precisament, de la intel·ligència dels bots a mi em preocupa molt.

Faig la tesi doctoral sobre l'ètica de la intel·ligència artificial, i em sembla a mi que qualsevol intel·ligència artificial que sigui capaç d'alterar la teva voluntat, la teva llibertat, és tan perillosa com aquella que pot amenaçar la teva vida, perquè, en un moment donat, aquesta és la finalitat de la guerra. Al final és canviar el teu punt de vista, canviar la teva llibertat i, per tant, és tan perillós *hackejar* la màquina com *hackejar* l'usuari final, *hackejar* les dades de l'usuari final. Les plataformes ens han introduït en les cambres d'eco que tendeixen a donar-nos la raó, que tendeixen a polaritzar també les nostres opinions, que tendeixen a retenir la nostra atenció a còpia de donar-nos la raó, i aquí és on entren en joc el que anomeno en el meu llibre els *egoritmes*. Els algorismes ens agrupen i després els *egoritmes*, és a dir, el nostre desig de ser rellevants dins el grup, acaben polaritzant aquests grups, acaben portant-nos cap als extrems i deixant-nos gairebé sense cap marge de maniobra per parlar, per al diàleg, per a la interacció, que és tan important. Acaben gairebé deixant-nos amb llenguatges diferents, amb la impossibilitat de comunicar-nos amb aquells que no pensen exactament igual que nosaltres.

Respecte a les falses notícies, que no notícies falses –en realitat són relats amb format de notícia–, l'important no és si és veritat o és mentida, és la seva capacitat per manipular i per dirigir el relat. L'important no és qui mata el mamut, l'important és el que el compta, perquè vint mil anys després, ningú es recorda de qui es va menjar el mamut, del que es recorden és de la pintura que va deixar.

Això ho utilitzen els terroristes. Aquesta és la portada de la revista Rumiya després dels atemptats que va haver-hi aquí a Barcelona: 146 morts en una portada gòtica, en una façana gòtica. No va haver-hi 146 morts. Hi posa morts i ferits.

Algú havia llegit "morts i ferits"? Oi que no? L'important és el 146 en vermell, que destaca una xifra elevada que doni la sensació d'una gran victòria per a ells, d'una gran humiliació, d'una gran derrota i d'una gran pèrdua per a nosaltres.

És el control del relat el que realment importa. Deia que hem perdut seguretat i hem perdut llibertat, i ho hem fet a canvi de guanyar comoditat. Ara mateix comprem qualsevol cosa que sigui còmoda, que sigui immediata i que sigui aparentment gratuïta.

Hem de revertir aquesta tendència, tenim, i es comentava ahir, la necessitat que els mateixos usuaris tinguem l'educació i posem l'esforç suficient per construir la nostra llibertat. La llibertat no és gratis. El diu el memorial de la guerra de Corea a Washington "Freedom is not free". La llibertat i la seguretat s'aconsegueixen amb l'esforç de cadascun de nosaltres. Si pretenem viure sense esforç, viurem normalment sense llibertat. Hem d'afegir aquesta faceta responsabilitat per evitar convertir-nos en éssers transparents, i una cosa molt important també, l'última:

Som a Barcelona. Recordeu que Barcelona, totes les ciutats al final no són més que una part del que veiem, també hi és, en el ciberespai. Que som en aquest món físic, però que som també en aquest món digital simultàniament. En aquesta Barcelona i en aquesta Barcelona que ens estan creant les xarxes socials. Moltíssimes gràcies.

Josep M. Argimon

Director gerent de l'Institut Català de la Salut

Molt ràpidament. En els sistemes sanitaris de la majoria de països en el món desenvolupat jo crec que hi ha dues característiques que es comparteixen.

La primera és que la inversió que es fa en salut va tenint cada vegada menys retorn. No és com fa un segle, quan es podia estar invertint en vacunes, en mesures preventives, en mesures higièniques i que, per tant, l'esperança de vida o els indicadors de salut anaven creixent. Actualment, aquest retorn és menor, i encara que els nostres fills i adolescents seran la generació dels 100 anys, és a dir, que les persones que avui tenen 14, 15, 16 anys viuran 100 anys, tots sabem que això no és exclusivament, ni només principalment, pels serveis sanitaris que rebran. Ho és perquè hi ha tota una sèrie de factors econòmics, socials que tenen tant o més importància que el mateix sistema sanitari sobre la salut.

I, des d'aquesta perspectiva, des d'aquests factors socials i econòmics, cada vegada també, i és una altra característica dels sistemes sanitaris, hi ha més desigualtat, i podem pensar en més desigualtat en països com els Estats Units, on la cobertura sanitària no és universal, i és cert, però també als països europeus, on tenim sistemes sanitaris amb cobertura universal, la desigualtat augmenta.

La segona característica important és que amb la digitalització s'està generant i s'ha generat una quantitat massiva de dades i que va creixent exponencialment.

Només per tenir-ne una idea, i donaré dades molt, molt locals, d'un país petit com és Catalunya, de 7 milions i mig d'habitants, i que per tant no són res comparables amb les dades que ahir va proporcionar Ulises, l'investigador del BCC, que allò et tirava endarrere.

Per exemple, amb 7 milions i mig, hem de pensar que cada vegada que algú a Catalunya veu el seu infermer, el seu metge, o ingressa en un hospital, o va a la farmàcia, va amb una targeta identificadora, i que per tant tenim informació sobre 150 milions de prescripcions electròniques, sobre 50 milions de visites que s'han fet en l'atenció primària, i tenim informació estructurada sobre salut i diagnòstic, i altra molta que no ho està. O podem tenir tota la imatge radiològica, digitalitzada des de fa deu anys, i més o menys ben categoritzada.

Tot això, i amb aquest identificador únic, et dona un potencial clar per poder utilitzar en investigació, però també sempre pensant si aquesta investigació et proporcionarà o no aquests beneficis que dèiem abans.

Primer, reduir les desigualtats, sobretot reduir també els errors diagnòstics que tenim, perquè hi són; els errors de tractament, perquè els tenim, hi són, i des d'aquestes premisses ja la Unió Europea l'any 2002 treu un informe, Redesign health in Europe for 2020, en què s'estableix precisament la necessitat de reutilitzar la informació que disposem dels sistemes sanitaris, bàsicament, per a investigació, innovació i planificació sanitària.

I aquestes dades es poden utilitzar d'una manera més o menys tradicional, com hem estat fent en la investigació durant dècades. L'única cosa que, en comptes de recollir la informació dada a dada, o persona a persona, doncs reutilitzar les nostres dades, però cercant una arquitectura similar d'estudis. Fem els estudis de talls, etc., i aquí podem trobar molt ràpidament elements de millora. Per exemple, per a la farmacovigilància, veure els efectes secundaris que poden tenir medicaments molt utilitzats en el nostre àmbit, i que segurament en aquesta sala hi ha més d'una persona que ahir, avui o demà prendrà doncs un protector gàstric, l'omeprazol, i el seu risc, que el té, de fractura de fèmur. Això fer-ho amb una investigació tradicional és complicat, perquè significa passar molts anys, reunir informació prospectiva de moltes persones, reutilitzant les dades d'una manera segura, anònima. Després, si voleu, parlarem de l'anonimització, i sabent que mai no podràs aconseguir l'anonimització 100%. És a dir, el risc existeix i la capacitat de desanonimitzar existeix, hem de deixar-ho tot clar, però et permet aquesta rapidesa en el que és obtenir uns resultats.

També és veritat que es pot anar utilitzant tota aquesta informació de dades amb una perspectiva molt més d'intel·ligència artificial, i la intel·ligència artificial en els sistemes sanitaris fa anys que s'hi ha introduït. Ara podria pensar en els mateixos robots que ens ajuden, que ens assisteixen en les intervencions quirúrgiques o una cosa que és molt important, i que redueix dràsticament els riscos d'una equivocació en el tractament quan tu estàs ingressat en un hospital, que són tots els armaris robotitzats que et permeten dispensar medicació.

Això fa anys que s'hi ha introduït, però en els últims anys el que s'hi està introduint són sistemes d'ajuda a la decisió o de suport a la decisió. Per exemple, en l'àmbit de la radiologia o en l'àmbit de l'anatomia patològica o salut mental, cardiologia o en qualsevol dels àmbits.

De fet, dissabte em vaig entretenir a cercar al Pubmed. Pubmed és una base de dades que és la més utilitzada i la que té un repositori més ampli d'articles mèdics indexats en revistes, en què hi ha avaluació d'experts, i simplement cercant les paraules clau: *intel·ligència artificial*, *deep learning*, *machine learning*, etc., el títol i *abstract*, no més enllà, any 2010, any 2019, se n'ha multiplicat per 17 la publicació.

Tots aquests algorismes que ara s'estan comercialitzant i que estan oferint, entre cometes, a la majoria de centres i centres sanitaris són algorismes que poden realment ajudar i que ens ajudaran. Un exemple molt senzill pot ser algorismes que en aquests moments estem provant al mateix Hospital de la Vall d'Hebron, per poder diagnosticar o descartar 21 patologies en una radiografia simple de tòrax. Perquè en vegeu la utilitat, la radiografia de tòrax no s'informa, és a dir, quan un metge demana una radiografia simple, no s'informa, normalment té les capacitats, les habilitats per poder dir si hi ha alguna cosa o no, però en alguns moments hi ha dubte i, davant el dubte, evidentment es pot sol·licitar aquest informe, que trigarà uns dies.

Si nosaltres som capaços de posar aquesta eina a l'abast dels professionals d'atenció primària, haurem guanyat una mica, sempre que aquesta eina sigui una eina vàlida, sigui una eina no esbiaixada, i després em referiré als biaixos, i sigui una eina transparent.

L'article que també es comentava ahir en la primera sessió de la tarda d'aquest simposi. La idea és que el camp és molt prometedor, encara es troba molt en l'àmbit de la investigació, però ja els hospitals introdueixen això: algorismes d'ajuda a la decisió. Un altre àmbit que és vital en què la digitalització va més endarrerida, almenys aquí a Europa, i que ara comencem a digitalitzar, és tot el que fa referència a l'anatomia patològica, i encara que no us sembli, o us semblarà una dada que és una mica xocant, però la concordança avui dia en sistemes analògics d'anatomia patològica és baixa. Tenim fins a un 25%, 30% de discordances. Parlem de diagnòstics, en alguns casos doncs de tumors; per tant, la digitalització ens permetrà reduir aquesta discordança, augmentar la concordança i, al seu torn, associat, poder implementar aquests algorismes d'ajuda a la decisió.

Ara bé, no tot són flors i violes. Hem de tenir present que aquests algorismes primer s'obtenen a partir de dades de poblacions que a vegades no sabem molt bé d'on procedeixen, i això és una cosa molt diferent d'un assaig clínic. En un assaig clínic tu saps perfectament a qui inclouen i a qui no, i els mateixos assajos clínics moltes vegades són massa estrictes a l'hora d'incloure pacients. Són massa estrictes i, per tant, són d'alguna manera, entre cometes, criticables, perquè després, quan els utilitzarem en la pràctica clínica, aquell resultat espectacular que es produeix en l'assaig clínic publicat, queda reduït a una mínima expressió quan s'utilitza en un món real. Però sabem que en el món dels algorismes, això és més complicat de saber.

D'un assaig clínic, i crec que això és important també destacar-ho, hi ha tota una sèrie de normes molt protocol·litzades de com s'ha de fer, de com s'ha de publicar, de com s'ha de comunicar i que, si no s'ha fet així, una revista, o almenys revistes en què hi ha avaluació d'experts, no admet aquell assaig clínic, mentre que aquestes normes d'estandardització en els algorismes, aplicables després al suport en la decisió mèdica, no existeix encara, i aquesta és una primera tasca que cal fer.

Segon, quan em refereixo als biaixos, tenim exemples clars com algorismes per a l'anàlisi dermatològic –i, concretament, en melanoma– que no inclouen determinats grups de pacients de determinades races, primer element. Hi ha sempre potencialment un biaix de gènere. En els assajos clínics també hi és, però és clarament identificable. Per tant, aquests elements d'estandardització són clau, i després lògicament quan ens equivoquem individualment, el dany és petit. Si m'equivoco jo, puc fer mal a poca gent. Al final puc visitar 15, 20 persones. Si m'equivoco en un algorisme, el dany pot ser força superior. I després, quan hi ha bretxes, cosa que ara comentava Àngel, bé, si algú treu del meu despatx una història clínica és una bretxa, i molt punible, però és un cas. Si hi ha una bretxa en un algorisme, una bretxa amb la intenció de danyar, el dany pot ser molt superior. Moltes gràcies!

Liliana Arroyo

Experta en transformació digital; Institut d'Innovació Social (ESADE)

Bon dia, i moltes gràcies, també, per la invitació i l'oportunitat de compartir una mica algunes de les idees que, bé, que floten dels despatxos de l'Institut d'Innovació Social d'ESADE.

La primera idea que m'agradaria traslladar és que tant avui com en la premsa com en molts fòrums sempre es parla de la intel·ligència artificial i el biaix. Moltes vegades es parla, afortunadament, d'intel·ligència artificial i biaix, però de vegades ens oblidem que la intel·ligència artificial o els algorismes en concret han nascut per discriminar, i discriminar entenent-ho des del punt de vista de crear categories, identificar patrons i, a partir d'aquí, aplicar solucions diferents per al grup A, grup B, grup C.

Lavors resulta sorprenent que ens agafi desprevinguts el fet que la intel·ligència artificial discrimina. Entre altres coses perquè és una cosa que fem cada dia. Discriminem quina roba ens posarem al matí en funció del temps; discriminem quan obrim la nevera, segons si volem seguir una dieta saludable o no aquest dia; discriminem quan triem col·legi per a les nostres criatures, etc., o discriminem quan viatgem i perquè funcionem amb estereotips que els sociòlegs en sociologia anomenem *l'imaginari social*.

Lavors, la intel·ligència artificial és un mirall enorme de 1.500 augments del que és l'imaginari social. El problema que ens trobem és que, ara mateix, els imaginaris socials que desenvolupen aquesta intel·ligència artificial són molt pocs i massa homogenis. Dit això de partida, que seria com una mica el diagnòstic sociològic del problema que comentem avui, m'agradaria aprofundir en exemples sobre tipus de biaixos, especialment envers col·lectius vulnerables, amb perspectiva de gènere, raça, però també classisme, capacitisme, edadisme, una sèrie d'eixos de discriminació que són

històrics permanents, però abans d'entrar en aquests exemples, i també acabaré amb alguna bona pràctica, per si podem agafar idees, m'agradaria parlar de tres tipus diferents de biaixos. Segur que n'hi ha més, però crec que aquests tres ens permeten situar-nos una mica.

El primer és el que trobeu a l'esquerra (assenyala imatge) per partida doble, perquè és el que més abunda, és el biaix cognitiu, que bàsicament és les ulleres que portem, les ulleres que ens hem construït per cultura, per educació, per experiència directa, i això són biaixos que estan directament traslladats a aquest mirall que dèiem, aquest mirall d'augment. Són els biaixos cognitius que portem amb nosaltres, el nostre bagatge, la nostra manera de mirar el món.

Després, a dalt a la dreta, hi ha aquest en què posa *code*. Seria una finestreta de codi, ja que hi ha el biaix tècnic, que l'agafo com una categoria àmplia no només del biaix que es pot introduir en el codi, sinó una cosa que s'apuntava també aquí, que és el biaix que poden tenir les dades. Si utilitzem les dades que tenim a mà i les que podem utilitzar, perquè al final per aplicar intel·ligència artificial, i això els experts ho saben millor que jo, les condicions de les dades han de ser molt específiques, que sembla que puguem posar-ho tot, però és molt diferent si les dades són estructurades, si no són estructurades o si no sabem l'origen de les dades, si no sabem quanta manipulació ha patit aquesta base de dades, etc., amb la qual cosa el biaix tècnic pot venir des de l'origen de les dades, la composició d'aquestes dades, unes dades que estiguin desactualitzades... Fins i tot a errors humans a l'hora d'escriure la línia de codi específica, i que no hi hagi validació darrere, aquest seria un altre capítol.

I després tenim el que es diuen els *biaixos emergents*, que són aquells que ni són exactament cognitius, ni són tècnics, és a dir, que es desenvolupen en la mesura que aquesta intel·ligència artificial aprèn, executa aquest aprenentatge automatitzat i es generen noves interaccions i noves possibilitats. I aquí en porto un exemple, que és el bot de Microsoft. Potser molts el coneixereu. Es deia Tai i va durar 24 hores, perquè la van posar a Twitter, crec que era el 2016. No van obrir el compte, intentava ser un perfil d'una noia d'uns 18 o 19 anys que bàsicament el que feia era aprendre, nodrir-se del que llegia a Twitter amb moltíssima capacitat i, a les 24 hores, se'n van adonar que el que havia pogut absorbir era el discurs de l'odi, discurs misogin, i van acabar amb un bot, el que alguns titulars van denominar *bot neonazi misogin*, quan en realitat simulava una adolescent de 19 anys.

Això ens reforça aquesta idea de dir, bé, de vegades, aquest bot en concret potser no tenia biaix cognitiu a l'hora de dissenyar-lo, potser no tenia biaix tècnic, però ens va presentar un biaix emergent, que al final tanca una mica el cercle d'allò que dèiem. No deixa de ser una mica l'imaginari que una plataforma com Twitter alimenta, tot i que, he de dir, nota a banda, que no només hi ha odi a Twitter, sinó que hi ha molts professionals que ens relacionem en aquesta plataforma de forma civilitzada.

Però, bé, el cas és que és un exemple clar de biaix emergent. Dit això, esmento alguns clàssics de passada, per no allargar-me, però aprofundint en aquesta repetició, en aquesta situació repetida que si tu ets una persona de pell blanca, Microsoft Face Plas o IBM seran molt més capaços d'identificar el teu gènere que si ets de pell fosca, però no només si ets de pell fosca, sinó que les dades ens diuen que aquestes plataformes t'identifiquen molt pitjor, amb un error molt més alt, si tu ets dona, no només que tinguis pell fosca, o pell no blanca, sinó que, a més, siguis dona. Aquí tenim una doble discriminació.

Tenim també el cas d'Amazon. Sabeu que Amazon va dissenyar un algoritme de contractació, justament perquè tenien vacants obertes, nivells de recepció de currículums altíssims i difícils de processar en el temps que necessitaven cobrir aquestes vacants professionals, amb la qual cosa van dir: No hi ha problema, prendrem una base de dades on recollirem els perfils professionals de la plantilla d'Amazon dels últims deu anys i dissenyarem un algoritme que bàsicament s'inspiri en aquestes dades.

Què va passar? Que durant aquests deu anys la plantilla professional d'Amazon era bàsicament d'homes, era bàsicament masculina.

Lavors aquest algoritme en el moment en què en el teu currículum apareixia la paraula *dona*, ja no només que en la part de gènere t'identifiquessis com a dona o femenina, sinó que haguessis dit "he estat participant en estudis de desenvolupament tècnic de l'algoritme per a la dona menopàusica". M'ho invento, d'acord? Només perquè aparegués la paraula *dona* en el teu currículum, en alguna part del teu currículum, directament ja et descartava.

Òbviament, gràcies a la tasca periodística, això va sortir a la llum. Van cancel·lar l'algoritme, però fixeu-vos que aquí hi havia aquest biaix d'aquests deu anys. Un altre cas molt conegut i de la mà de *ProPublica*, que van fer una investigació molt a fons del cas de Compass, que és un sistema que intenta predir el risc de reincidència d'una persona que ha estat detinguda. Sabeu de sobres que hi ha un problema de racisme per part de la policia als EUA, i això òbviament es reflecteix també en l'algoritme.

Per posar-vos-en un exemple. A dalt a l'esquerra (assenyala imatge), teniu dues persones a qui es va detenir per possessió de drogues, per tinença de drogues. A l'esquerra de l'esquerra hi ha Dylan Fugett, a qui es va detenir en aquest moment i l'algoritme va diagnosticar o va predir que tenia un risc molt baix de reincidència. Li va posar un número 3, mentre que al costat una persona racialitzada, amb les mateixes condicions, mateixa detenció, sense antecedents penals, etc., tenia un risc directament de 10.

Al cap d'un temps, i això és el que *ProPublica* va aportar amb la seva investigació, van veure què havia passat amb aquestes dues persones, ja que a Dylan, la persona no racialitzada que tenia un risc baix, un risc de reincidència número 3, el van detenir tres vegades més amb tinença de drogues, mentre que a la persona que tenia el risc número 10, la de la dreta, Bernard Parker, se la va detenir una vegada i no va fer falta detenir-la mai més perquè no van trobar que tingués ni drogues, ni cap altre delicta a posteriori.

Malgrat això, si mireu els dos gràfics que hi ha avall, el de l'esquerra presenta la corba de distribució dels riscos de reincidència per part de les persones blanques. Veieu que hi ha com una mena de tobogan que fa així (assenyala una corba amb la mà), que tindria a veure amb una distribució esperable i habitual, mentre que si ets una persona racialitzada, doncs això sembla que ho fan amb daus, perquè és igual on et situïn perquè en realitat la proporció de persones que poden reincidir entre una i deu vegades és pràcticament la mateixa, i aquí és on hi ha la discriminació del cas.

Més exemples. Això va sortir fa uns dies a *The Guardian*. Per exemple, ara hi ha moltíssim interès a ser capaços de llegir les nostres emocions a partir de reconeixement facial, però, bé, també sabem que arrufar el nas no significa el mateix aquí que a Papua Nova Guinea o que a la Xina, o el que sigui, o qualsevol altre gest. Tots sabem que nosaltres, per indicar que estem d'acord amb alguna cosa, fem (mou el cap de dalt a baix) així amb el cap. Si heu treballat amb persones procedents de l'Índia us faran així (mou el cap cap als costats) i no us diuen més o menys, us diuen que sí. Bé, això ho podem extrapolar a molts altres exemples, però novament si assumim que la forma d'expressar les emocions és universal, cometem un error, perquè creurem que eliminem el biaix cognitiu d'aquest algoritme de reconeixement facial, però en realitat partim d'una assumpció errònia. I aquí ja vaig apuntant algunes coses que David va dir ahir per als ombudsman, no pensar que és un problema tècnic, és un problema de drets, i és un problema que fa falta debatre'l.

Més coses. Bé, hi ha un estudi per Bolukbasi i altres investigadors que es titula "L'home és a programador informàtic el que la dona és a mestressa de casa". Bé, doncs això, si poseu a Google les cerques d'infermeria, medicina o desenvolupador web, doncs veureu que tipus d'imatges us surten.

També volia parlar de cas de Joy Buolamwini, no sé si el coneixeu. Bé, aquesta noia porta una màscara blanca a la mà, tres màscares a les imatges si ens fixem en la seva imatge al centre. Ella és una investigadora, fa el seu doctorat –crec que ja l'ha acabat– als EUA, sobre reconeixement facial. El problema era que la base de dades que utilitzava no tenia mostra de persones racialitzades, amb la qual cosa per desenvolupar el seu propi algoritme necessitava utilitzar una màscara amb faccions d'una persona no racialitzada per poder desenvolupar la investigació.

Imagineu-vos que esteu desenvolupant un algoritme, l'algoritme que forma part de la vostra tesi doctoral, i no us reconeix. Imagineu-vos-ho. Això en el nivell científic té moltíssima implicació, però psicològicament, personalment, aquesta persona podria haver decidit abandonar la seva carrera, tret que se li hagués ocorregut dir: A veure, què passa? I és que resulta que es va trobar que l'algoritme sí que reconeixia una màscara, que no és una persona, que no té drets, etc., però no la reconeixia a ella.

Hi ha un documental, el cas d'Holanda que van comentar ahir i comentar tres bones pràctiques, perquè ja estic enfocant la recta final. Aquestes bones pràctiques tenen a veure, per exemple, amb un estudi que acaben de finançar a Austràlia sobre temes de diversitat. Tendim a pensar moltíssim en raça i gènere, o encara pensem poc en això, però ja comencem a parlar-ne més. Però pensem poc en termes de capacitats i les persones discapacitades, l'accés a la tecnologia, per exemple, i la capacitat d'estar representades en els imaginaris socials són molt limitades. L'aspecte interessant d'aquest estudi bàsicament és que generen una aliança on hi ha codisseny per part d'aquestes persones, és a dir, no es desenvolupa l'algoritme per un costat, la solució en un laboratori en un context industrial i després es testarà, sinó que des de l'inici es fan les preguntes de què és el preocupant, què és el rellevant, etc.

A Finlàndia, a Espoo, hi ha un assaig d'intel·ligència artificial molt semblant al que van denunciar a Holanda fa unes setmanes, però el que s'està fent és justament vigilar que el seu algoritme no intensifiqui la vigilància i la recollida de dades en els barris amb la població més pobra, perquè el tema de la qüestió econòmica també és important. Finlàndia, a més, és exemple perquè entén que fa falta conèixer i entendre una mica com funciona la intel·ligència artificial perquè puguem fer *opt in-opt out*, o sigui l'opció que agafem, amb la qual cosa ha fet un curs obert, que qualsevol de vosaltres el pot fer, sobre quins són els elements bàsics de la intel·ligència artificial, una mica en la línia d'alfabetització digital.

Després aquí, més pròxim, tenim exemples locals, i això és d'un informe que va treure Carme Peiró, també fa escasses setmanes. No sé si és avui aquí, però ahir sí que hi era, hi va estar participant. Recull diversos exemples, s'ha assegut amb diverses administracions i, per exemple, n'hi ha una que es diu Riscanvi, de la Generalitat, que és un protocol d'avaluació i gestió del risc de violència en població penitenciària, semblant al Compass, i el que fa bàsicament és que, en comptes de recollir, no sé si el Compass recollia com 300 variables, aquest només en recull 43. Es concentra la informació rellevant. Si una persona amb cert nivell de risc detectat demana un permís penitenciari per sortir a fer el que sigui, se li permet sortir, però s'activen determinades mesures. Aquest algoritme, a més, s'actualitza cada tres anys, s'hi incorporen millores, és a dir, es va revisant, i hi ha una junta de tractament que el pot validar.

L'àrea de drets socials de l'Ajuntament de Barcelona també es dedica a discriminar per decidir a qui se li concedeix una ajuda i el que permet aquest algoritme és, en comptes que prengui la decisió l'algoritme, doncs assistir el professional per poder adjudicar aquestes ajudes, amb la qual cosa un dels efectes que està aconseguint és l'homogeneïtzació de les respostes, per la qual cosa hi ha una certa millora en la resposta i la qualitat d'aquest servei.

I la meua última idea per tancar seria: Per què insistim a aconseguir el que seria el biaix zero quan, en realitat, hem dit a l'inici, si utilitzem la intel·ligència artificial per discriminar, llavors, la pregunta és, o la qüestió és, com podem avançar cap a un biaix que sigui conscient, que sigui consensuat i, sobretot, que sigui responsable?

Moltes gràcies

Itziar de Lecuona

Doctora en Dret i Professora Agregada, Facultat de Medicina i Ciències de la Salut; subdirectora de l'Observatori de Bioètica i Dret-Càtedra UNESCO de Bioètica de la Universitat de Barcelona

Moltes gràcies al Síndic de Greuges i a l'organització per haver aconseguit organitzar aquesta jornada que trobo molt adient i oportuna. I gràcies a tots els participants per les seves aportacions des d'aquest enfocament interdisciplinari. He seguit el treball del Síndic, també a través de la figura del Defensor del Pacient, i crec que és un actor fonamental. Un exemple és el recurs d'inconstitucionalitat sobre la possibilitat de fer perfilaments per orientar la intenció de vot. En aquest sentit, ja deu conèixer les implicacions del cas Cambridge Analytica. Us convido a reflexionar sobre l'ètica o la bio-data-ètica de la intel·ligència artificial. Les seves implicacions ètiques, però també jurídiques i socials.

L'ètica hi és present sempre, l'ètica com a suport al dret i l'ètica també ens condueix a reflexionar sobre les decisions polítiques. Em preocupa tremendament que estiguem aquí com si l'ètica i la regulació de la intel·ligència artificial estigués en el tema zero, com si fos el futur, i és que no és el futur. Jo diria que ja de manera rutinària la intel·ligència artificial s'aplica, i en alguns casos m'atreveria a parlar de passat. Així que crec que els que estem asseguts aquí, siguin ombudsman o siguem diferents perfils, hem d'intervenir-hi, i especialment fer aterrar la legislació, allò amb què ja comptem, doncs el que tenim –que és molt–, s'ha de processar i és el moment d'aplicar-ho. Aquest discurs, aquesta presentació, el 2017, hauria estat parlar del futur, però és que des del maig de 2018 la situació ha canviat perquè des de llavors s'aplica el Reglament general de protecció de dades, a què em referiré, i tenim una decisió política, com l'Estratègia digital i el Llibre blanc de la intel·ligència artificial. A més, Europa defensa una intel·ligència artificial “de confiança”, així ho ha posat de manifest el Grup d'Experts d'Alt Nivell sobre Intel·ligència Artificial.

Europa, per primera vegada, decideix trencar la dependència de l'imperi tecnològic estatunidenc GAFAM, és a dir, de Google, Apple, Facebook, Amazon i Microsoft (excloent-ne aquí als gegants asiàtics). Sobre aquesta excessiva dependència ja alertava Angela Merkel fa uns mesos. Ni els estats membres ni Europa disposen d'infraestructures digitals pròpies. Comprendre les conseqüències d'aquesta situació és fonamental, com ho és assumir que, de moment, necessitem tercers, les *bigtech*. Aquesta dependència ha de ser estudiada detalladament i, en particular, en l'àmbit de la salut, per portar endavant els projectes d'investigació que es financen a través de pagament d'impostos i sobre els quals s'espera un resultat, bé sigui directe o amb impacte en les generacions futures.

La meua aportació a aquest debat és acadèmica, però també pretén ser pràctica, a causa de l'experiència acumulada com a membre de comitès d'ètica de la investigació de centres d'investigació i universitats de caràcter públic. Comitès d'ètica encarregats d'avaluar els aspectes metodològics, ètics i socials de projectes d'investigació i d'innovació basades en el desenvolupament d'algoritmes que es nodreixen de dades personals i que apliquen intel·ligència artificial i altres tecnologies com el *big data* o la realitat virtual, per posar-ne alguns exemples. Des de la filosofia del dret, entenc la bioètica com una bioètica d'impacte, que no només serveixi per repensar la vida i la mort i les implicacions dels avenços científics i tecnològics. En

aquest sentit, el debat intel·lectual és apassionant, i subscriu les paraules de la taula d'ahir de Daniel Innerarity i de José María Lasalle.

Cal prendre decisions, i cal contribuir a avançar en el desplegament i l'aplicació de la intel·ligència artificial amb unes finalitats prèviament consensuades. La meua presentació està orientada, des d'una perspectiva crítica, a entendre el problema i les qüestions a què ens enfrontem, perquè no podem esperar que el ciutadà que cedirà les seves dades personals per desenvolupar algoritmes en l'àmbit mèdic i no ho qüestionarà pugui prendre decisions lliures i informades, si primer no concretem de què parlem i identifiquem els reptes i els problemes. La situació asimètrica que es genera és lògica: qui té les dades, té el poder, i qui les cedeix des de les bones intencions, el perd si no es genera primer un debat social informat i si, havent-hi normes, no s'apliquen adequadament.

I és que hem deixat de ser anònims per ser reidentificables, i tenim un problema si no acceptem que l'anonimat ja no es pot garantir. Actualment està en boga la intel·ligència artificial, però fa res no es parlava del *big data*, quan sobre el que s'hauria de parlar és de la convergència de les tecnologies esmentades. Hem d'entendre que, en la societat digital guiada per la dada, nosaltres som la primera matèria, proveïdors de les dades personals, i també destinataris del resultat de les aplicacions i els serveis basats en els tractaments de dades personals. Les nostres dades nodreixen la intel·ligència artificial i, per tant, és una qüestió que ens ha de preocupar a tots. Hem d'entendre les implicacions que té, especialment, pel que fa al disseny de polítiques i normatives.

Davant els processos d'investigació i innovació basats en l'aplicació de tecnologies com la intel·ligència artificial, el *big data*, la biometria, etc. especialment en l'àmbit de la salut, m'agradaria reflexionar sobre l'ètica de les decisions automatitzades, perquè d'això va el nostre temps. Decisions automatitzades que estan determinades per un algoritme programat per humans, que beu de dades de la realitat i que incorpora biaixos. Per això cal evitar la discriminació algorítmica. Es tracta de decisions que, segons estableix la normativa aplicable, requereixen una revisió humana abans de ser preses. Imagineu-vos les implicacions que això té en l'àmbit mèdic i sapiguen que des de 2018 existeix el dret a reclamar contra aquestes decisions. La intel·ligència artificial ha arribat per prestar un suport, per ajudar a millorar la presa de decisions, però no pot substituir l'humà, i menys en qüestions tan importants com la salut de les persones. El cert és que resulta divertit plantejar aquest futur distòpic, com sempre ha passat, pensem en el poder de la genètica primer i ara el poder de la intel·ligència artificial.

Continuem analitzant els secrets de la intel·ligència artificial com la justícia de l'algoritme i l'equitat de les dades, perquè la qüestió ja no és programar l'algoritme de manera justa. La qüestió és amb quines dades s'alimenta. Després, la responsabilitat que tenim tots els actors. Hem de treballar interdisciplinàriament, des d'un enfocament proactiu i de gestió dels riscos pel que fa als tractaments de dades personals. Sovint es plantegen problemes com si fossin ètics, i no tècnics, quan el que són és precisament això, problemes tècnics com, per exemple, garantir que no es pugui atribuir personalitat al conjunt de dades personals que es gestiona mitjançant les tècniques de pseudonimització adequades. Si no entenem primer els problemes tècnics, la reflexió ètica serà inútil.

La qualitat de les dades personals que nodreixen els algoritmes no és una qüestió intranscendent. Tampoc no ho és la seva gestió i com es comparteixen. Abans ja he fet al·lusió a la falta d'infraestructures per emmagatzemar, desar i explotar dades per part d'Europa, especialment la seva interoperabilitat i reutilització. D'altra banda, és lògic que hi hagi interès a accedir a bases de dades de salut i d'una altra índole, atès que qui té conjunts de dades personals avui, ben estructurades i responen a criteris de qualitat, té el poder. Poder que també s'aconsegueix, amb el temps i l'esforç necessari, combinant bases de dades de qualitat amb unes altres de qualitat relativa

per així dir-ho, per extreure patrons de comportament, predir tendències i millorar així la presa de decisions. I aquest interès en l'accés el té tant la iniciativa pública com la privada i amb finalitats diverses.

En particular, m'interessa reflexionar sobre el valor i el preu de les dades personals en la societat digital; i, sobretot, com hem de combinar, i això no cal negar-ho, la iniciativa pública i la privada, el bé comú, amb els interessos personals. Qui separi en el públic i el privat s'equivoca. La separació en molts casos no és possible, les sinergies entre la iniciativa pública i privada són necessàries i formen part dels plans d'investigació, innovació i desenvolupament dels estats. Així, avui, hauríem de recordar que Europa va aprovar el juny de 2019 la Directiva sobre reutilització de dades del sector públic. No és una idea en fase de desenvolupament, és una aposta política, científica i econòmica amb una regulació aplicable. Així, per avançar en l'aplicació de la intel·ligència artificial és necessari que les dades es puguin trobar, siguin accessibles, interoperables i reutilitzables. En aquest sentit, també hauríem de analitzar els models de consentiment informat per a l'explotació de dades personals amb fins mèdics i d'investigació, per exemple. Només si la persona consent determinats tractaments de dades personals, aquests es podran utilitzar. A aquest model, hi hem d'afegir la possibilitat de consentiment presumpte (*opt out*), és a dir, que, per defecte, estiguem en bases de dades que es destinarien a determinades finalitats com la investigació biomèdica, entre d'altres. Es tractaria d'evitar usos secundaris i discriminacions encobertes en funció de qui pugui tenir accés a les dades i per tant els qui estarien autoritzats i en quines condicions.

Si repasseu la política recent –bé, recent per dir-ho d'alguna manera, perquè del *big data* es parla des de principis dels 2000–, la Unió Europea ho deixa ben clar, “Què pot a fer el *big data* per tu?” I després, com he comentat, l'aliança europea sobre intel·ligència artificial. Fa poc llegíem en la premsa que Margaret Vestarger, una de les vicepresidentes del nou executiu europeu, tenia només 100 dies, i ho ha complert, per presentar la seva nova estratègia, l'he esmentada a l'inici.

Una altra qüestió que m'agradaria portar a debat és la necessitat que els projectes d'investigació i innovació en intel·ligència artificial en l'àmbit de la salut disposin de l'aval d'institucions de reconegut prestigi com a centres d'investigació i hospitalaris. Per a què? Per al que deïeu abans, validar aquelles intervencions, aquelles innovacions, aconseguir el segell de qualitat i les necessàries certificacions. Aporta seguretat i permet un aval ètic de la intervenció que es proposa. Aquesta no és una qüestió menor, atès que el pes d'aquests avals recau sobre els comitès d'ètica de la investigació.

L'ètica per a què hi és? L'ètica parla de la felicitat, de la llibertat, dels valors que una societat vol assolir i sobre els límits i el llindar de risc que estem disposats a suportar.

Realment, ens hem de plantejar si aquesta societat digital ens fa més o menys lliures que l'analògica; i si, efectivament, prenem nosaltres les decisions o les prenen d'altres per nosaltres. Un ciberatac es pot identificar, però no tal discriminació encoberta per raó de l'accés a dades personals perquè no s'han protegit adequadament dades de salut, per exemple, que, com sabeu, són categories especials de dades personals. Vegem al que em refereixo: Quan el meu fill cerqui feina i li diguin que no té el perfil, potser serà perquè té alguna predisposició genètica a patir alguna malaltia, i potser ni ell ho sàpiga. Potser a causa de la nostra ignorància i falta de comprensió del problema s'hagin alliberat aquestes dades personals i estiguin en mans de tercers que no haurien d'haver-hi accedit i així l'afectat sigui discriminat de forma encoberta.

D'altra banda, són desitjables sistemes sanitaris més eficients, una medicina personalitzada, i poder predir els efectes adversos dels medicaments en menys temps, o fins i tot, sistemes de predicció de la COVID-19. Aquestes qüestions, que ja són una prioritat per a la UE des del punt de vista polític, han de donar lloc a intervencions i aplicacions que siguin proporcionals a les finalitats que es persegueixen i respectuoses amb els drets dels implicats.

Per tant, davant aquests escenaris, torno a recalcar que hem de deixar de referir-nos que les dades estan anonimitzades, perquè no és així. La combinació entre la ingent quantitat d'informació emmagatzemada en bases de dades que conté dades personals i l'estat del coneixement científic tècnic actual condueixen a la reidentificació per defecte com a base sobre la qual determinar què volem, què prioritzem i com ho farem perquè respongui als reptes socials del nostre temps, altrament seria maleficient. Es tractaria de combinar aquests conjunts de dades per determinar patrons de comportament, predir conductes i millorar la presa de decisions, ajudats per la intel·ligència artificial.

Les meves línies d'investigació se centren en l'estudi de les implicacions ètiques, jurídiques i socials de la biomedicina i la biotecnologia, i des de fa ja diversos anys analitzo l'aplicació de la intel·ligència artificial i el *big data*, etc., i també el desenvolupament de dispositius digitals i apps en l'àmbit de la salut. En aquest context, i en particular, pel que fa a projectes d'investigació i innovació en salut, la meua aportació consisteix a contribuir al desenvolupament de pautes que permetin millorar els processos d'avaluació dels esmentats projectes per part de comitès d'ètica de la investigació i, especialment, identificar mercats de dades disfressats d'investigació i innovació.

Són nombroses les propostes que es fan per accedir a bases de dades, a bases de dades sanitàries, per exemple, amb interessos que no són pròpiament els de la investigació, però que ho semblen. És obvi que vivim en una societat de mercat en què existeix una clara tendència a la mercantilització del cos humà i les seves parts, i afegiria la monetització de les dades personals, situació gens desitjable. Però aquí ja s'observen tensions. Volem una medicina personalitzada i a la vegada la Unió Europea aposta per un mercat únic digital, basat en dades, que la converteix en més competitiva. Ara us mostraré una sèrie de diapositives perquè vegeu al que em refereixo:

El projecte europeu NESTORE, finançat pel programa marc de la Unió Europea Horitzó 2020, en què participo, és un exemple que s'emmarca en aquestes decisions de política científica i econòmica, que té per objectiu desenvolupar un company virtual per a les persones de certa edat (gent gran, m'atreviria a dir) i no dependents a través de sistemes de monitorització i recomanació per a un envelliment actiu i saludable. És un company virtual amb un finançament gens menyspreable i desenvolupat per un consorci públicoprivat. Fa un any, el 4 de març de 2019, vaig participar en la presentació a Granada de l'Estratègia espanyola d'intel·ligència artificial d'R+D+I, que és un altre exemple de decisió política en l'àmbit de la investigació i innovació amb impacte econòmic.

Si viatgeu a Amsterdam amb avió, en transitar per l'enorme aeroport de Schiphol us n'haureu adonat, i si no aquí ho podreu veure a través d'aquestes fotografies que vaig fer en diverses ocasions, que està absolutament ple d'anuncis d'un dels grans gegants tecnològics. Són anuncis sobre un núvol que canviarà el futur del golf o les capacitats d'anàlisi de l'ADN de manera exponencial. Aquest núvol, i fonamentalment els serveis prestats en núvol –que és on hi ha el valor –, serà capaç d'ajudar a millorar la presa de decisions. Són coneguts els problemes personals que té un dels millors golfistes del nostre temps, si aquesta informació personal es combina amb la posició i la velocitat del cop, el problema és obvi. En aquestes bases de dades que explotaran tercers acreditats és possible que hàgim de començar a tancar informació per complir el principi de minimització (a què s'ha referit el Dr. Argimon) i necessitat de les dades.

Aquest exemple (assenyala diapositiva) és interessant, no? perquè a ningú no li interessa si té problemes amb el sexe, amb l'alcohol, però aquí hi ha la informació, preparada per combinar-se. Ho hauríem de fer? I aquesta és la qüestió: determinades dades personals no poden estar a l'abast de qualsevol ni són necessàries. Continuem amb les fotografies. Fer un escanejat de 7 trillions d'ADN en menys de 12 hores, això abans era impensable. Però novament són les *bigtech* les que ofereixen aquesta possibilitat.

Catalunya és pionera. Catalunya es va prendre molt seriosament seguir les directrius europees, i es va alinear amb el sistema *opt-out*, que es basa en el consentiment presumpte. Pel que fa a aquesta imatge (foto de la targeta sanitària) acumula molta

informació: els fills que depenen de mi, si cotitzo en la seguretat social per ells, el meu consum farmacèutic, les meves voluntats anticipades, si les tinc, el meu codi postal, el sexe, la data de naixement, que com sabeu, permet la reidentificació de la persona.

Llavors, tenim l'obligació de protegir les persones a través de la protecció de les seves dades, la seva dignitat, honor i llibertat estan en joc en funció dels usos. Convé afegir que el model d'història clínica informatitzada i compartida català (HC3) és un referent per la qualitat de les dades, la seva fiabilitat i seguretat. Per això, no és possible parlar de només de futur, perquè ve d'abans. Convé revisar la situació per integrar les qüestions ètiques des del principi i prendre la direcció adequada.

En aquest sentit, el Programa d'anàlisi de dades per a la investigació i la innovació (PADRIS), el programa abans conegut com a VISCA+ (per les seves sigles en català "més valor a la informació en salut", és a dir, més valor a la informació de salut), que va néixer amb una sèrie de problemes i qüestions no resoltes que van fer que es paralitzés per decisió parlamentària precisament i, en resum, perquè no s'havia integrat l'ètica des del principi. Convé no deixar de costat les lliçons apreses, i no repetir errors que poden ser colossals i d'impossible solució quan les dades personals estan en joc.

És necessari fomentar un debat social sobre la intel·ligència artificial i els seus aspectes ètics, per prendre decisions lliures i informades, des de la transparència i la rendició de comptes. Es tracta de determinar quins projectes i iniciatives es prioritzen; si es tracta d'un sistema *opt-out*, com s'informa sobre els usos i les finalitats de manera clara i transparent i com es desenvolupen sistemes de governança en línia de la investigació i innovació responsable que propugna Europa. En resum: a més informació i més transparència, menor serà l'obsessió pel consentiment informat com a garantia. En el nostre sistema, que està basat en la solidaritat, les dades personals no es poden concebre com a moneda de canvi, ni tampoc lliurar-les al millor postor. És una concepció errada de les regles del joc establertes.

Sembla bona idea tenir descarregada una aplicació connectada a la nostra història clínica que ens avisi que tindrem un infart abans de tenir-lo, que previngui i estalviï temps, recursos, i sobretot que salvi vides (assenyala la següent diapositiva d'una aplicació mòbil de salut que anuncia *The Economist*). Però aquest tipus d'aplicacions requereixen, per crear-les i utilitzar-les, l'accés a dades personals, dades de salut, entre d'altres. És necessari alliberar informació confidencial a tercers aliens al context sanitari que poden tenir diferents interessos en els conjunts de dades.

Com es fa investigació avui? Es fan *hackatons* per desenvolupar algorismes (per exemple de predicció de determinades malalties), en el marc d'una competició en què participen programadors que possiblement mai no van rebre formació en ètica mèdica, o en tractament confidencial de dades.

Un altre exemple: empreses dedicades la genètica directa al consumidor. Aquests models de negoci que avui proliferen consisteixen a analitzar mostres d'ADN enviades per correu postal per predir el risc a patir determinades malalties i també (sumant un suplement al preu) informar sobre els teus ancestres, a canvi un preu (força assequible, per cert). Després les farmacèutiques fan acords amb aquestes empreses per accedir a les seves bases de dades (que inclouen informació no només de salut, sinó també d'un altre tipus com ara gustos, aficions, etc.) per desenvolupar intervencions i tractaments. Aquest model d'accés a dades personals per investigar no hi és en els manuals, com tampoc no es pot afirmar rotundament que qui va enviar la mostra de saliva per saber més sobre el seu ADN pensés a ser participant en investigació.

On vull arribar? No hauríem de pensar que les nostres dades no interessin a ningú. Nosaltres, les persones, de forma aïllada i, en grup i les dades que som i generem són d'especial interès, tant per a la iniciativa privada com per a la pública. La nostra societat digital s'alimenta de dades, com hem vist, i cal crear ontologies per poder prendre decisions més personalitzades, per a això cal comptar amb moltes dades.

Tornant a la necessitat de fer aterrar la regulació existent, tinc un especial interès a destacar aquí que, al costat dels principis i els drets que estableix, el Reglament general de protecció de dades europeu també incorpora una novetat del tot rellevant des de l'enfocament de gestió de riscos i la responsabilitat proactiva, que és l'avaluació de l'impacte dels tractaments de dades personals de les persones participants. Aquesta avaluació, que s'ha de dur a terme mitjançant unes metodologies, és de compliment obligat en determinades circumstàncies, com, per exemple, quan es tracten dades de salut, genètiques etc., i permet una anàlisi conjunta de les qüestions tècniques i de seguretat, juntament amb els drets i la perspectiva ètica ja analitzada. D'aquesta avaluació, se'n deriva un pla d'acció a implementar que serà determinant per al tractament adequat de les dades

La proposta concreta és, com deia al principi, que deixem de pensar que la regulació encara ha de venir i fem aterrar el que tenim. Però es tendeix a demandar regulació: regular la intel·ligència artificial, regular el *big data*, i no és possible fer-ho així. El problema és que el Reglament no s'està aplicant, fonamentalment perquè no identifiquem el problema perquè no es comprèn adequadament el no tan nou paradigma digital en què estem immersos. És complex, els canvis científics i tecnològics succeeixen a una velocitat vertiginosa i a més implica canviar l'enfocament per garantir la protecció de dades davant l'aplicació intensiva de noves tecnologies. Aquest es basa a tenir accés a les mesures tècniques i organitzatives dels projectes i iniciatives, des del seu disseny, per poder analitzar l'impacte dels tractaments de dades proposats en les persones afectades.

En aquestes jornades s'està comprovant que tenim referents jurídics, i ètics, que hi ha treballs rellevants sobre la matèria de diferents organismes i organitzacions. Si no som capaços de processar, mai més ben dit, i aplicar aquests marcs de referència, no podem esperar comptar amb ciutadans apoderats per la tecnologia.

Res més, moltes gràcies per la vostra atenció.

José María Sánchez Bursón

Consultor

Moltes gràcies a l'Oficina del Síndic; personalment, al seu equip per permetre'ns compartir amb tots vosaltres un conjunt de reflexions generals sobre el context dels infants, la infància, en els nous escenaris de la intel·ligència artificial. Habitualment, es parla dels riscos i de les dificultats que ens reportarà la intel·ligència artificial a la nostra societat, a les persones. A mi m'agradaria més incidir en les oportunitats que la intel·ligència artificial reportarà als infants en els nous contextos socials regits per noves tecnologies.

Si m'ho permeteu, prenc dos minuts per analitzar l'impacte que l'emergència tecnològica, a través de la intel·ligència artificial i altres tecnologies, realment reportarà a la nova societat a què ens incorporem. Primer, s'apropa una emergència tecnològica descomunal. Us poso una primera diapositiva, basada en la teoria de singularitat, que explica com anirà creixent el potencial dels nostres ordinadors personals i la seva relació amb la intel·ligència humana. Per entendre'ns, el nostre cervell també és un processador d'informació. Bàsicament disposem d'un processador natural, químic/biològic, que processa informació i acaba analitzant al voltant de 10 elevat a 16 operacions per segon. Aquesta és la capacitat intel·lectual de processament que desenvolupa el cervell humà.

Si veieu l'evolució d'un ordinador personal, aquí es pren de referència de valor de mil dòlars, com un mòbil superior, l'any 2010 tenia la capacitat de processar la intel·ligència sintètica del cervell d'una mosca. Més tard, sobre actualment el 2020, assolirà la intel·ligència de processament d'una rata, i s'estima que al voltant de la dècada dels anys 2030 al 2040 disposarà de la capacitat de processament d'una persona humana. És a dir, un ordinador de mil dòlars, com un mòbil nostre, tindrà la mateixa capacitat intel·lectual d'una persona, processarà un volum d'informació

similar a un cervell humà. I s'espera que al voltant dels anys 2050 al 2060 tingui la mateixa capacitat del conjunt de la humanitat, dels 10 mil milions de persones que visquem al món. És a dir, un ordinador de mil euros ens brindarà més capacitat intel·lectual que el conjunt de la humanitat.

Això presentarà un impacte sever en tota la nostra vida, en com vivim, en allò que fem, en què treballem, evidentment, descomunal, inevitable. És curiós perquè tots els processos de desenvolupament de la intel·ligència artificial s'estan construint sobre la replicació del mateix cervell humà. Bàsicament, el que fem és acumular informació i després, a través d'un processador, prenem decisions, discriminant, seleccionant la informació que estimem més rellevant, i els nostres ordinadors artificials fan el mateix, i per això els nostres ordinadors artificials també aprenen, també dedueixen i també podran incorporar l'element ètic, i previsiblement en poc temps seran molt superiors a nosaltres mateixos. Això crec que és important.

A més, tots els processos tecnològics es produeixen d'una forma cada vegada més apressada. La següent diapositiva ens mostra els temps d'adhesió de les tecnologies a la cultura social. Mentre l'electricitat va trigar 50 anys a adherir-se al 25% de la població, avui dia hi ha tecnologia com els jocs de Pokémon que va assolir en dos dies 300 milions de persones. És a dir, la capacitat d'adherència de la societat de la nova emergència tecnològica és explosiva. Això generarà una renovació permanent del nostre coneixement.

S'estima que el 2040 el coneixement humà, és a dir, tot el que sabem, tot el nostre coneixement enciclopèdic, es dupliqui cada dos anys. Fixeu-vos en el potencial que això significa i els grans reptes que s'apropen. No és ja tant el coneixement, sinó el que els prospectivistes i futuròlegs parlen d'anticoneixement, perquè moltes de les coses de què nosaltres estimem que sabem, en realitat són falses, és anticoneixement. Possiblement, gairebé tot el que estudiem al nostre col·legi posteriorment s'ha revelat com a erroni o fals, tot ha estat substituït. En els estudis de química al col·legi, jo em vaig quedar en l'àtom, avui dia se sap que sota de l'àtom hi ha cinc unitats menors. Estudiem els set planetes, ja no sabem molt bé què és un planeta o exoplaneta, planeta gasós, etc. És a dir, el coneixement rellevat és tan potent que realment està transformant d'una forma molt intensa tota la presència de l'home en la nova societat.

I en aquests nous contextos d'acceleració tecnològica se situen els nostres infants. Cal advertir primer que les nostres societats tradicionalment han postergat els drets dels infants. Històricament i fins i tot actualment. Els infants són les persones més invisibles i silencioses de la societat. Al final, si ens dona temps, us presento una petita referència de la invisibilitat que pateixen els nostres infants europeus, no africans, d'Europa. No voten, no són representats, no participen en els mitjans de comunicació, realment són invisibles socialment.

Els sotmetem a infinites normes de protecció, de prevenció de riscos i seguretat sense tenir-ho en compte. És més, quan a occident presumim que els infants són els primers davant una emergència i catàstrofe, això també és fals, són els últims de ser auxiliats. Les poques investigacions de què es disposa referides als infants són ocasionals, queden infrarepresentats en les investigacions comunes, s'adverteix el que deduïm, que realment en els incidents catastròfics massius els infants són els que moren més i en primer lloc. Així que ni això és cert. Realment, són els últims d'aquesta societat, i per això tenim l'esperança que en els nous escenaris es generaran oportunitats sobre la valoració de la infància en les properes societats, en especial per l'enfortiment de les àrees de l'educació i la salut.

Necessitem definir una nova governança i comprometre'ns amb una ètica per al futur, perquè som la primera generació de la humanitat que tindrem la capacitat d'impactar en el conjunt del desenvolupament humà posant en risc la seva pròpia sostenibilitat. Així, les actuals generacions tenim la capacitat d'impactar en el planeta, a través de l'escalfament global, la sobreexplotació, el creixement demogràfic, la

contaminació, etc., i posar en crisi la mateixa sostenibilitat humana per a les recents generacions.

Els infants són molts al món, però cada vegada són molts menys, i Espanya és un cas paradigmàtic del que parlem. Fa 100 anys els infants representaven el 40% de la població humana i ara només són al voltant del 28% de la població, i seguiran reduint-se, i a més la majoria d'aquests infants viuen als països en desenvolupament. En països més avançats com el nostre la preeminència dels infants s'ha reduït ostensiblement. De fet, les nostres piràmides poblacionals estan truncades. El Japó representa avui el cas més extrem, per bé que a Espanya som els tercers més vells del món, segons els indicadors que utilitzem, però la població mitjana espanyola està al voltant dels 44 anys d'edat, i la població votant, que és molt significatiu, perquè són els que prenen les decisions, està entre 52 a 53 anys. I per això estímem que els infants cada vegada seran menys, però en les noves societats seran molt significatius, amb una sèrie de variables en què no podem ara aprofundir, però que van connectades amb la capacitat per a la innovació i la creativitat, l'adherència a l'apropiació tecnològica, la piràmide del coneixement invertida, la importància de l'educació i altres raons.

D'altra banda, d'alguna manera la intel·ligència artificial possibilitarà que els nostres infants siguin infants amb superpoders. Són infants que naixeran amb el anomenem una *personalitat enriquida*, una personalitat ampliada. Gaudirem d'una realitat molt augmentada. Mai no s'ha donat una bretxa biològica i tecnològica tan severa entre nosaltres i els nostres fills o els nostres nets. És totalment diferent com neixen, creixen i viuen. Ells, al cap de poc temps de ser concebuts, als tres mesos, tenen ja un vídeo en 3D de l'ecografia de la mare. Abans de néixer, ja un percentatge important són presents a les xarxes socials, tindran un accés immediat a les tecnologies més avançades en molt poc temps i neixen en un procés caracteritzat per una bretxa biològica radical entre com s'ha conformat la nostra vida i com s'ordenarà la dels nostres infants.

L'obtenció de les habilitats digitals marca aquest rumb. Un estudi de l'empresa d'antivirus AVG analitza les habilitats digitals dels infants i adverteix que són més accelerades les digitals que les manuals i tradicionals. L'infant aprèn a navegar a Internet molt abans que a nedar en una piscina, a fer les cerques abans que caminar amb bicicleta, el ratolí abans que lligar-se els cordons, i així diverses funcions. Les habilitats digitals dels nostres infants són molt més avançades que fins i tot les habilitats manuals ordinàries, i això generarà uns superinfants del segle XXI amb superpoders d'informació, comunicació, coneixement, connexió, etc. La investigadora catalana Dolors Reig va publicar fa uns anys un llibre sobre els joves hiperconnectats, i ens indica que es tracta d'infants connectats 24 hores, més ràpids en les seves comunicacions i desenvolupament llenguatge de símbols, practiquen la multitasca, no distingeixen món en línia de fora de línia (real-virtual), fan servir noves formes d'expressió i, paradoxalment, són cada vegada hiperindividualistes, però al seu torn més sociables.

Resumint, el previsible és que el desenvolupament tecnològic acabi promovent un gran potencial dels nostres infants, que siguin més desenvolupats intel·lectualment i socialment, més apoderats, i que gaudeixin d'un gran potencial de vida. L'element capital serà el procés d'educació i aprenentatge que emprenguin. En aquesta àrea igualment canvien les tendències i els procediments de producció.

D'entrada, l'educació també es globalitza. Un informe de l'Observatori Internacional de la Infància d'UNICEF a Florència ens explica com la capacitat educativa dels infants també s'ha globalitzat, s'ha fet molt extensa. Abans educàvem les famílies, els col·legis i els contextos propers, avui dia l'educació s'ha obert, eduquen els mitjans de comunicació, Internet, la televisió... Incorporen més valors els grans líders esportius, els líders del joc, més que la mateixa relació personal amb els pares. I, en general, una sèrie de variables que canviaran transcendentalment l'evolució dels models educatius dels nostres infants i el valor del coneixement. Evidentment, l'educació serà vital en les noves societats.

Per formular un raonament, segons les prospectives, els infants que avui neixen viuran al voltant dels 100 anys. Com ja s'ha comentat, en un món poblat per 12.000 milions de persones, i els infants durant la seva vida desplegaran diferents identitats personals, familiars, professionals en un context en què el coneixement es doblarà en un període curt de temps. És a dir, no tenim ni idea de com serà el món el 2040 o 2050. Imagineu com una professora d'escola bressol o de preescolar enfoca l'aprenentatge, el coneixement d'aquests infants, amb l'escenari d'incertesa que hi ha.

I la intel·ligència artificial modelarà tot el procés educatiu, això és un dels grans reptes que assumim els qui intentem activar els drets dels infants. És a dir, com integrem en el sistema educatiu les noves realitats i no aplicar-los els nostres patrons educatius de les èpoques tradicionals. Canviaran la pedagogia, la dieta cognitiva, els plans d'estudi. Fins i tot la mateixa supervivència de l'escola, segons manifesta la UNESCO, ja que entre diferents tendències, algunes consideren la seva desaparició.

El mateix ocorre en el món de la salut, que canviarà radicalment. Altres companys que m'han precedit ja ho han indicat, i no fa falta que ho reiteri. De la mateixa manera, tota la governança, tot el que significa la incorporació dels infants en els espais públics i en la planificació estratègica dels països, l'aprovació dels pressupostos i altres assumptes públics. Actualment Noruega és l'únic país del món que aplica una anàlisi de l'impacte dels pressupostos públics en les noves generacions. I això aporta un element capital en els resultats de societat que construïm. A la part final de l'exposició intentaré exposar-ho.

I també tenim la necessitat de construir una nova ètica del futur, d'un futur que tingui en compte la sostenibilitat a llarg termini de la nostra humanitat, la convivència entre les noves generacions i les generacions que han viscut, i construir un planeta i una cultura que sigui realment respectuosa, amb una cultura d'aprenentatge, amb l'atenció i el desenvolupament dels infants i amb els drets dels infants.

I respecte als reptes que assumim. Cal partir d'una breu reflexió. Assistim a un procés, no de confrontació amb les tecnologies, sinó al contrari, un procés de fusió, d'absoluta integració tecnològica. És a dir, ara vivim el que es diu *l'hivern negre*. Pensem més en el control, la protecció, en tot allò de què hem estat parlant anteriorment. Ens fan por la inseguretat i la incertesa, però les tecnologies ens portaran a abraçar-nos als homes, els electrons s'incorporaran en la nostra pròpia biologia, i anem a processos de transhumanisme, de realitat augmentada, d'humanisme reinventat, humanisme augmentat, que és l'escenari en què viuran els nostres infants. I els grans reptes, evidentment respecte a la infància, impliquen trencar en aquests nous contextos amb la invisibilitat dels nostres infants, que segueix sent molt severa malgrat les circumstàncies i l'evolució que s'ha produït. La invisibilitat de la infància continua molt ferma en la nostra societat, i queda reflectida fidelment, per exemple, en els nostres pressupostos públics: el 57% de les despeses públiques van dirigides a les classes majors, i només el 7% als infants. La dedicació dels nostres propis pressupostos públics manifesta el sentir de les nostres decisions actuals, i això evidentment cal corregir-ho, i encara més respecte a les noves generacions a què hem de confiar tota la nostra pròpia sostenibilitat social.

En la realitat espanyola, l'efecte es fa pervers. Tenim un deute públic pròxim al 100%, que estem finançant amb crèdit a 20, 30 i fins a 40 anys a vista, que hauran de pagar els nostres fills i els nostres nets, i part del deute que pagaran els nostres nets obeeix als nostres creditors actuals. N'és un exemple el Fons Sobirà de Noruega, que es nodreix de la venda del petroli extret al seu país, que l'inverteix en benefici de les seves noves generacions per a d'aquí a 25 anys. Per tant, els nostres nets hauran de finançar la vida dels nets dels noruecs, perquè hauran de pagar els nostres deutes, mentre exploten els crèdits dels noruecs presents, perquè han aplicat el contrast de l'anàlisi dels pressupostos en les noves generacions i han tingut en compte la sostenibilitat generacional.

Realment, la sostenibilitat en el nostre desenvolupament humà es construeix sobre la base del tractament que li oferim als nostres fills, i els nous contextos socials ajudaran a valoritzar la infància com a remesa intel·lectual imprescindible per construir les societats del futur. Sense infància no hi ha futur, i encara menys progrés. Avui dia hi ha més de mil milions de nens i nenes interactuant en múltiples xarxes socials, de forma oberta, lliure, acrítica i no jerarquizada, i ho fan malgrat els riscos, prodigant un comportament basat en la col·laboració i cooperació mútua. Confiant en ells, esperem que el futur proveeixi un pronòstic molt més significatiu dels drets dels infants en noves societats més integratives i creatives. Res més, estic a la vostra disposició per a tot allò que vulgueu preguntar.

Moltes gràcies!

Intel·ligència artificial i drets humans

reptes, rols i eines de les institucions d'ombudsman

Dilluns, 2 de març de 2020

Drets, reptes: Tractats internacionals i definicions

Moderador: Miquel Molina, director adjunt de *La Vanguardia*

Gregor Stojin, president del CAHAI

Martha Stickings, analista de polítiques, Agència de Drets Fonamentals (FRA)

Geraldine Mattioli, assessora al Comissari Europeu de Drets Humans (Consell d'Europa)

Peter Bonnor, jurista principal de l'Ombudsman Europeu

Miquel Molina

Director adjunt de *La Vanguardia*

Per a mi és un honor participar en aquesta trobada. Soc Miquel Molina, soc director adjunt de *La Vanguardia* i escric normalment cada diumenge sobre temes de cultura i ciutats. I en el marc d'aquestes coses que escric, en sintonia amb algunes institucions i administracions, estic interessat, i per això m'agrada molt parlar en aquest àmbit, que Barcelona es converteixi d'alguna manera en una de les capitals del debat sobre l'humanisme tecnològic o la tecnoètica.

Aquí tenim diverses coses a favor: tenim centres d'investigació molt rellevants amb àrees com la intel·ligència artificial, la robòtica, la fòtica o la supercomputació, entre d'altres; tenim el Mobile World Congress, com sabeu, i tenim també una energia política important que es pot canalitzar de moltes maneres, i una d'elles és que Barcelona tingui aquest paper rellevant. Amb això coincidim, doncs, a banda del meu diari, com us deia *La Vanguardia*, la Sindicatura de Greuges, que també està compromesa amb aquesta idea de convertir Barcelona en un referent, en un dels referents mundials d'aquest debat, i també institucions com la Digital Photo Society, l'Ajuntament, la Generalitat i el Govern espanyol, que també tenen aquest propòsit.

Avui, del que parlarem en aquest debat és sobretot del marc legal, sobre què és aquesta realitat, aquest futur, com s'ha dit molt bé a l'anterior xerrada, aquesta realitat que ja és passat, realitat i futur de la intel·ligència artificial. No cal a aquestes alçades, després de tots els debats que heu tingut des d'ahir, que entrem en detall sobre de què parlem quan parlem dels avantatges de la intel·ligència artificial o quan parlem dels problemes que incorpora, els riscos, les amenaces que incorpora la intel·ligència artificial.

Amb aquest debat d'avui es planteja un repte important i interessant: com la Comissió Europea i la Unió Europea en general, els països membres, hi han de participar, han d'estar a l'alçada d'aquesta cursa tecnològica que hi ha al voltant de la intel·ligència artificial, com han d'evitar el risc que es creï un mercat fragmentat que faci menys competitives les empreses europees, i al mateix temps mantenir aquests estàndards europeus, que són molt més alts que els d'altres economies avançades, estàndards europeus de garantia de drets als seus ciutadans.

Per respondre aquestes qüestions, tenim aquí la Martha Stickings, analista política i responsable de cooperació institucional en la Fundamental Rights Agency, una agència europea; tenim la Geraldine Mattioli-Zeltner, que és assessora de la Comissària Europea de Drets Humans, i tenim en Peter Bonnor, que és jurista principal de l'Ombudsman Europeu. Teníem també en Gregor Stojin, que era el president del Comitè Ad-Hoc creat per la Comissió Europea per a la Intel·ligència Artificial, però ha excusat la seva presència i ens ha enviat un vídeo amb la seva primera intervenció. Ho passarem primer perquè així ja entrem després en el debat humà.

Gregor Stojin

President del CAHAJ

Hola i salutacions des d'Eslovènia.

Lamento no poder ser al seminari personalment i saludar-vos a Barcelona, però viatjava via l'aeroport de Venècia i, a causa de les mesures preses amb motiu del recent brot de coronavirus al nord d'Itàlia, podria trobar-me amb complicacions a la tornada, i també després per desplaçar-me a algunes institucions europees, perquè estan limitant l'accés als viatgers procedents del nord d'Itàlia.

En qualsevol cas, he de dir que em sento una mica incòmode per haver de fer la meua intervenció per vídeo preenregistrar, perquè no puc debatre interactivament amb vosaltres sobre aquest tema. No sé què és el que ja s'ha dit sobre intel·ligència artificial durant les sessions d'avui. He revisat els punts de l'agenda i veig que és molt interessant, i m'hauria agradat participar-hi. Intentaré explicar esquemàticament el que estem fent al Comitè Ad-Hoc sobre Intel·ligència Artificial al Consell d'Europa, i també per què ho estem fent, què estem intentant aconseguir i quins són els resultats esperats.

Actualment, sembla que gran part del desenvolupament global en el camp de la intel·ligència artificial es basa en l'autoregulació. L'autoregulació, al seu torn, es basa en la confiança, i la confiança s'ha infringit diverses vegades en el passat recent pel que fa a la intel·ligència artificial i als seus efectes en els drets humans. Necessitem restablir la confiança, també com a condició per a un ús més responsable de la tecnologia per part de les autoritats i, a través d'això, garantir el foment d'un major desenvolupament de la tecnologia responsable i fiable i que les seves aplicacions puguin incrementar els beneficis per a la societat.

Però, a què ens referim quan diem *regulació*? La falta de regulació representa una regulació en si mateixa. Podríem substituir la paraula per *política* o *governança*, o simplement limitar-la a la legislació, però què és més important tenir en compte? També necessitem comprendre l'impacte autònom de les forces del mercat, les normes socials i la tecnologia en el procés.

No tot el que és factible des d'una perspectiva tecnològica pot considerar-se acceptable des d'una perspectiva de drets humans. Actualment, molts organismes nacionals, internacionals i supranacionals estan desenvolupant o avaluant diversos tipus de mecanismes reguladors per abordar els problemes que plantegen l'aparició i la proliferació de les tecnologies d'IA.

Actualment, hi ha més de dos-cents instruments de dret no vinculants. És adequat el marc legal actual? Poden els documents no vinculants, com les recomanacions i les directrius, protegir eficaçment els nivells desitjats de drets humans? Hi ha algunes qüestions que val la pena tenir en compte per aconseguir les solucions més efectives que equilibrin de manera adequada i apropiada els aspectes beneficiosos del desenvolupament tecnològic amb la protecció simultània de les persones i la societat davant els aspectes negatius, tant intencionats com involuntaris. Els impactes que són contraris als estàndards de drets humans existents ja recauen sota la responsabilitat preventiva dels estats. Tanmateix, s'aborden de manera eficaç?

Els poders reguladors són principalment competència dels estats. Dins les tecnologies i les activitats amb efectes extraterritorials significatius o fins i tot globals, és desitjable tenir regles comunes per garantir la claredat i impedir una deriva al'abisme en posar massa èmfasi en les possibilitats tecnològiques. Les organitzacions internacionals i supranacionals en què sovint es delega aquesta tasca poden diferir i difereixen en la seva jurisdicció, les competències i el mandat, i poden produir documents i normes amb diferent influència o força vinculant, la qual cosa és rellevant, per exemple, per al compliment en casos individuals. El Consell d'Europa és un dels pocs que pot elaborar instruments jurídicament vinculants. Els diferents

enfocaments no han de considerar-se competència, sinó contribucions complementàries als marcs en desenvolupament. Per exemple: les definicions i classificacions comunes són més valuoses per garantir la claredat i la seguretat jurídica potencial i permetre una plataforma comuna per a l'avaluació de riscos i oportunitats. La duplicitat, d'altra banda, pot causar confusió i possibles divergències en les pràctiques d'implementació, i s'ha d'evitar llevat que estigui ben fonamentada.

Probablement tots estem familiaritzats amb els exemples relacionats amb la justícia penal per a l'avaluació de delinqüents o de la reincidència, que han demostrat que el biaix en les dades utilitzades per als models de capacitat es replica i augmenta quan la tecnologia s'aplica a escenaris de la vida real.

Els problemes també poden sorgir dels mateixos algoritmes, ja sigui a causa del treball de dissenyadors i programadors o a causa de la naturalesa del procés de presa de decisions. Aquest és probablement l'aspecte més rellevant de la IA que encara cerca una solució adequada: les caixes negres no són prou transparents o no mostren o no permeten una explicació adequada i el consegüent escrutini legal.

De vegades, els models es basen en suposicions que tenen poc o cap valor científic, com ara l'avaluació de les emocions basada en el reconeixement d'imatges. A vegades, aquestes eines s'utilitzen en un entorn privat, com en l'àmbit de recursos humans, ja sigui per seleccionar candidats per a una feina o per al cessament d'una relació laboral. De vegades, també és impossible saber si s'ha utilitzat la IA en el procés. L'opacitat es pot produir a diferents nivells.

Un altre tipus de problemes sorgeixen de l'ús sense escrúpols de la tecnologia sense una avaluació adequada dels riscos o impactes, un exemple destacat del qual és la tecnologia de reconeixement facial. Poden haver-hi instruments legals per protegir-nos, però la pregunta és si són prou efectius i si tenen la suficient capacitat de resposta. Aquestes van ser les conclusions sobre els reptes als règims de dret de responsabilitat civil existents, recopilades en un estudi recent de la UE sobre responsabilitat en la intel·ligència artificial. I, si m'ho permeteu, llegiré un extracte de l'estudi. Diu així: "És possible aplicar els règims de responsabilitat existents a les tecnologies digitals emergents, però en vista d'una sèrie de reptes i a causa de les limitacions dels règims actuals, fer-ho pot deixar les víctimes amb una compensació insuficient o sense cap compensació. L'adequació de les normes de responsabilitat existents pot ser, per tant, qüestionable, tenint en compte especialment que aquestes lleis van ser formulades fa dècades o fins i tot segles, basades en conceptes encara més antics i incorporant un model principalment antropocèntric i monocausal d'infligir dany"¹.

Un altre exemple d'aplicació dels mecanismes legals existents també el va donar una recent decisió holandesa sobre el programa d'aprenentatge automatitzat anomenat SyRi, que va ser utilitzat per les autoritats holandeses per predir quins ciutadans és probable que cometin algun tipus de frau en matèria d'habitatge o en prestacions socials. La sentència, que es va emetre a principis de febrer, és històrica perquè es va basar en el dret a la vida privada que estableix l'article 8 de la Convenció europea de drets humans. Al tribunal li va semblar problemàtica la falta de transparència sobre com funciona el sistema de puntuació de riscos algorítmics, i va determinar que la legislació SyRi no superava la prova d'equilibri, que requereix que es ponderi qualsevol interès social davant la violació de la vida privada de les persones amb un equilibri just i raonable. En opinió del tribunal, però, el sistema automatitzat d'avaluació de riscos no va passar aquesta prova. Era una caixa negra i, com a tal, no era legal.²

Les diferents etapes del cicle de vida del sistema d'IA poden produir diferents riscos i implicacions, amb diversos nivells de gravetat, tant en termes d'actors com de parts interessades. A més, la definició, la tecnologia, els impactes, l'ús de la tecnologia...

¹ Vegeu *Liability for Artificial Intelligence and other emerging digital technologies*, informe del Grup d'Experts sobre Responsabilitat i Noves Tecnologies, Comissió Europea, p. 19; <https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608e>

tots són un objectiu en moviment. La comprensió clara dels efectes, l'impacte i els riscos requereixen tant la classificació com la categorització dels sistemes i aplicacions d'IA, un seguiment continu i una avaluació interdisciplinària de les conclusions. Això eventualment pot conduir a una verificació, validació o certificació. L'avaluació de l'impacte i el risc en els drets humans sembla gairebé necessària i requeriria no només una anàlisi tècnica, sinó també diversos coneixements i aportacions humanístiques. Hi ha diverses opcions que es poden aplicar i és clar que no pot haver-hi una solució que s'adapti a totes.

Llavors, com ho fem? El CAHAI, que és un acrònim de Comitè Ad-Hoc d'Intel·ligència Artificial, ha rebut l'encàrrec del Comitè de Ministres del Consell d'Europa de preparar un estudi de viabilitat d'un marc legal i els seus elements potencials per al disseny, el desplegament i l'aplicació de la IA en el context de les normes vigents al Consell d'Europa per a la protecció dels drets humans, la democràcia i l'estat de dret. Va tenir el seu primer plenari el novembre de 2019 i s'espera que finalitzi l'estudi de viabilitat a finals de 2021.

Durant la primera etapa, el CAHAI treballa en la identificació de riscos i oportunitats de la intel·ligència artificial tenint en compte els estàndards de drets humans i en el mapaatge d'instruments vinculants i no vinculants. L'objectiu és establir una visió general integral del marc legal existent i l'avaluació de les seves capacitats per abordar els efectes més urgents. Amb el temps, això pot conduir a la identificació de llacunes i àrees que han d'abordar-se de nou.

El Consell d'Europa té un recorregut substancial en l'aspecte més ampli del tema. A més, el Conveni Europeu de Drets Humans és el bressol dels convenis 108 i 108+ sobre protecció de dades, que de vegades es consideren antecessors del Reglament general de protecció de dades, i també és el bressol de la Convenció de Budapest sobre la ciberdelinqüència. Actualment hi ha més de quaranta iniciatives més al Consell d'Europa, inclosa, per exemple, la recomanació del Comissari de Drets Humans, que va elaborar el document "Unboxing AI: 10 steps to protect Human Rights" (Desxifrant la IA: 10 passos per protegir els drets humans).

També cal mencionar que la CEPEJ, la Comissió Europea per a l'Eficiència de la Justícia, va ser el 2018 la primera a adoptar la carta ètica sobre l'ús de la IA en els sistemes judicials. Què es podria considerar també parcialment i històricament rellevant? Tenim un exemple de 1964 amb la convenció que va establir la Farmacopea Europea, que és una institució que verifica la idoneïtat dels productes i procediments mèdics que es comercialitzen al mercat europeu. Abans d'això, la reglamentació dels productes mèdics estava descentralitzada i era d'interès per a la salut pública establir un mecanisme d'aquest tipus.

Pot ser que aquest sigui el camí que també hàgim de prendre amb les tecnologies modernes, com la intel·ligència artificial. Per descomptat que no es tracta d'un producte estàtic, com comprimits o píndoles o procediments mèdics. Estem davant d'un sistema dinàmic. Però hi ha alguns aspectes que són similars; per exemple, analitzem: El sistema fa el que diu que fa? També fa alguna cosa més? De què està compost? Quins són els seus efectes? Quins són els seus efectes secundaris? Quines són les indicacions? Quins són les contraindicacions?

El que volem fer és identificar les aplicacions i els efectes problemàtics de la intel·ligència artificial des de la perspectiva de garantir el respecte dels drets humans, l'estat de dret i la democràcia. Volem identificar els valors i principis clau que estan sorgint des d'una perspectiva ascendent des d'enfocaments sectorials i des de pautes ètiques i després des de dalt cap a baix: drets fonamentals, democràcia i requisits de l'estat de dret. Hi poden haver línies vermelles que sembla que no s'haurien de permetre creuar, per

² Vegeu "Blackbox welfare fraud detection system breaches human rights, Dutch court rules", Natasha Llomes, *Techcrunch*, 6 de febrer de 2020; <https://tcm.ch/376Jq8a>

exemple: vigilància massiva impulsada per la intel·ligència artificial, qualificació social com la coneixem en certes parts del món, deepfake (manipulació de vídeos utilitzant tècniques d'intel·ligència artificial) i els seus efectes en la democràcia. Com abordem això? I molt probablement la prohibició o no permetre el debat sobre la personalitat jurídica de la IA.

I després volem establir possibles opcions d'un marc legal del Consell d'Europa tenint en compte el contingut, les direccions, el valor afegit i les expectatives dels estats membres i de les múltiples parts interessades sobre les diferents opcions que es poden utilitzar. I aquestes poden ser: una actualització dels instruments jurídicament vinculants existents. Podria ser una nova convenció, podria ser una nova convenció marc, podrien ser instruments de dret no vinculant, com ara recomanacions o directrius. També podria ser un altre tipus de suport als estats membres. També hem de valorar la possible complementarietat entre els elements horitzontals i transversals que podrien formar part d'un instrument tipus convenció i el món sectorial vertical que podria donar lloc a instruments específics de diferent naturalesa. A més, el que és de vital importància: l'estudi de viabilitat haurà d'abordar els mecanismes operatius per assegurar l'efectivitat del marc legal.

Això es fa mitjançant una àmplia participació de múltiples parts interessades. A més dels quaranta-set estats membres, comptem amb una sèrie de representants d'uns altres òrgans del Consell d'Europa, però també d'uns altres països que serveixen com a observadors del Consell d'Europa, organitzacions no governamentals, societats civils i també empreses.

En la meua opinió, és important aprofitar el treball existent, especialment en termes de desenvolupar un grup de definicions i classificacions complet i mútuament acceptable. És per això que també incloem diverses organitzacions internacionals com a participants i observadors. Tanmateix, és massa aviat per predir si els enfocaments eventualment convergiran i en quin grau, però hem d'evitar la distracció i hem d'evitar els retards. Hem d'evitar la deriva cap a l'abisme per intentar-ho durant massa temps o per intentar cercar un consens mínim. Això probablement no seria suficient per al nostre repte.

El que és necessari a llarg termini quan es tracta de l'impacte de la IA en els drets humans és: en primer lloc, acceptació de la responsabilitat política pels riscos i impactes no regulats; en segon lloc, coneixement de possibles instruments de política, i en tercer lloc, proporcionalitat reguladora i agilitat quan s'utilitzen, cosa que requereix (en quart lloc) una avaluació constant, contínua i objectiva de l'impacte i l'efectivitat dels mecanismes. En conclusió, aquest és el resum general del que estem intentant fer al CAHA. Encara estem en les primeres etapes de l'exercici de mapatge, però esperem poder contribuir al desenvolupament del camp i a la protecció dels drets humans, l'estat de dret i la democràcia.

Amb això acabo el meu missatge i us desitjo una molt bona continuació del taller.

Martha Stickings

Analista de polítiques, Agència de Drets Fonamentals (FRA)

Moltes gràcies, president, gràcies també als organitzadors per convidar l'Agència de Drets Fonamentals a participar en aquest important esdeveniment. És un gran plaer ser aquí i poder parlar una mica sobre la feina que hem estat fent en intel·ligència artificial i les consideracions sobre drets fonamentals que entren en joc com a part d'aquest debat.

Vull explicar-vos quatre coses sobre l'organització per a la qual treballo. Som una agència independent de la Unió Europea amb el mandat de recopilar dades i proporcionar evidències i anàlisi a les institucions de la UE i als estats membres quan apliquen la llei de la Unió. Com tots, crec que la intel·ligència artificial i una agenda

digital més extensa és una part cada vegada més important de la nostra feina, i tenim diverses iniciatives que us explicaré una mica i veurem on posa el focus la IA en particular.

Abans de continuar, he de dir breument que, com heu sentit en la meua presentació, som una agència independent, de manera que no puc parlar sobre el posicionament de la Unió Europea o el de les comissions, però em centraré en alguns dels nostres punts de vista des d'una perspectiva de drets fonamentals, en termes d'algunes de les principals inquietuds i consideracions en aquest debat.

Per començar, potser és útil reflexionar ràpidament sobre quina és la situació actual. Ja hem sentit molts dels ponents en aquest esdeveniment dir que hi ha molts desenvolupaments de polítiques diferents en l'àmbit europeu, i veiem que les consideracions sobre drets fonamentals entren en joc com a part de molts d'aquests debats. Hi ha massa iniciatives per esmentar-les totes i potser val la pena destacar-ne algunes que perquè us feu la idea de com els debats al voltant de la IA van arribar a l'agenda política.

El 2017, el Parlament Europeu va adoptar dues resolucions que van destacar la importància de les macrodades, la intel·ligència artificial i la robòtica, i també es van referir a la importància dels drets fonamentals. També hi va haver conclusions del Consell el 2017, després de les quals la Comissió Europea va publicar la Comunicació sobre IA per a Europa i va establir el grup d'experts d'alt nivell sobre IA de què ja hem sentit a parlar una mica. Totes dues iniciatives inclouen una clara referència als drets fonamentals, un dels nostres caps d'unitat és membre d'aquest grup d'alt nivell. Posteriorment, per descomptat, més recentment, al febrer, es va publicar el Llibre blanc de la Comissió sobre intel·ligència artificial, de què ja heu sentit a parlar bastant. Però també val la pena recordar que aquest va ser només un element d'un paquet més ampli de documents que es va publicar al febrer i que hi va haver altres informes sobre qüestions de responsabilitat, hi va haver una estratègia europea de dades i crec que és útil situar la IA en aquest context més ampli.

En l'àmbit de Consell d'Europa, ja heu sentit a parlar del Comitè CAHAI, però també hi ha el treball de CEPEJ i altres comitès, i estic segura que sentireu més coses sobre això més endavant.

Atès el nombre d'iniciatives que hi ha, hem mirat de mantenir-nos al dia amb aquesta proliferació i tenim una llista en la nostra pàgina web de moltes de les diferents iniciatives de la UE, a més de molts altres temes, que actualment ascendeix a 216 i escaig, i com està en constant increment duem a terme actualitzacions periòdiques d'aquestes dades.

Algunes de les activitats recents fan referència explícita als drets fonamentals i també hem de reconèixer que moltes d'aquestes iniciatives prefereixen marcs ètics menys exigents i hem sentit moltes coses sobre l'ètica en els debats que s'han fet fins ara aquí. Per descomptat, és bastant atractiu centrar-se en l'ètica, la qual té forts vincles amb la confiança, i novament hem sentit a parlar molt sobre la confiança dels consumidors en els productes. Es considera que fomenta la innovació i dona suport als models comercials per aprofitar els avantatges del mercat únic i, per descomptat, sabem que hi ha una competició entre el que està succeint en l'àmbit europeu i altres desenvolupaments, com ara als EUA i la Xina, per exemple.

Des de la nostra perspectiva, desconfiaríem una mica de l'ús de l'ètica, almenys de manera aïllada, sobre la base que s'infereixen connotacions de voluntarietat que no són exigibles pels règims legals i que no estan necessàriament estandarditzades, per la qual cosa podria haver-hi una espècie de diferents models ètics que s'apliquen en diferents contextos. Com a conseqüència, des de la perspectiva de l'Agència de Drets Fonamentals, recomanem encaridament l'ús dels drets fonamentals com a marc i com a punt de partida per a l'evolució de qualsevol llei i política en matèria d'IA.

Això es deu en part al fet que, a diferència de l'ètica, els drets humans són instruments jurídicament vinculants. També els drets humans situen la responsabilitat de manera molt clara en els estats, perquè els estats tenen la responsabilitat última de protegir i promoure els drets humans. També perquè nosaltres tenim una gran experiència en termes d'ús dels drets humans per garantir proteccions i en la utilització d'aquest marc legal per donar suport legislacions posteriors i també per fer-hi esmenes. I aquí crec que val la pena reflexionar sobre el fet que els drets humans ja s'han utilitzat com un marc molt útil en algunes àrees relacionades. Llavors, per exemple, podem pensar en la guia sobre empreses i drets humans, i en el fet que alguns dels principis que s'han establert en aquest context també es poden incloure en el debat sobre la IA.

I, finalment, vull dir que l'avantatge de l'enfocament basat en els drets humans és que ens permet centrar-nos en el dany particular que s'infligeix a les persones, i hi ha una referència a això també en el Llibre blanc de la Comissió.

Breument comentaré la nostra tasca en intel·ligència artificial. Intentem donar suport a alguns dels processos de polítiques que he esmentat anteriorment, amb evidències i consells que segueixin aquest enfocament basat en els drets fonamentals. Fins ara hem publicat tres articles, el primer dels quals analitza la discriminació en el context de la presa de decisions automatitzada, el segon analitza la qüestió de la qualitat de les dades. De tots dos temes, n'hem sentit a parlar molt aquest últim dia i, més recentment, en un article sobre la tecnologia de reconeixement facial, que és un ús particular de la intel·ligència artificial que realment pot causar moltes inquietuds en matèria de drets fonamentals importants. Aquests documents són part d'un projecte més ampli sobre l'ús de la intel·ligència artificial en els estats membres de la UE i els resultats d'aquest treball es publicaran a finals d'aquest any. La intenció darrere d'aquesta investigació és realment poder identificar casos d'ús pràctic en què la IA s'estigui utilitzant en els diferents sectors de l'economia, ja sigui en ocupació, comerç al detall, assegurances, sector públic, específicament en ordre públic, etc. i comprendre realment fins a quin punt els drets fonamentals formen part dels debats quan es desenvolupen i s'implementen eines d'IA.

Finalment, com ja he dit, també vam participar activament en alguns dels processos polítics en curs, tant en l'àmbit de la Comissió com del Consell d'Europa.

La feina que hem fet fins ara realment destaca la gamma completa de drets fonamentals que es poden veure afectats per la intel·ligència artificial i les diferents tecnologies relacionades. Hem sentit a parlar molt en aquests debats sobre protecció de dades i privacitat, i també cada vegada més sobre la no-discriminació i, per descomptat, tots dos drets són extremadament importants en qualsevol debat d'aquest tipus. Però també és molt important que pensem en tota la sèrie de drets que potencialment s'hi poden veure afectats, i realment és difícil pensar en drets que no puguin ser potencialment afectats per la intel·ligència artificial, per la qual cosa això inclou: llibertat d'expressió, llibertat d'informació, llibertat de reunió, però també la igualtat de gènere, els drets de grups i societats particulars, els drets dels infants, els drets de les persones grans i les situacions específiques que puguin afectar aquests grups. Suposo que en el context de les organitzacions d'ombudsman és particularment important pensar en els drets d'accés a la justícia que es veuen afectats en qualsevol debat sobre intel·ligència artificial i també en el dret a una bona administració.

També és important reconèixer que es plantejaran dubtes sobre l'equilibri de drets, i ahir vam veure algunes pistes que anaven en aquesta direcció. Així, per exemple, la qüestió del dret a la privacitat i la protecció de dades i el dret a la no-discriminació. Per tant, si se supervisa un algoritme per veure on discrimina en relació amb grups religiosos particulars, serà necessari recopilar informació sobre l'afiliació religiosa per determinar si aquesta discriminació s'està produint.

Llavors, què significa això per a qualsevol possible regulació futura de la intel·ligència artificial? La nostra investigació suggereix alguns elements que podrien ser els fonaments desitjables per a lleis i polítiques futures. Aquí el primer que he de dir, i crec que el nostre

col·lega de CAHA també va reflexionar sobre això, és que a Europa ja tenim un marc legal sòlid, de manera que hem de reconèixer que no partim de zero, que ja hi ha fortes proteccions. La Carta dels drets fonamentals cobreix molts drets diferents que s'hi poden veure afectats, tal com vam veure en la diapositiva anterior, i, per descomptat, aquests drets són vinculants per a la UE i els estats membres quan apliquen el dret de la Unió. A més, tenim l'RGPD (Reglament general de protecció de dades) i disposem de les directives de qualitat existents. Llavors, per exemple: si la paraula *discriminatori/òria* apareix com a resultat en una eina de contractació que utilitza intel·ligència artificial, això ja seria il·legal segons la legislació de la UE. També és realment important abordar la gamma completa de drets afectats, de manera que hem d'assegurar que qualsevol debat realment intenti rastrejar les implicacions de qualsevol ús de la IA en particular en tota la sèrie completa de drets.

En tercer lloc, crec que és molt important que tinguem casos d'ús concrets sobre com la IA s'està implementant realment en la pràctica, i això es deu d'alguna manera que gran part del debat en aquest moment pot romandre sovint en un nivell bastant teòric, i no tenim necessàriament el coneixement i la informació sobre algunes de les formes pràctiques en què les tecnologies d'IA ja s'estan utilitzant i, en aquest sentit, la sessió d'ahir crec que va ser realment útil en termes de fer-nos una idea d'alguns dels usos que ja estan implementats. I això permet un debat una mica més contextualitzat sobre les implicacions dels drets fonamentals en qualsevol dels usos de la IA. Llavors, per exemple, en el nostre informe recent sobre la tecnologia de reconeixement facial, destaquem que en realitat no hi ha un sol tipus d'aquesta tecnologia, però depenent de com es faci servir les implicacions sobre drets fonamentals canviaran. Llavors, veiem que és molt diferent si la tecnologia de reconeixement facial s'usa amb finalitats de verificació, per exemple, si passem per les portes automàtiques de control de passaports a l'aeroport, és a dir, d'identificació, i un altre exemple de molts serà l'ús en termes de reconeixement facial en viu, per exemple, per motius de seguretat.

En quart lloc, la qüestió de l'alfabetització de les persones implicades i afectades per l'ús de la intel·ligència artificial és particularment important. Sovint, simplement no hi ha prou coneixement sobre la tecnologia, sobre la forma en què es podria usar i sobre les inquietuds particulars que podria plantejar. Però aquesta alfabetització va en els dos sentits; per tant, és qüestió que les persones que estan al costat tècnic estiguin més alfabetitzades també pel que fa a les normes i les inquietuds pels drets fonamentals, i també que les persones que provenen d'una perspectiva de drets fonamentals siguin més capaces de comprendre els elements tecnològics.

Finalment, també és molt important que garantim la cooperació en les diferents àrees i disciplines. Un cop més, crec que aquest és un punt que hem sentit comentar de maneres diferents en aquesta sessió, però realment el fet d'assegurar-nos que s'uneixen aquestes veus diferents és una cosa particularment important, i potser també valgui la pena afegir que, a mesura que aquestes tecnologies s'implementen cada vegada més, escoltar les veus de les persones que poden haver estat afectades negativament per les decisions preses per la IA seria una part certament important d'aquest debat.

I, finalment, només vull acabar pensant en alguns aspectes que serien útils perquè les organitzacions d'ombudsman els considerin com a part d'aquest debat. És interessant saber que el Llibre blanc de la Comissió sí que destaca els riscos dels drets fonamentals com alguns dels reptes més importants que planteja la intel·ligència artificial, i el Llibre blanc també destaca que hi ha riscos molt reals sobre els quals serà més difícil buscar reparació i obtenir compensació per a les víctimes, en comparació amb els damnificats per danys causats per productes i serveis més tradicionals. Llavors, un dels problemes és si les persones són realment conscients que les dades estan sent processades per eines d'intel·ligència artificial. Per descomptat, el remei efectiu correcte requereix que sapigueu realment el que succeeix, i aquest no serà necessàriament el cas. Per tant, pensar en com les persones poden ser més conscients de com es processen aquestes dades és particularment important.

En segon lloc, hem sentit a parlar sobre els reptes que planteja l'opacitat dels algorismes i la presa de decisions automatitzada i això, per descomptat, també es vincula amb qüestions de complexitat i, novament, el Llibre blanc aborda aquests casos amb un cert detall, mirant no només si podem comprendre realment el que està fent l'algoritme, sinó també si podríem, per exemple, accedir a les dades que serien necessàries per poder construir una acusació, en el cas que un algoritme causi un dany.

En tercer lloc, hi ha la qüestió de l'assignació de responsabilitats. Hi ha diferents drets fonamentals que poden sorgir en diferents etapes del procés de disseny, implementació i ús d'algorismes, i és realment important pensar en quins podrien ser aquests riscos i què significarien en última instància per a qui podria ser el responsable dels danys ocasionats, i veiem que en aquest informe adjunt sobre responsabilitat de la Comissió s'aborda aquesta qüestió amb força detall.

I, finalment, hi ha un problema d'escala i de la quantitat de persones que es veuen potencialment afectades per l'ús d'algorismes, i crec que hem sentit una referència sobre això al principi de la sessió d'ahir al matí, d'algú que parlava sobre l'impacte de l'RGPD en els recursos per a organismes que podrien estar gestionant queixes relacionades amb l'RGPD, i atesa l'escala d'aquesta aplicació o l'aplicació potencial de tecnologies d'intel·ligència artificial, serà particularment important tenir els recursos per poder garantir que les persones obtinguin reparació del dany.

Per tant, la meva última petició seria que potser aquest i altres debats que hem tingut puguin aportar alguna cosa en què pensar a les organitzacions d'ombudsman en termes de la seva contribució a la consulta sobre el Llibre blanc de les comissions. Serà realment important obtenir la perspectiva dels òrgans nacionals de drets humans com a part d'aquest debat, perquè crec que tenen una experiència i unes perspectives molt rellevants que realment s'haurien de tenir en compte com a part d'aquest procés. Moltes gràcies.

Geraldine Mattioli

Assessora al Comissari Europeu de Drets Humans (Consell d'Europa)

Moltes gràcies, i gràcies per convidar-nos. És un veritable plaer ser aquí per presentar el treball del Comissari de Drets Humans del Consell d'Europa sobre intel·ligència artificial i drets humans. Per als qui no estiguen gaire familiaritzats amb la institució del Comissari, és un òrgan no judicial independent i imparcial del Consell d'Europa establert per promoure la consciència i el respecte dels drets humans als quaranta-set estats membres del Consell d'Europa. Malgrat el títol de "Comissari", no és membre de la Comissió Europea a la Unió Europea. És important diferenciar entre la Unió Europea i el Consell d'Europa. El nom de l'actual comissària és Dunja Mijatović i va assumir el càrrec l'abril de 2018.

La meva presentació constarà de dues parts. Primer, presentaré àmpliament la Recomanació sobre intel·ligència artificial que el Comissari va publicar l'any passat, que va ser esmentada anteriorment pel Sr. Stojin, i després abordaré els punts de la Recomanació que tracten específicament del possible paper de les institucions d'ombudsman i altres estructures nacionals de drets humans en matèria d'intel·ligència artificial i drets humans.

Des que va començar el seu mandat, la Sra. Mijatović, va identificar la intel·ligència artificial com una de les seves prioritats. La raó d'això va ser que va reconèixer des del principi que la intel·ligència artificial afectarà tots i cadascun dels drets humans que té el mandat de vigilar i protegir en tots els quaranta-set estats membres del Consell d'Europa. Per tant, tenia molt clar que aquest desenvolupament seria fonamental en tot el treball de la seva oficina. L'enfocament de la comissària pot descriure's de la manera següent: eventualment pot ser necessari identificar les llacunes en el marc legal existent i adoptar nous instruments per regular la intel·ligència artificial. Però això portarà temps i, mentrestant, s'està implementant la tecnologia. Per tant, hem de

començar a analitzar les implicacions de la intel·ligència artificial en els drets humans, i hem de començar amb les eines que tenim. No oblidem que la normativa de drets humans, la normativa de discriminació i la normativa de protecció de dades són eines molt poderoses, que ja estan a la nostra disposició i que hem d'aprofitar per garantir que la intel·ligència artificial s'implementi de manera compatible amb els drets humans.

Tenint en compte aquestes consideracions, la comissària va publicar una recomanació basada en les normes existents l'estiu passat. És un document molt breu. N'he dut algunes còpies i us animo que us n'emporteu una. La Recomanació identifica en només deu passos què han de tenir en compte les autoritats nacionals per assegurar que la intel·ligència artificial compleixi amb els drets humans. En altres paraules, la Recomanació descriu un enfocament basat en els drets humans per a l'ús de la intel·ligència artificial.

No esmentaré aquests deu passos ara, però permeteu-me que n'il·lustri alguns de particularment rellevants. Si una autoritat pública està pensant a utilitzar un algoritme similar als exemples que vam escoltar ahir, és a dir, per a prestacions socials, reconeixement facial, pronòstics policials, llavors hi ha una sèrie d'accions que aquestes autoritats haurien de fer abans d'utilitzar-lo. La recomanació suggereix, per exemple, que es desenvolupin avaluacions d'impacte sobre els drets humans. Si és probable que l'ús de l'algoritme tingui un possible impacte en els drets humans dels usuaris, les autoritats han d'avaluar aquest possible impacte per avançat. També han de repetir aquesta avaluació d'impacte una vegada s'estigui utilitzant l'algoritme per assegurar-se que no s'estiguin produint violacions de drets humans.

Un altre pas suggerit en la Recomanació és que les autoritats públiques informin la gent i siguin transparents sobre l'ús dels algoritmes. Actualment, la majoria de les vegades, les persones no saben que són objecte d'una decisió automatitzada, no se'ls diu que no és una persona física la que realment revisa la seva sol·licitud de prestacions socials. Això no hauria de ser així, la transparència és molt important. A més dels seus propis deures, les autoritats públiques també tenen l'obligació d'encoratjar les empreses a implementar les normes en matèria de drets humans quan desenvolupen algoritmes.

Un altre punt esmentat en la Recomanació és la necessitat de supervisió. També vam sentir parlar molt sobre aquesta qüestió durant les conferències d'ahir. Mantenir la supervisió humana sobre els sistemes d'intel·ligència artificial és realment important, i aquesta supervisió ha de dur-se a terme independentment de les empreses que desenvolupen els algoritmes i de les autoritats que els utilitzen. Com cal desenvolupar exactament aquesta supervisió? Aquesta és una gran pregunta que haurem d'abordar. Finalment, la Recomanació també cobreix un punt destacat per la Martha fa un moment: els estats tenen la responsabilitat de desenvolupar l'alfabetització en IA entre la població. En conclusió, la Recomanació de la comissària pretén ser un full de ruta perquè les autoritats nacionals puguin evitar que la intel·ligència artificial es desplegui de forma salvatge i descontrolada.

La Recomanació esmenta en diversos moments el possible paper de les estructures nacionals de drets humans, per la qual cosa ens referim a les institucions d'ombudsman, però també a les institucions nacionals de drets humans, organismes d'igualtat, oficines de protecció de dades...; és a dir, a tots els organismes de l'àmbit nacional que ja tenen el mandat de controlar que es respectin els drets humans. Aquesta és la segona part de la meua intervenció: presentaré les possibles funcions que es preveuen per a les institucions d'ombudsman en relació amb la intel·ligència artificial en la recomanació de la comissària, i estic desitjant escoltar les vostres reaccions i les vostres opinions sobre aquestes idees.

D'entrada, vull subratllar que som conscients que les institucions d'ombudsman poden ser molt diferents, tenir mandats diferents segons els països. Per descomptat, les possibles funcions dependrien de les funcions que duguin a terme aquestes institucions d'ombudsman a cadascun dels diferents països.

Però, en general, m'agradaria convidar-vos a reflexionar sobre el paper de les institucions d'ombudsman en quatre grans categories diferents. En primer lloc, què poden fer en el camp de la prevenció, per prevenir violacions de drets humans per part de la intel·ligència artificial? En segon lloc, què poden fer per augmentar la sensibilització del públic i educar la gent en aquests temes? En tercer lloc, què poden aportar les institucions d'ombudsman en termes de supervisió? I, finalment, què poden fer en termes d'accés a la justícia i reparació? Per a cadascuna d'aquestes quatre categories, veiem dos tipus d'activitats: el que jo anomenaria les *activitats clàssiques* de l'ombudsman, i després algunes coses més noves que cal tenir en compte, específicament en el context de la intel·ligència artificial.

Per *activitats clàssiques* em refereixo a la tasca que les institucions d'ombudsman ja estan fent cada dia en el seu paper de supervisar el respecte pels drets humans, però aplicades a l'ús de la intel·ligència artificial. Per exemple: les institucions d'ombudsman revisen periòdicament els projectes de llei i les polítiques públiques per assegurar-se que compleixen amb els drets humans. Això pot fer-se per a la legislació i les polítiques relatives a la intel·ligència artificial. Ahir sentíem que hi ha una carta de drets digitals a Barcelona. Es va consultar el Síndic? Va revisar ell la carta? Aquesta és una pregunta important. En aquest moment, molts estats estan preparant estratègies nacionals sobre intel·ligència artificial o estan adoptant una nova legislació en relació amb la intel·ligència artificial. És essencial assegurar-se que les institucions d'ombudsman revisin aquests documents per garantir que compleixen els drets humans.

Altres tipus d'activitats "clàssiques" inclouen: llançar una campanya d'informació pública sobre intel·ligència artificial i drets humans, iniciar investigacions sobre denúncies de violacions sistèmiques que puguin estar ocorrent en el context de l'ús de sistemes d'intel·ligència artificial, publicar informes, publicar recomanacions per a les autoritats i després, per descomptat, la tramitació de les queixes individuals quan es presenten. Aquest és el nucli del mandat de les institucions d'ombudsman. Les institucions d'ombudsman també podrien facilitar assistència legal a les víctimes i dur a terme litigis estratègics, ja que és possible que no puguin abordar-se totes i cadascuna de les queixes si n'hi ha milers.

El que acabo de dir pot semblar obvi, però és important reconèixer que la feina de les institucions d'ombudsman segueix sent la mateixa independentment de si la violació dels drets humans la comet un algoritme o una persona. És fonamental reconèixer que la intel·ligència artificial no és un concepte de ciència ficció, sinó que ja és a les nostres societats, ja està succeint. Un primer pas que cal valorar és que les institucions d'ombudsman facin un mapatge dels sistemes d'intel·ligència artificial utilitzats als seus països. Els organismes d'igualtat de Suècia i Alemanya han posat en marxa estudis d'aquest tipus que els permetran comprendre millor els possibles impactes i riscos de la intel·ligència artificial en els drets humans en l'àmbit nacional.

Més enllà d'aquestes activitats "clàssiques", m'agradaria esmentar breument dos rols específics. En primer lloc, les institucions d'ombudsman podrien participar en les avaluacions d'impacte dels sistemes d'intel·ligència artificial en els drets humans que he esmentat al principi. De fet, les institucions d'ombudsman tenen molta experiència rellevant que pot aprofitar-se per ajudar les autoritats a dissenyar aquestes avaluacions d'impacte. Les institucions d'ombudsman poden proporcionar orientació sobre quines preguntes han de fer-se i qui ha de consultar-se. Potser els ombudsman haurien de participar fins i tot a fer les avaluacions d'impacte. Això és una cosa que cal valorar.

L'altra idea es refereix a la qüestió de la supervisió independent. La recomanació de la comissària és deliberadament imprecisa sobre com podria ser un mecanisme d'aquest tipus. Ens fem moltes preguntes sobre qui hauria de ser el responsable d'aquesta supervisió. Un mecanisme de certificació? Una institució pública o una institució privada? Hi han de participar diferents òrgans, en la supervisió, aportant coneixements i perspectives diverses? O hauríem d'acostar-nos a la creació d'un mecanisme de supervisió específic per a la intel·ligència artificial, seguint el model de les oficines de

protecció de dades? Sigui quin sigui el model escollit, el Comissari de Drets Humans creu que és fonamental que el mecanisme de supervisió aporti experiència en drets humans, fins i tot mitjançant la participació de les institucions d'ombudsman.

Vull acabar amb dos punts positius. En primer lloc, la comissària ha comprovat amb molta satisfacció que diverses institucions d'ombudsman ja han començat a treballar en intel·ligència artificial i drets humans. Hi ha institucions d'ombudsman que estan marcant el camí i fent una tasca important. Us puc donar alguns exemples: a França, l'Ombudsman s'ha fet càrrec de diversos casos vinculats a l'ús de la intel·ligència artificial, per exemple, pel que fa a la contractació en universitats o a l'ús d'algoritmes en el sistema judicial; a Finlàndia, l'Ombudsman va guanyar un litigi estratègic sobre la discriminació per edat mitjançant algoritmes; a Eslovènia, l'Ombudsman va abordar un cas constitucional sobre la lectura automàtica de matrícules i la Llei de privacitat. Ahir també vaig tenir una conversa amb Catherine de Bruecker, la qual em va donar diversos exemples de Bèlgica en què l'Ombudsman està analitzant processos de presa de decisions automatitzades. Estan succeint moltes coses i aquesta conferència és una altra demostració que les institucions d'ombudsman estan assumint el desafiament. Això és molt positiu.

El segon punt positiu que vull esmentar és que no hem d'oblidar que la intel·ligència artificial es pot utilitzar per promoure els drets humans. La intel·ligència artificial ens ajudarà a documentar i demostrar casos de discriminació, per exemple, d'una manera que no vam poder provar abans, mitjançant l'ús de *big data*. La Comissió Nacional de Drets Humans d'Irlanda, per exemple, va poder utilitzar un algoritme per analitzar el discurs d'odi en línia a Twitter en el context específic d'Irlanda i comprendre millor d'on venia. Hi ha moltes possibilitats que la intel·ligència artificial ajudi a la tasca de les institucions d'ombudsman sobre els drets humans, i cal tenir-ho en compte.

La comissària té previst continuar la seva tasca sobre intel·ligència artificial i drets humans. El següent pas probablement serà intentar analitzar els usos específics de la intel·ligència artificial en diversos països del Consell d'Europa. La comissària ja va examinar l'impacte de la intel·ligència artificial en els drets humans de la gent gran a Estònia en un informe publicat el setembre de 2018. Els membres de l'oficina de la comissària també van dur a terme missions d'investigació a Suècia per estudiar l'ús de la intel·ligència artificial en el benestar social; i a Sèrbia, per examinar la implantació de la tecnologia de reconeixement facial. Vull acabar animant tots els assistents que es posin en contacte amb l'oficina de la comissària si consideren que l'ús de la intel·ligència artificial al seu país planteja qüestions específiques de drets humans. Moltes gràcies.

Peter Bonnor

Jurista principal de l'Ombudsman Europeu

Gràcies per donar-me la paraula i gràcies per la invitació.

Unes paraules sobre l'Ombudsman Europeu. L'Ombudsman Europeu és el defensor de l'Administració pública de la Unió Europea. I què és això? Són les institucions que coneixem de Brussel·les, la Comissió Europea, el Parlament Europeu, el Consell i el gran nombre d'agències que tenim a tot Europa, especialitzades en diferents camps, com ara els drets fonamentals, o que participen en la regulació de la salut i la seguretat, la seguretat dels productes químics, la seguretat dels productes mèdics, etc. L'Administració pública europea inclou entre cinquanta i seixanta mil persones. Es tracta d'un nombre molt reduït de persones per a una administració pública de la Unió Europea. És el nombre de persones que més o menys correspon al nombre de funcionaris públics d'una ciutat europea de mida mitjana. El motiu d'això, per descomptat, és que no tenim una administració directa, no tenim agents de policia i no tenim un hospital de la Comissió Europea ni res per l'estil.

L'Administració pública de la Unió Europea és un mecanisme regulador. Això és molt important en aquest context.

L'oficina de l'Ombudsman Europeu dona feina a una vuitantena de persones. Som entre vint-i-cinc i trenta persones que participem en el treball d'investigació diari. Fins ara no hem rebut queixes relacionades amb la intel·ligència artificial, no actuem d'ofici en aquest camp i crec que per ara soc l'empleat que té més coneixements sobre intel·ligència artificial. Així doncs, de quina manera som rellevants?

Voldria destacar dos camps en què l'Ombudsman Europeu pot ser força rellevant en relació amb la intel·ligència artificial, i probablement ho serà.

La Unió Europea, tot i que és relativament petita, distribueix grans sumes de diners. Sens dubte, invertirà molts diners en aquest camp de la intel·ligència artificial. Mitjançant aquesta activitat, hi haurà convocatòries de concursos, convocatòries de propostes, contractes, es crearan consorcis amb desenes de socis ubicats per tota Europa. L'experiència ens demostra que hi haurà problemes, hi haurà malentesos i hi haurà situacions en què la Comissió Europea digui a algú: "No, no et podem pagar perquè no ens has lliurat els productes", i llavors tindrem una crisi. I aquestes són les situacions en què l'Ombudsman Europeu pot intervenir i treballar com una mena d'obrellaunes i ajudar mitjançant avaluacions basades no només en la legislació, sinó també en la justícia i la bona administració. Per tant, això és rellevant des d'un punt de vista pràctic però encara relativament limitat.

Des d'un altre punt de vista, cal recordar que l'actual ombudsman europeu, que va ser reelegit per cinc anys el desembre de l'any passat, ha adoptat un enfocament que té molt en compte el fet que l'Administració pública europea és una maquinària reguladora. L'Ombudsman Europeu no pot examinar i estudiar la legislació com a tal, però pot aproximar-se bastant als processos reguladors. Hem sentit a parlar d'ètica i la intel·ligència artificial. Estem examinant l'ètica en la regulació a escala de la UE. Podem mirar si hi ha situacions de conflicte d'interessos en els processos que condueixen a les normes reguladores. Podem mirar fins i tot si hi ha un equilibri adequat, per exemple, en els grups d'experts, els tallers, les conferències, aquest tipus de coses. Si algú ens diu: "Escolteu, la Comissió Europea només convida la indústria a participar en aquesta o aquella activitat", llavors podem fer-hi una ullada (per il·lustrar-ho, crec que el primer panell de conferenciants que vam veure ahir va ser força equilibrat, de manera que va ser un bon exemple d'un grup equilibrat sobre la qüestió.) L'Ombudsman Europeu ha dut a terme aquestes avaluacions en diversos camps, inclosa l'agricultura, i ha comprovat qui són els membres d'aquests grups d'experts que s'incorporaran en el procés d'elaboració de polítiques que finalment donaran lloc a legislació europea.

A la UE, tenim dos comissaris molt poderosos i competents a Brussel·les encarregats d'aquesta àrea de la intel·ligència artificial, la digitalització i tota l'estratègia digital, i les coses aviat podran avançar molt ràpidament.

No puc deixar de destacar, com han fet els meus col·legues d'aquest panell, que, si us plau, aprofiteu l'oportunitat actual per introduir-vos en el procés d'elaboració de polítiques responnent o contribuint a la consulta pública organitzada pel Comissari Europeu. Només heu de buscar a Google "European Union public consultations" ("consultes públiques de la Unió Europea"), i trobareu una llista de consultes públiques obertes sobre intel·ligència artificial i estratègia de dades. Requereix una mica de feina llegir els dos llibres blancs, sobre intel·ligència artificial i estratègia de dades, però contribuir en aquests processos és molt important. No és un exercici consistent a marcar caselles, no és un simple tràmit, és un procés que s'està prenent molt seriosament. I forma part de la cultura de Brussel·les, de la cultura de la Unió Europea, fer consultes públiques i examinar detingudament les aportacions. Fins i tot ho fem com a oficina de l'Ombudsman. Hem fet diverses consultes públiques sobre diferents qüestions i ha estat interessant veure que obtenim aportacions de gairebé totes les parts: de persones individuals, d'investigadors i fins i tot de governs o institucions de la UE. Per tant, és un exercici molt ric que influirà en allò que integrarà la legislació. Fins i tot alguns de vosaltres us podeu reunir ad hoc i fer una contribució com a grup.

A continuació, voldria dir un parell de coses, que no són més que la meua opinió, per enllaçar amb la sessió següent: quins tipus d'eines podem recuperar, els participants?

Crec que la presentació del grup que vam veure ahir la va encertar de ple en una de les diapositives que contenia una referència a la contractació pública. És cert que tenim un mandat general com a ombudsman, que podem investigar moltes coses, però cal que no oblidem els límits que també tenim: no podem encarregar a l'Administració pública que faci alguna cosa, només podem recomanar-ho, oi? Per tant, és molt important que estiguem presents en aquest camp en una fase molt primerenca, a fi d'influir en les coses. No és més que una norma bàsica de negociació.

A l'oficina de l'Ombudsman Europeu no tenim experts en intel·ligència artificial, i crec que la majoria d'administracions públiques europees tampoc no tindran experts reals en intel·ligència artificial. Seran contractats per grans empreses. Els salaris són millors, l'oportunitat de ser creatius serà més gran, etc. Per tant, la qüestió serà haver de seguir els diners. On aniran els diners? Ara mateix l'Ombudsman té l'oportunitat d'asseure's, parlar amb l'Administració pública i dir: "M'agradaria saber què teniu previst fer" i "vull veure què incloueu a les especificacions funcionals i tècniques de les licitacions que publicareu". I després, segons els estàndards dels quals hem sentit parlar per exemple durant aquesta conferència, examinar les licitacions públiques i els punts d'avaluació que es donen a les empreses que després responen a aquestes licitacions públiques. El nivell de transparència de la intel·ligència artificial ha de ser bo? Les proves relacionades amb la intel·ligència artificial han de tenir un nivell molt alt? Per tant, fins a cert punt es tracta de l'antiga norma: seguïu el rastre dels diners quan vulgueu fer investigacions eficients.

Tornant a la consulta pública que actualment s'està duent a terme amb la Comissió Europea, aquí teniu una o dues idees sobre el que es podria plantejar proposar.

Tenim notícies de moltes coses que van sortir malament en algun lloc i que la gent de dins en tenia coneixement. Per exemple, l'escàndol de les emissions dièsel, o la crisi actual de Boeing, en què l'empresa coneixia les coses que no anaven bé. No he vist en els llibres blancs que s'hagin proposat normes estrictes per als denunciants, per exemple els programadors. Si hi ha algun programador en algun lloc que sap que alguna cosa no funciona correctament, aquesta persona hauria de poder denunciar-ho de manera segura a l'organisme adequat i mitjançant els mecanismes de protecció escaients.

Finalment, per acabar, només puc insistir en el que ja s'ha dit, que la intel·ligència artificial és simplement una cosa que ja està passant, i que molta de la normativa vigent ja s'aplica a aquest camp. Així, tornant a la caixa d'eines, quina acció concreta podríem dur a terme? Una d'elles és assegurar-nos que interioritzem la comprovació periòdica de la igualtat d'oportunitats, la protecció de dades, etc., i que tenim les xarxes adequades a les quals podem accedir immediatament. Per exemple, com a investigador de l'Ombudsman, voldria poder accedir a una xarxa d'experts en intel·ligència artificial, però si la setmana vinent rebo una queixa relacionada amb la intel·ligència artificial, voldria no haver de passar setmanes o mesos fent un concurs públic per ser capaç d'aprofitar els coneixements rellevants.

M'adono que a Barcelona sou molt privilegiats, esteu envoltats de persones que saben de què parlu. Jo, com a investigador de l'oficina de l'Ombudsman Europeu, en aquest moment no sabia a qui trucar si tinc un cas sobre intel·ligència artificial. Per tant, la creació de xarxes a les quals poder recórrer immediatament pot ser una part important de la nostra caixa d'eines.

Moltes gràcies.

Intel·ligència artificial i drets humans

reptes, rols i eines de les institucions d'ombudsman

Dimarts, 3 de març de 2020

Reptes, rols i eines de les institucions d'ombudsman

Moderador: Joan Manuel del Pozo, síndic de la Universitat de Girona

Robert Behrens, ombudsman del Servei de Salut i Parlamentari del Regne Unit

Elisabeth Rynning, ombudsman parlamentària de Suècia

Quirine Eijkman, comissària i vicepresidenta de l'Institut Neerlandès de Drets Humans

Eric Houtman, ombudsman federal belga de l'energia i membre de la NEON (Xarxa Europea d'Ombudsman de l'Energia)

Joan Manuel del Pozo

Síndic de la Universitat de Girona

Molt bona tarda, senyores i senyors. Comencem la taula prevista per a aquesta hora, amb un petit retard pel qual demanem excuses, amb el tema de "Rols, reptes i oportunitats per a les institucions d'ombudsman".

Primer de tot, una presentació de la taula, en què registrem l'absència involuntària de la Sra. Madise, ombudsman d'Estònia, que serà substituïda per un vídeo inicial de la senyora Elisabet Rynning, ombudsman parlamentària sueca. I de les persones presents tenim: el Sr. Rob Behrens, ombudsman parlamentari i del sistema sanitari del Regne Unit; la Sra. Quirine Eijkman, comissària i vicepresidenta de l'Institut Neerlandès de Drets Humans, i el Sr. Eric Houtman, ombudsman federal belga de l'Energia i membre de NEON.

Agraeixo, en primer lloc, al Síndic de Greuges de Catalunya l'organització de la jornada i la invitació de la meua persona per participar-hi, i la presència de tots vostès, que dona sentit òbviament a la reflexió col·lectiva que tots estem fent.

En aquesta taula ens toca, com si diguéssim, condensar orientativament cap a la responsabilitat dels ombudsman tot allò que s'ha estat reflexionant amb gran riquesa de perspectives i amb molt bona aportació d'informacions i de criteris, i intentar parlar nosaltres de quins són els rols o papers que hem de fer, quins són els reptes que tenim al davant i quines són les oportunitats que ens convé aprofitar com a ombudsman.

Sobre això voldria fer una petita paraula introductòria, deixant clar que com s'ha dit en alguna intervenció anterior, plenament d'acord amb ella, crec que els ombudsman mai no hem de renunciar a allò que està en la nostra naturalesa com a tals. Primer de tot, nosaltres no som ni jutges, ni legisladors, ni governants. Som defensors dels ciutadans i ciutadanes i aquesta és una funció de contrapès al pes enorme que tenen les administracions i els poders clàssics que acabem d'esmentar. I en aquesta defensa de drets tenim un paper democràticament principal i essencial del qual ens sentim orgullosament responsables.

És veritat que, a part de totes les possibles violacions de drets humans que la història humana per desgràcia ens ha ensenyat, ara està emergint un altre problema que pot afectar, com s'ha anat veient aquests dos dies, els drets humans de les noves societats. En aquest sentit, crec que nosaltres hem de mantenir molt clar el paper de restitució de qualsevol violació personal de drets humans. Aquesta és la nostra primera i fonamental funció. El ciutadà que es dirigeix a nosaltres i diu: jo em sento alterat, vulnerat, violat en algun dret, el que sigui, de privacitat, de llibertat, d'igualtat per culpa, en aquest cas, d'un ús indegut de la intel·ligència artificial, és la nostra primera missió essencial.

La segona és anticipar-nos amb iniciatives dites d'ofici o d'iniciativa pròpia per tal que, conscients com som de l'evolució de la societat, i en aquest cas de l'evolució ja molt forta, ja present entre nosaltres, de la intel·ligència artificial, puguem fer tasques d'anticipació a la possible vulneració d'aquests drets. De fet, és el que estem fent aquests dos dies. I jo gosaria dir que derivadament d'aquesta capacitat d'iniciativa pròpia que tenim, hauríem d'assumir-ne una altra, que també ha aparegut en alguna intervenció anterior, que jo anomenaria de pedagogia social del problema. Aquesta pedagogia social que consisteix a fer créixer la consciència del risc de la intel·ligència artificial en relació amb els drets de les persones. Tenim una situació en la qual la intel·ligència artificial no presenta l'aspecte d'una bèstia abominable que es veu venir de lluny, sinó que la intel·ligència artificial es presenta amb una subtilitat que és difícil de percebre per part de molts ciutadans, de tal manera que probablement, com a institucions d'ombudsman, hauríem de contribuir a explicar a la societat que té una ètica, que té una necessitat de protecció de la seva dignitat, de la seva llibertat i de la seva igualtat que podria veure's afectada per determinats usos de la intel·ligència artificial. I voldria remarcar que en la preocupació ètica que estem tenint aquests dies hauríem de ser conscients que l'ètica no és imposició de ningú.

En altres èpoques històriques l'ètica, o la moral, era imposició d'una autoritat, o religiosa o política. No, avui l'ètica és una necessitat de les nostres persones i de les nostres societats. No és un luxe, no és tampoc una moda. De fet, des dels anys seixanta que va aparèixer una de les ètiques més conegudes en la seva aplicació a la vida humana, que és la bioètica, han anat apareixent ètiques aplicades que responen a nous reptes socials. I així, per exemple, sentim a parlar de l'ètica de la comunicació, sentim a parlar fins i tot de l'ètica dels negocis. I em sembla bé que es faci. Però el fet que sigui una moda no hauria de fer perdre de vista que sota de la moda hi ha un reclam ètic que ve amb la història de la humanitat mateixa. Hi ha uns valors que són consubstancials a la condició humana, que són també anteriors a la nostra època i que reclamen, òbviament, la nostra atenció permanent.

Voldria fer una brevíssima lectura de dos pensadors que estan avui dia interpretant bé la nostra situació, que ens diuen alguna cosa tan interessant per reflexionar i que proposo que els nostres ponents també puguin opinar i dir el seu criteri. El pensador israelià Yuval Noah Harari, molt conegut per *Homo Deus* o per *21 lliçons per al segle XXI*, diu concretament en aquest últim llibre una frase molt curta però que és tot un repte per a la nostra reflexió: "Quan la intel·ligència artificial decideixi més i millor que nosaltres, els nostres conceptes de la humanitat i de la vida hauran de canviar." O sigui, ens està dient que l'autocomprensió de la condició humana està en procés de canvi. Algú podria, si és pessimista, pensar que de canvi a pitjor, té el dret de fer-ho. Altres, potser més optimistes, poden pensar que de canvi a millor.

En qualsevol cas, hem d'estar atents al canvi, perquè el que sembla que com a ombudsman no hauríem d'acceptar és que aquest canvi fos a costa d'alguns dels drets que considerem essencials. I per dir les paraules, diguem-ho així, bàsiques: dignitat, llibertat, igualtat. Si aquest canvi s'ha de fer a costa d'aquests valors, òbviament nosaltres hauríem de presentar alguna resistència.

L'altre pensador és Byung-Chul Han, d'origen Coreà però afincat a Alemanya, i es pot dir pràcticament que és un pensador alemany, d'inspiració filosòfica alemanya, amb una cita un pèl més llarga però que crec que val la pena escoltar amb atenció. Diu: "El poder intel·ligent, amable, no opera frontalment contra la voluntat dels subjectes sotmesos, sinó que dirigeix aquesta voluntat al seu propi favor. És més afirmatiu que negador, més seductor que repressor, s'esforça a generar emocions positives i a explotar-les, sedueix en comptes de prohibir, no s'enfronta al subjecte, li dona facilitats. La present crisi de llibertat consisteix en el fet que estem davant d'una tècnica de poder que no nega o sotmet la llibertat, sinó que l'explota". És a dir, estem davant, com deia abans, no d'una bèstia horrorosa sinó d'una essència subtil, que entra en les nostres ments de manera que gairebé ni ens n'adonem. D'aquí el reclam que jo feia abans d'una certa pedagogia social, de fer emergir la consciència del repte que tothom té a davant, que és que davant de determinats usos indeguts de la intel·ligència artificial no quedem adormits, no quedem desorientats, no quedem despistats.

Bé, davant d'això crec que els reptes per a les institucions d'ombudsman han d'anar en moltes direccions. Jo en citaria alguna de molt important que també ha sortit en intervencions anteriors.

El primer de tot, l'atenció, tenir atenció com a ombudsman als biaixos de legislació i de codis ètics que van apareixent. El biaix és un risc, i hem d'estar atents que en la construcció d'algoritmes i en la construcció de programacions no es produeixin biaixos que acabin deformant l'ús humà de la intel·ligència artificial.

I, en segon lloc, el repte és participar com a ombudsman en processos de reflexió social entorn del bon ús de la intel·ligència artificial. Podríem dir que aquesta trobada és un exemple d'això, és un exemple de voluntat d'anticipació i de procés de participació en processos de reflexió ètica.

I finalment, i amb això acabo, crec que la gran oportunitat que tenim, puix que s'ha plantejat aquesta trobada com una relació de la intel·ligència artificial amb els drets humans, tenim l'oportunitat d'aprofitar els drets humans com un marc de referència que podria considerar-se una mena de codi d'ètica bàsica mundial. Ja hi ha hagut algun pensador que s'hi ha referit, i penso que és útil considerar que, davant del perill d'un cert relativisme ètic, que qualsevol ètica val, i una ètica pot ser canviada per una altra ètica (com deia l'humorista: "Jo tinc uns principis, però si no li agraden, aquí en tinc uns altres"), hem de tenir els drets humans com una mena de reclam ètic, de caràcter bàsic per a tota la humanitat.

I fins aquí, si em permeteu, la presentació d'aquest modest moderador, i donem pas ara a les intervencions dels nostres ponents. En primer lloc, la del Sr. Rob Behrens, ombudsman parlamentari del sistema sanitari del Regne Unit.

Robert Behrens

Ombudsman del Servei de Salut i Parlamentari del Regne Unit

Espero que pugueu sentir-me, no soc un holograma. I aquesta sessió en particular serà participativa, de manera que espereu a ser cridats a intervenir durant els transcurts del que explicaré. Vull començar agraint a Rafael Ribó i el seu equip per convidar-me a contribuir en aquest seminari. Ha estat una veritable experiència d'aprenentatge.

He vingut aquí amb un discurs preparat i diapositives associades. He escoltat amb molta atenció durant els dos dies del seminari i, en conseqüència, he trencat tant el discurs com les diapositives, per a alleugeriment vostre, perquè haguéssiu tornat a sentir el que ja us han explicat, i això seria avorrit i una pèrdua de temps. Així doncs, vaig escriure una presentació més curta que reflecteix el que he sentit i enfocada en les implicacions de la intel·ligència artificial en la meua oficina d'ombudsman i en d'altres, en el Servei Nacional de Salut al Regne Unit i en els usuaris de serveis i professionals de la salut. I acabaré amb una sèrie de breus conclusions emergents.

Vull començar amb una història que tracta sobre la primera ciutat industrial del món. Llavors, qui pot dir quina és la primera ciutat industrial del món? Manchester, correcte. Quanta gent d'aquí ha estat a Manchester? Excel·lent. Sabia que estava davant una audiència sofisticada i culta. Més de la meitat de vosaltres heu estat a Manchester, on hi ha la seu de la meua oficina. I la història que vull contar-vos és sobre el que va succeir el 1830 quan l'empresari del sector privat, George Stephenson, va inaugurar el primer ferrocarril interurbà públic del món. Es deia ferrocarril de Liverpool a Manchester. L'estació de Manchester ha estat restaurada recentment des de 1830 i bé mereix una visita quan vingueu a la meua oficina. La màquina de vapor d'aquesta línia es deia The Rocket (El coet) i tenia una velocitat mitjana, que era la més ràpida del món, de dotze milles per hora i una velocitat màxima de trenta milles per hora. Per celebrar la inauguració de la línia, van acudir alguns dignataris, entre ells el cèlebre i prometedor ministre del Govern, William Huskisson. Com que el projecte era nou, no hi havia barreres de seguretat i ningú estava acostumat a les altes velocitats del nou tren. Molt tràgicament, Huskisson, amb sobrepès i recuperant-se d'una greu malaltia, no va poder moure's prou ràpid i fou atropellat i mort pel tren que s'aproximava. Després de la seva mort, hi van haver predictibles crides en els diaris per demanar la prohibició dels ferrocarrils, la tornada de les diligències impulsades per cavalls i la regulació de la seguretat ferroviària.

Suggereixo que aquesta és una història d'intel·ligència artificial al Regne Unit i a la resta d'Europa. El tren, el coet, és una metàfora de la intel·ligència artificial. George Stephenson és una representació de les empreses que promouen la intel·ligència artificial, com vam veure ahir, que ens asseguren que tot és brillant. I per a mi, el punt més important és que el polític amb sobrepès i inepte que va ser atropellat és una metàfora de l'ombudsman en aquest moment, l'ombudsman el 2020. A més, després de la tragèdia cap de les parts va parlar amb les altres sobre el que havia succeït. Simplement van fer suposicions.

En la meua opinió, i havent estat assegut aquí durant dos dies, aquesta és una crisi existencial per als ombudsman. No soc un romàntic. No podem retrocedir negant la realitat de la intel·ligència artificial. Estem captivats, encantats pel que la intel·ligència artificial pot fer per nosaltres, igual que Huskisson sabia que una màquina de vapor podria revolucionar el transport. El vell món mor i no tornarà. La intel·ligència artificial no desapareixerà. I té grans avantatges per als tres sectors de què ara vull parlar breument.

La meua oficina, l'Ombudsman del Regne Unit, té un pressupost d'aproximadament 30 milions de lliures l'any. Tenim 125.000 consultes telefòniques ciutadanes sobre errades en el servei, rebem 30.000 queixes l'any i per telèfon cursem 70.000 queixes. No podem tractar-les directament, perquè la legislació exigeix que siguin ateses primer per

l'organisme que està a primera línia. El 80% de les queixes tenen a veure amb el nostre Servei Nacional de Salut, i morts evitables i errors greus en el servei de salut. El mandat també ens permet analitzar el judici clínic dels metges en el servei de salut, per la qual cosa encarreguem l'assessorament independent de metges i infermers.

La intel·ligència artificial té un enorme potencial per ajudar-nos, com a ombudsman al Regne Unit, en termes de procés, en termes de conversa, en termes d'arribar a dictàmens sobre decisions clíniques i si aquestes són apropiades. Però la veritat és que al Regne Unit encara no ens sentim còmodes amb el seu ús.

Pel que fa a l'evolució de l'oficina de l'Ombudsman, ens hem concentrat durant els últims tres anys a obtenir els beneficis no només d'actualitzar el nostre sistema informàtic, sinó també de mantenir converses presencials amb les persones que presenten queixes després d'anys d'investigacions en paper. Així, la gent ens ha dit "no escoltes correctament perquè no tens una conversa cara a cara".

Un gran problema per a totes les institucions d'ombudsman és que els usuaris dels seus serveis no es beneficiaran necessàriament de la introducció de la intel·ligència artificial. Aquest és particularment el cas, i ja ho hem sentit, de les persones més vulnerables, les persones amb problemes de salut mental i la gent gran, que tenen menys probabilitats d'utilitzar el nostre servei perquè ja els demanem que utilitzin ordinadors quan ens presenten una queixa. El repte per a tots nosaltres és dissenyar un esquema d'ombudsman que compregui aquells que no se senten còmodes amb els ordinadors i la tecnologia de la informació, i si poguéssim fer això, amb la quantitat de temps i recursos que podríem estalviar, i implicar les persones, podria ser revolucionari per al nostre servei i per a la nostra eficiència.

Hem pensat sobre alguns exemples potencials reals. Agafem el cas d'un col·lega, el professor Stephen Perkins, que és director de la Junta del Govern Local i ombudsman d'Atenció Social al Regne Unit. Fa dues setmanes, va ser entrevistat per un robot que estava programat per fer preguntes durant mitja hora i respondre a les reaccions de Perkins d'una manera en què els que en saben diuen que podríem utilitzar amb els usuaris del servei que acudeixen a nosaltres o que ens truquen. Les preguntes que sorgeixen són: això desanimaria la gent? Quin valor tindria i com sabríem si ha sortit malament? Tanmateix, el que sí que sabem és que, en general, les capacitats de parla i llenguatge de la intel·ligència artificial, inclosa la traducció, acceleraran la nostra comprensió del que diuen les persones que presenten queixes i ens ajudaran a prestar un servei més eficient.

També podríem utilitzar el reconeixement de l'escriptura a mà per convertir el contingut de les lletres escrites a mà i eliminar la necessitat d'introduir dades manualment. I hi ha moltes possibilitats en la intel·ligència artificial que no hem de descartar simplement perquè siguin estranyes i siguin noves.

En segon lloc, em refereixo als beneficis per al nostre sistema de salut, i hem sentit algunes declaracions audaces del sector privat sobre això. Una mica d'això és cert, i el punt destacable és que l'estatu quo que tenim en aquest moment no és ideal i necessita una reforma dràstica; no hem de ser romàntics sobre això. La gestió de dades i la mala gestió al Servei Nacional de Salut Britànic són notòriament vacil·lants. M'ocupo tota l'estona de documents que es perden abans que es dugui a terme una investigació quan han mort persones. Això inclou documents que desapareixen i documents que són falsificats pels metges després de la mort d'un pacient per encobrir una errada del servei i la mort evitable. El present no és ideal, i no hem de tenir por d'invertir en intel·ligència artificial si es vol fer front a aquesta situació.

Gran Bretanya gasta una mitjana de 1.800 milions de lliures esterlines en inversió en intel·ligència artificial, però no estem a l'avantguarda dels països europeus per trobar una solució a això. Sabem per la investigació sobre llibertat d'informació que es va fer en els últims anys que potser només la meitat dels hospitals al Regne Unit han invertit realment

en intel·ligència artificial fins ara. I s'han mostrat molt reticents a unir-se al Govern pensant que aquest serà el camí a seguir. Però el Govern té raó d'entendre que més d'un milió de pacients cada trenta-sis hores fan transaccions amb el servei de salut, i com més puguem utilitzar la intel·ligència artificial per donar-li sentit a això, millor.

A més, hi ha alguns avenços clínics molt importants que salven vides de persones com a resultat de la intel·ligència artificial. L'escaneig de l'interior de l'ull per a pacients a Cambridge condueix a una cura del càncer de pròstata al Regne Unit més ràpidament que abans. Les tomografies computeritzades de pacients amb sospita de malaltia coronària s'utilitzen d'una manera molt millor que les angiogrames. Hi ha moltes innovacions amb Moorfields Eye Hospital a Londres, per exemple, i amb metges de capçalera i infermers en exercici a Londres, que salven vides concretament com a resultat de la intel·ligència artificial. Però hi ha un problema, i Elisabeth Rynning s'hi ha referit, i és que els usuaris del servei han d'estar convençuts. Llavors, ho estariau si us convidessin a una cirurgia major realitzada per un robot? Estariau al costat dels nigerians si us dic que el 69% d'ells estarien a favor de sotmetre's a la cirurgia, o estariau amb els tímids britànics en què només un 27% probablement acceptaria la cirurgia robòtica? Quants de vosaltres us posicionaríeu al costat nigerià? Dos, tres, quatre. Quants del costat britànic? Vinga, ja ha passat la votació del Brexit, de manera que no passa res, podeu dir el que penseu. D'acord, més persones es mostrarien escèptiques. Aquesta va ser una enquesta realitzada per YouGov a 12.000 persones. No s'està fent prou per convèncer la ciutadania que l'ús de la intel·ligència artificial és segur, creïble i eficient i, per tant, la confiança pública, que és el tema clau en tots els serveis públics, no és tan sòlida com hauria de ser. També hi ha la qüestió que els metges compreguin quin valor té i com han de ser formats per estar d'acord amb això, i això no succeeix en aquest moment.

Vull acabar parlant de quatre temes emergents molt ràpidament:

En primer lloc, la confiança pública és el tema clau per al segle XXI en termes de prestació de serveis públics. Si l'ús de la intel·ligència artificial ha de socavar la confiança del públic, tenim problemes més seriosos que abans, en què la confiança del públic ja s'estava erosionant.

En segon lloc, és necessari que hi hagi una regulació. En aquest moment tenim, segons ens van dir, potser entre 84 i 85 marcs reguladors a Europa. I això simplement no és creïble, particularment quan hi ha un sector privat que vol guanyar diners amb el que està fent, en comptes de promoure l'interès públic. El problema per al servei públic és que no és tan coneixedor de la intel·ligència artificial com el sector privat. Per tant, ha d'actuar en conjunt per gestionar les conviccions del sector privat. Altrament, seran superiors al servei públic. No podem reproduir la situació de Groucho Marx dient: "Aquests són els meus principis i si no t'agraden en tinc d'altres". L'interès públic ha de defensar la confiança pública en aquest cas.

En tercer lloc, hi ha la suposició en el món acadèmic que els ombudsmen que únicament es concentren a ser ombudsmen són més efectius que aquells a qui s'assignen altres funcions, pel que fa a ser un organisme normatiu o també a ser un comissionat de la informació, cosa que ocorre en molts països. En la meua opinió, això es capgira amb la intel·ligència artificial. Són les persones, els meus col·legues, els qui tenen la responsabilitat de la protecció de dades i dels drets humans, els que van per davant en el món dels ombudsmen, més que aquells de nosaltres que simplement ens ocupem de les queixes ciutadanes.

I l'últim punt és aquest, i és molt important: tot el tema de la intel·ligència artificial subratlla la importància que els ombudsmen siguin professionals en el que fan. Com han assenyalat Peter Bonnor i Rafael Ribó, llevat d'honroses excepcions, els ombudsmen no estan prou preparats o capacitats per treballar amb la intel·ligència artificial en aquest moment. Si no abordem això, acabarem com William Huskisson, morts a la via del tren. Moltes gràcies.

Elisabeth Rynning

Ombudsman parlamentària de Suècia

Senyores i senyors, benvolguts col·legues,

En primer lloc, permeteu-me que us expressi com lamento no poder reunir-me avui amb vosaltres a Barcelona. Tot i això, estic molt agraïda per aquesta oportunitat digital de compartir alguns dels meus pensaments. Atès que òbviament no soc amb vosaltres en temps real, he de disculpar-me per endavant si repeteixo algun punt que ja s'ha esmentat en els debats d'avui.

El meu nom és Elisabeth Rynning, soc ombudsman parlamentària de Suècia, i m'han demanat que aporti algunes reflexions des de la perspectiva dels defensors suecs.

El Govern suec, com molts ciutadans suecs, s'enorgulleix d'estar molt a favor de la innovació i de la tecnologia. Ja en la dècada de 1990 es van introduir algunes decisions automatitzades en el sector públic en els camps dels impostos i de la seguretat social. Actualment tenim una disposició general en la llei de procediment administratiu que permet decisions automatitzades, però no especifica garanties addicionals. S'ha qüestionat si això realment satisfà els requisits del reglament de protecció de dades.

Com ja deveu haver comentat, l'ús de la intel·ligència artificial pot tenir efectivament molts efectes positius, com ara rendibilitat, accessibilitat als serveis públics, coherència, assegurar la qualitat de les decisions, etc.

Fa dos anys, el nostre Govern es va proposar que Suècia fos líder mundial en l'aprofitament de les possibilitats que ofereix la intel·ligència artificial, i va crear l'Agència Sueca per al Govern Digital. Des del setembre de 2018 aquesta agència té la funció de coordinar i donar suport a la digitalització de l'Administració pública. L'Agència ara està mapant i analitzant les aplicacions actuals de la intel·ligència artificial, i investigant les possibilitats futures. Els aspectes importants que cal observar inclouen, òbviament, qüestions tant ètiques com de legalitat, de protecció, i qüestions relacionades amb la transparència, la responsabilitat i la rendició de comptes.

Moltes agències de supervisió sueques i altres organismes també estan estudiant diferents aspectes sobre l'ús de la intel·ligència artificial. L'Oficina Nacional d'Auditoria està investigant l'eficiència de l'ús de les agències estatals en la presa de decisions automatitzades sense infringir l'estat de dret, l'Agència de Protecció de Dades està valorant diferents implicacions en dades personals i la Junta Nacional de Salut i Benestar està estudiant l'ús de la intel·ligència artificial en el sector de la salut, només per esmentar-ne algunes.

Per tant, és just dir que l'Administració pública sueca s'està movent, tot i que encara amb força cautela, pel que fa a la intel·ligència artificial. Però, què en pensen els ciutadans suecs? Un estudi de 2018 mostra que el 80% de les persones enquestades en aquell moment desconeixia que algunes decisions en el sector públic es prenen per processos automatitzats. La majoria creia que aquestes decisions eren, d'una banda, menys esbiaixades, però de l'altra, menys adaptades a les circumstàncies individuals, menys transparents i no més precises que les decisions preses per humans. Suposo que el recent desenvolupament en aquesta àrea i la major consciència pública sobre la intel·ligència artificial en general poden haver afectat l'opinió pública des de llavors. Dos anys és força temps quan es tracta d'intel·ligència artificial.

Passaré a descriure alguns exemples de problemes, o reptes, si se'ls vol anomenar així, a què ha donat lloc el desenvolupament de la intel·ligència artificial en l'Administració pública a Suècia.

Pel que fa a la protecció de dades personals, l'Agència de Protecció de Dades va portar, ja fa quinze anys, un cas davant el Tribunal Suprem Administratiu, relatiu a l'ús de la identificació biomètrica en una escola. El Tribunal va determinar que sol·licitar

un escaneig d'empremtes dactilars perquè els alumnes tinguessin accés a un plat al menjador escolar no era legal sense el consentiment exprés del processament de dades. Avui dia, l'Agència té un cas pendent sobre el tema dels controls d'assistència mitjançant escàners de reconeixement facial, que l'Agència considera il·lícits independentment de qualsevol consentiment, ja que l'assistència pot controlar-se fàcilment mitjançant altres mètodes menys intrusius. Encara hem de veure com acaba aquest cas.

En relació amb la precisió o la qualitat de les decisions automatitzades, hem vist exemples de problemes tant petits com grans. Per exemple, sembla que el sistema automàtic de peatge a les carreteres que identifica les plaques de matrícula no sempre pot distingir les plaques letones de les sueques, cosa que pot donar lloc a peatges indeguts que han de pagar-se mentre s'impugnen.

Un altre exemple el va proporcionar l'Agència Pública de Treball el 2018, quan es va constatar que un gran nombre de persones (estem parlant de milers) havien perdut les seves prestacions de desocupació per error a causa d'un algoritme defectuós. Les persones afectades finalment van cobrar les prestacions, però, per descomptat, havien patit una injustícia que hauria d'haver-se evitat.

El meu últim exemple és un cas que va portar un dels meus col·legues de la institució de l'Ombudsman de Suècia, que té a veure amb decisions automatitzades sobre prestacions socials en un municipi suec. La queixa no es refereix a la presa de decisions automatitzada com a tal, sinó a la falta de transparència de l'algoritme que s'utilitzava, que per descomptat és una qüestió de gran importància. Presumiblement el Govern local no ha estat disposat a revelar l'algoritme, el qual ha estat desenvolupat per una empresa privada. El meu col·lega encara no ha pres una decisió sobre aquesta qüestió. Aquest últim exemple em porta a fer algunes observacions finals sobre el paper dels ombudsmans del Parlament suec, tant de caràcter general com en particular pel que fa a la intel·ligència artificial.

Els ombudsmans suecs, com ja sabeu, som quatre, tots independents i designats pel Parlament, i tenim la missió principal de supervisar l'aplicació de les lleis i altres estatuts en l'Administració pública. En particular, hem de garantir els requisits constitucionals d'objectivitat i imparcialitat i que els drets i les llibertats fonamentals dels ciutadans no siguin infringits en l'Administració pública. L'objectiu general és promoure l'estat de dret. Els ombudsmans tramiten les queixes dels ciutadans, però també actuen per iniciativa pròpia i fan inspeccions. Durant les nostres investigacions tenim àmplies competències establertes en la Constitució, però quan es tracta de les nostres decisions, són formalment consultives i no podem canviar la decisió d'un tribunal o d'una agència pública. Podem proposar esmenes a la legislació per resoldre les deficiències legislatives, però el Govern no té l'obligació legal d'actuar sobre les nostres propostes. Pel que fa a l'ús de la intel·ligència artificial, els defensors podem supervisar el sector públic i facilitar orientació sobre bona governança, transparència, accés a la justícia i, en general, promoure l'estat de dret. A més, podem emetre opinions sobre proposicions de llei, en què les implementacions proposades en matèria d'intel·ligència artificial poden implicar riscos per als drets fonamentals, i podem proposar canvis en la legislació existent quan sigui necessari.

La intel·ligència artificial, com tots sabem, no és un fenomen que pugui separar-se, per exemple, dels drets fonamentals. Quan l'Administració pública la utilitza com a eina, la intel·ligència artificial forma part intrínseca de l'ús dels poders públics i ha de ser revisada com a tal.

Els nostres reptes impliquen, com sempre, prioritzar l'ús de recursos, però en aquest cas també per trobar la competència adequada per supervisar eines digitals tècnicament complexes. Amb la privatització dels serveis públics en constant increment, també és un repte la divisió pública/privada en la nostra supervisió.

En efecte, tenim reptes en la nostra funció d'investigació. Les nostres activitats internes també han d'estar actualitzades. Fins i tot les institucions més antigues han d'adaptar-

se als canvis en la societat, malgrat que encara respectin principis bàsics importants. Per tant, hem d'intentar ser més eficients mitjançant l'ús de solucions digitals i d'eines d'intel·ligència artificial dins les nostres pròpies oficines. Potser mitjançant un robot que classifiqui les queixes entrants, com han fet els nostres col·legues noruecs. A més, hem d'assegurar-nos que les habilitats del nostre personal s'adaptin a les noves necessitats.

Per afrontar millor els reptes i les possibilitats d'un futur més digital, he suggerit al Parlament suec que pot ser que els ombudsmen necessitem un mandat lleugerament més ajustat. Com a resultat, els meus col·legues i jo celebrem que el Parlament hagi decidit recentment que es revisaria el marc regulador de la institució de l'Ombudsmen, tant en relació amb aquest aspecte com en d'altres. Estic segura que estarem més ben equipats per continuar promovent l'estat de dret en la nostra societat actual, amb la intel·ligència artificial formant-ne part.

Amb aquest comentari acabo, gràcies per escoltar-me. Us envio els meus millors desitjos per a un debat fructífer i espero poder veure-us en persona la propera vegada.

Quirine Eijkman

Comissària i vicepresidenta de l'Institut Holandès de Drets Humans

Gràcies, moltes gràcies. Gràcies per convidar-me a ser aquí. Avui represento l'ENNHRI, la Xarxa Europea d'Institucions Nacionals de Drets Humans, i també l'Institut Holandès de Drets Humans. Moltes gràcies. M'ha agradat molt la teva presentació (amb referència a la presentació de la ponent per part del moderador), per cert, especialment la part en què has dit que l'Institut Nacional de Drets Humans o altres col·legues dels ombudsmen podrien tenir un paper per desenvolupar, potser també podrien col·laborar amb unes altres institucions i parlaré una mica més sobre la situació holandesa en un minut.

Però primer l'ENNHRI. Què fa l'ENNHRI? Aquesta Xarxa Europea d'Institucions Nacionals de Drets Humans, que té quaranta-cinc membres. Nosaltres donem suport a altres instituts nacionals de drets humans en el seu establiment, en l'intercanvi de bones pràctiques, en el desenvolupament de capacitats i en la capacitació, intervenim en documents legals i de polítiques, com ara la IA, i ens comprometem amb els mecanismes regionals. I estem molt contents de col·laborar també amb l'IOI. És molt important per a l'ENNHRI, en particular perquè la meitat dels nostres membres tenen mandats a l'Institut Nacional de Drets Humans i com a ombudsmen. Això no succeeix igual en tots els països. Bé, moltes gràcies, senyor Rafael Ribó, per organitzar aquest esdeveniment, ens sentim honorats de participar-hi.

Vull explicar-vos una mica sobre els instituts nacionals de drets humans abans d'explicar què fem en el context holandès i per què estem compromesos amb la digitalització i els drets humans com un programa estratègic, i també amb la intel·ligència artificial. Els instituts nacionals de drets humans són organismes amb mandat estatal, independents del Govern. Tenim mandats amplis per promoure els drets humans en l'àmbit nacional. Se'ns avalua periòdicament sobre la gestió de les queixes davant les Nacions Unides pel que fa a diversos principis que garanteixen la independència, el pluralisme, la responsabilitat i l'eficàcia. I no sé vosaltres, però jo solia treballar a Amnistia Internacional i quan treballava a Amnistia Internacional als Països Baixos mai no vaig haver d'explicar què eren els drets humans o quines actuacions desenvolupàvem, però tan aviat com vaig arribar als Països Baixos i vaig parlar amb ciutadans holandesos, sempre em preguntaven quins haurien de ser els drets humans al nostre país. Està demostrat que és una cosa que trebalem més en política exterior que dins el nostre país. Per tant, de vegades certs problemes no es perceben realment com un problema de drets humans, i això també s'aplica a la tecnologia i a l'ús de la tecnologia per part del nostre Govern i de les empreses privades.

Els instituts nacionals de drets humans aborden tots els diferents drets humans, inclosos els drets civils, polítics, econòmics, socials i culturals, i ens enfrontem cada dia als reptes, però també a les oportunitats, com acaba de dir Elisabeth Rynning, sobre les oportunitats que

la tecnologia i la intel·ligència artificial també aporten al nostre propi treball, al treball de l'oficina de l'Ombudsman als Països Baixos, i també a la ciutadania.

Per això, l'ENNHRI està molt interessada en aquests temes i, igual que els Països Baixos, també som nous en aquesta matèria, n'hem d'aprendre molt, molt treball per fer i molt treball per dur a terme pel que fa a posar-nos al dia. No obstant això, dins la nostra xarxa intentem compartir les nostres experiències, per la qual cosa si teniu alguna cosa a aportar, si us plau participeu també a les reunions de l'ENNHRI, hi ha fins i tot un comitè especial que ara s'ocupa de la intel·ligència artificial. Crec que és només compartint i sent un col·lectiu que podem ser més eficaços.

L'Institut de Drets Humans dels Països Baixos, com a membre de l'ENNHRI, estem encantats d'intervenir en aquest taller i hi estem participant com a experts, i al mateix temps puc dir que en aquest últim dia i mig he après moltes coses que puc compartir amb tot el meu equip, així que estic molt, molt feliç de ser aquí. L'Institut de Drets Humans dels Països Baixos no és només un institut de drets humans, també som un organisme d'igualtat de tracte, per la qual cosa amb els altres nou comissionats, inclòs el nostre president, tenim una política, observem les lleis, la política, prestem assessorament com a experts, etc., però també ens ocupem de les queixes individuals en relació amb la igualtat de tracte. La gent pot presentar queixes individuals i després podem emetre una resolució, en fem al voltant de 150 per any.

Crec que són igualment importants les trucades telefòniques que rebem cada dia de persones que simplement estan atrapades en el sistema. I aquestes trucades telefòniques solen començar amb persones que diuen: "S'estan vulnerant els meus drets, si us plau ajudeu-me, no entenc el sistema, pot ser que sigui un analfabet digital", amb relació al que també deïeu. De vegades són gent gran o també persones que no parlen holandès o que tenen altres limitacions, els resulta força difícil interactuar amb l'Estat o amb empreses privades per, per exemple, sol·licitar un compte bancari, etc. Així que rebem moltes d'aquestes trucades en el nostre servei d'atenció a les persones i sovint intentem ajudar les persones a entrar en el sistema o a fer-lo anar. Ara també ens trobem amb aquest tipus de casos a l'Institut Nacional de Drets Humans, on intentem defensar els drets humans, oferir assessoria experta en projectes de llei, informes parlamentaris, etc.

Durant els anys, i també a causa de les limitacions pressupostàries, hem intentat crear programes estratègics. Durant un parell d'anys treballem en aquests programes, per això intentem ser més efectius. Per exemple, l'últim parell d'anys vaig dirigir un programa sobre educació en drets humans, que no és obligatori als Països Baixos, i intentem advocar, incorporar això a la legislació, i treballem amb un equip d'aproximadament deu persones per aconseguir-ho. Un dels nostres nous programes és Digitalització i Drets Humans, i la raó per la qual seleccionem aquest programa és perquè sentim que hem de complementar el treball que està fent la nostra Autoritat de Protecció de Dades, perquè som conscients dels reptes que planteja la digitalització, i també per descomptat incloent-hi la intel·ligència artificial, no és només una qüestió de privacitat o de protecció de dades, sinó que també hi estan implicats molts altres drets humans.

Ara, en aquesta conferència aquí, això és molt evident, perquè ja parlem d'igualtat de tracte, discriminació, accés a la justícia, etc. Però crec que també hem d'adonar-nos que en molts debats públics i polítics, almenys al meu país, aquest problema només està relacionat amb la privacitat. Llavors, realment sentim aquesta urgència per formar els que prenen decisions i el públic en general pel que fa al fet que les amenaces des d'una perspectiva de drets humans no es relacionen principalment amb qüestions de privacitat i de protecció de dades, sinó que són molt més àmplies que això.

La segona raó per la qual hem seleccionat aquest programa és perquè ca-da vegada més rebem informes en què veiem que la igualtat de tracte en el mercat laboral és un tema molt important, també en el tema educatiu, especialment en la nostra formació professional, i sentim que havíem de fer alguna cosa sobre això. Perquè, ens agradi o no, la contractació de nous companys o la contractació de persones que vulguin treballar es farà cada vegada més per ordinador o per IA, algorismes o sistemes que

hi estan relacionats, siguin semiautomàtics o automàtics. I una de les coses en què volem implicar-nos i parlar amb els empresaris és sobre si aquests mecanismes són neutrals i si poden beneficiar a tots de la mateixa manera, perquè hi ha un gran risc de biaix.

Ara bé, si parlem de la segona part del nostre programa, es refereix al fet que a Holanda en teoria hi ha accés a la justícia. Així, per exemple, el programa SyRI, del qual s'ha parlat un parell de vegades durant aquesta conferència i del qual us explicaré una mica més en un minut. Hi ha aquesta creença, la qual no és certa, que la gent no pot oposar-se a la decisió que es va prendre o es va relacionar amb aquest sistema. Aquest no és el cas, però el problema és que les persones sovint no són conscients de com es pren la decisió que els afecta. Per tant, les persones assenyalades per aquest sistema automàtic per un frau social no reben una carta del Ministeri o del municipi que digui "t'avisem per aquesta i aquesta raó, i per això estem investigant el frau". No van obtenir aquest tipus d'informació. Llavors, una de les coses que volem aconseguir amb el nostre programa és fer que l'accés a la justícia a la pràctica sigui més transparent, que les persones siguin més conscients de quina és la base de la decisió d'investigar-les, perquè pensem que el sistema de risc i el sistema de negació són utilitzats per moltes autoritats públiques, per l'autoritat fiscal, etc. I per a mi, personalment, no dic això en nom del meu Institut, crec que aquests sistemes hi són per quedar-s'hi, ens agradi o no.

La qüestió és: són a prova de drets humans? Contenen drets humans per defecte? El sistema està creat de manera justa? I tot i que el Tribunal de Districte de l'Haia va dictaminar que el sistema d'eines de detecció havia de complir amb uns altres estàndards, no crec que s'hagi de modificar de manera directa, és molt probable que els legisladors holandesos modifiquin la llei i aquest sistema d'una forma o d'una altra s'utilitzi en el futur, i no dic que sigui bo o dolent, però també crec que hem de formar els nostres membres del Parlament i els legisladors que prenen les decisions sobre l'ús d'aquest tipus de sistemes perquè prenguin consciència que han de ser a prova de drets humans en comptes de dir que el sistema com a tal no s'ha d'utilitzar.

La segona qüestió és que actualment, i crec que és potser el tema més interessant per acabar, hi ha molts debats al Parlament holandès sobre la supervisió pública de l'ús de la intel·ligència artificial per part de les autoritats públiques. Això no només està relacionat amb els sistemes d'indicació social per a detecció, per tant, el sistema SyRI, sinó que també està relacionat amb l'autoritat fiscal. L'ombudsman sueca, Elisabeth Rynning, va explicar una mica sobre la intel·ligència artificial i les relacions públiques (PR), que la policia fa fotos d'automòbils i decideix automàticament si estan sancionats o si reben una multa o no, aquests sistemes s'utilitzen, i ara els membres del Parlament holandès diuen: "Hauríem de fer alguna cosa, potser hauríem d'elaborar totes aquestes mocions parlamentàries al·legant que hauríem de tenir un nou comissionat, un nou comissionat d'intel·ligència artificial per fer alguna cosa sobre això, haurien de registrar-se, haurien de fer això, haurien de fer allò".

El novembre, el ministre de l'Interior va decidir que s'havia de fer una investigació oficial sobre l'ús d'algoritmes per part de les autoritats públiques i va tenir lloc un debat sobre com s'ha de regular o si ja està ben regulada la fiscalització en aquest context. I penso que realment vaig gaudir de la investigació que van dur a terme Valerie Frissen i Marlies van Eck. Marlies van Eck és una investigadora molt interessant dels Països Baixos que ha fet molts estudis, acaba d'acabar el seu doctorat sobre la presa de decisions automatitzada per part de l'autoritat fiscal. I el seu consell al ministre de l'Interior és que no hauríem de crear més o nous comissionats o òrgans, sinó que els mecanismes de supervisió existents, com ara l'Ombudsman, l'Autoritat de Protecció de Dades, l'Institut de Drets Humans dels Països Baixos han d'utilitzar els seus mandats per investigar aquest tipus de qüestions, i també han de fer que les autoritats públiques retin comptes per explicar aquests usos amb més claredat i fer que es responsabilitzin si utilitzen aquest tipus de sistemes per a ús públic, de manera que siguin transparents, que siguin a prova de drets, etc., és a dir, l'ús d'algoritmes és part de teu deure de protegir, igual que succeeix amb uns altres sistemes que utilitzis.

Finalment, quan parlava d'aquest informe sobre l'ús d'algoritmes per part de les autoritats públiques, es va entrevistar el nostre institut i vam estar parlant amb molts dels organismes públics que són realment responsables de proporcionar benestar, responsables de tot tipus de serveis públics, més en concret de la part executiva del nostre Govern, va haver-hi molta gent implicada. L'analista científic de dades va dir: "Tenim un problema molt gros i un dels inconvenients és que hem de prendre moltes decisions ètiques sent comandaments intermedis, i els superiors en les nostres organitzacions (el president, els directors, etc.) no sempre entenen bé aquesta dificultat". I crec que això també és una cosa que ens hem de prendre molt seriosament per formar les persones, els que prenen les decisions, perquè realment compreguin i assumeixin la responsabilitat de les decisions que prenen basant-se en la intel·ligència artificial. Quan vaig estar en aquesta reunió d'experts, realment em vaig adonar que els experts en ciència de dades se senten molt incòmodes a l'hora de prendre aquestes decisions, i crec que realment n'hem de ser més conscients, perquè si com agència governamental utilitzes la intel·ligència artificial, has d'assumir la responsabilitat del que estàs fent i crec que la sentència del Tribunal de Districte de l'Haia ha fet que el Ministeri d'Afers Socials en sigui més conscient.

Per acabar, hauríem de fer molt més per formar els nostres legisladors quan aquestes lleis són aprovades pel Govern, perquè quan SyRI va ser aprovat pel nostre Parlament fa un parell d'anys, ni tan sols va haver-hi una votació a mà alçada, va ser simplement "per cop de martell" (per imposició o 'per cop de decret') i ningú no va fer preguntes, i crec que aquí és on hi ha el veritable problema, perquè m'agrada el litigi estratègic, però en realitat no és una alternativa per a una bona presa de decisions polítiques, igual com succeeix amb les lleis que s'aproven al Parlament en relació amb la intel·ligència artificial.

Llavors, aquests són molts reptes, no necessàriament només per a les oficines de l'ombudsman, però estic completament d'acord amb el que deia Rob, que no crec que els ombudsman o l'Institut Nacional de Drets Humans puguin fer aquestes coses sols, realment hem de cooperar i tenir una perspectiva estratègica conjunta sobre com hem de garantir la rendició de comptes, tant en el sector públic com en el privat.

Moltes gràcies.

Eric Houtman

Ombudsman federal belga de l'energia i membre de NEON (Xarxa Europea d'Ombudsman de l'Energia)

Hola a tothom, estimats companys, soc Eric Houtman, ombudsman belga de l'energia, i també membre de la NEON. La NEON és la Xarxa Europea d'Ombudsman de l'Energia, i actualment està formada pels ombudsman de l'energia de Bèlgica, el Regne Unit, Catalunya, Irlanda, Itàlia, Polònia i Geòrgia. I Bèlgica, el Regne Unit i Catalunya, amb el Síndic, en van ser els membres fundadors. Soc un ombudsman sectorial i tinc la missió legal d'atendre les queixes i reclamacions entre usuaris finals, persones i empreses d'energia privades. Per tant, estic molt agraït, per descomptat, com a ombudsman sectorial, que m'hàgiu convidat a aquest esdeveniment organitzat per l'IOI, i felicito el senyor Rafael Ribó per l'excel·lent organització d'aquest esdeveniment.

La intel·ligència artificial ja està definint els nostres patrons de consum d'energia. En pocs anys, l'assistència a la llar i els comptadors intel·ligents s'han generalitzat. Estem envoltats de molts dispositius connectats a la xarxa, començant pels nostres telèfons intel·ligents, i d'aquí a pocs anys les màquines i el programari d'autoaprenentatge podran utilitzar la quantitat increïble de dades que genera la Internet de les coses per prendre decisions i dur a terme accions sense supervisió humana. Potser demà el nostre cotxe elèctric autònom decidirà per si sol quan carregar la bateria després d'haver parlat amb la nostra casa intel·ligent.

Així doncs, us donaré una visió general de l'ambició verda i digital d'Europa en el sector energètic, perquè la digitalització és un dels pilars fonamentals de les noves polítiques energètiques europees. Permet descentralitzar fonts d'energia, com ara l'energia produïda per panells solars individuals, per integrar-les a la xarxa, i serà un excel·lent motor de descarbonització i eficiència energètica. Per exemple, la intel·ligència artificial i el *big data* poden detectar les necessitats energètiques amb més precisió i donar lloc a infraestructures energètiques més eficients i un consum adaptat. En alguns països, la gent ja té accés a preus dinàmics que coincideixen amb els fluxos generals i els preus majoristes, i els agregadors també actuaran com a intermediaris per als consumidors combinant dades de demanda i flexibilitat, venent-les al mercat majorista i compartint els beneficis amb els usuaris finals.

Hi ha reptes per als consumidors? Per descomptat, els consumidors s'enfrontaran a més dades, i totes aquestes dades seran gestionades i emmagatzemades de forma segura. En realitat, és molt fàcil: totes les dades abans de la lectura del comptador són gestionades pel gestor de la xarxa de distribució, i totes les dades posteriors a la lectura formaran part del sector privat, seran responsabilitat del sector privat. Hi haurà més actors. Qui serà el responsable de què? El consumidor ha d'estar actiu, s'ha d'activar, serà un consumidor proactiu, tindrà un acord de condicions contractuals amb el proveïdor, però també amb altres proveïdors de serveis. Ja passa actualment que els proveïdors que ofereixen productes com la fusta, o els pèl·lets, també ofereixen altres serveis com el manteniment i la reparació de la instal·lació de calefacció o de les plaques solars, això ja és així.

La pregunta més important, per descomptat, és: la intel·ligència artificial hauria de prendre totes les decisions? Però, sobretot, què passa amb les persones vulnerables que no tenen la capacitat financera, tècnica i digital per convertir-se en un consumidor actiu o proactiu? I encara una altra pregunta: què passa amb la discriminació provinent de les referències creuades de dades? Això ja està passant avui en dia en el sector energètic; s'anomena *intercanvi misteriós*. El vostre contracte energètic el pot assumir un proveïdor nou, i aquest proveïdor és el proveïdor del vostre veí, perquè utilitzava el codi d'identificació del vostre comptador en lloc del veí, i tindreu un contracte amb un proveïdor nou, i això s'anomena *intercanvi misteriós* o *segrest de dades*, la discordança de les dades. I totes aquestes preguntes són les raons per les quals el sector energètic és un dels principals objectius del llibre blanc de la Comissió Europea que es va publicar fa catorze dies, crec.

Hi ha reptes per a l'ombudsman? És clar. Hem de treballar amb noves relacions. I he esmentat la NEON perquè la feina de la NEON va ser molt, molt bona a Europa. Perquè l'any passat es va publicar una nova directiva de l'electricitat, la quarta, i en l'article 26 hi ha una disposició que també esmentava els ombudsmen. I per què és important això? Perquè hi va haver un debat entre la Direcció General de Justícia i la Direcció General d'Energia. La Direcció General de Justícia deia: "La directiva del 2013 ja existeix, de manera que només podem fer referència a aquesta directiva ADR (en anglès *alternative dispute resolution*, mètodes alternatius de resolució de conflictes)".

Però hi va haver un problema, i la Direcció General d'Energia entenia el problema perquè era conscient que amb tots aquests nous productes, tots aquests nous serveis, els algorismes utilitzats per operadors i agregadors, no només hi haurà una relació contractual entre un consumidor i un proveïdor, sinó que també hi haurà altres relacions entre empreses, entre organitzacions sense ànim de lucre, per exemple: les escoles amb panells fotovoltaics al terrat poden tenir problemes amb el proveïdor o amb un agregador o amb un consumidor individual, per la qual cosa era important que també es mencionessin els ombudsmen en la disposició d'aquesta directiva perquè cal destacar que no només tractem qüestions, problemes o queixes entre empreses energètiques i usuaris finals, sinó també relacions de B a B, de C a C, i fins i tot de plataformes d'igual a igual i també de serveis que van indirectament lligats al subministrament d'electricitat i gas natural.

I aquesta també és la raó per la qual a finals de 2019, l'any passat, vaig escriure un memoràndum, per assessorar el govern federal, que és bastant inexistent, tothom sap que tenim problemes a Bèlgica per tenir un govern federal amb plenes competències. Però ho vaig escriure per al proper govern federal, perquè permetés ampliar la missió legal de l'ombudsman per abordar també aquestes noves relacions, nous serveis i nous productes.

El que també és important i és un repte per a l'ombudsman és que també s'han de reforçar les obligacions socials, ecològiques i ambientals de servei públic. És important que s'actualitzin per garantir com a mínim el mateix nivell de protecció que en l'actualitat, perquè crec que és difícil d'explicar a un usuari final que li poden tallar el subministrament perquè té un conflicte amb un producte nou o amb un servei nou, i que se li pot tallar l'electricitat o el gas natural. Crec que seria molt difícil d'explicar als usuaris finals, però pot passar. Pel que sembla, a Alemanya ja han tingut casos d'aquest tipus. Un altre exemple és que ara a Flandes han començat a distribuir comptadors intel·ligents i ha quedat clar que ara el proveïdor vol utilitzar els comptadors intel·ligents com a comptadors de prepagament. Quan la gent no pugui pagar les factures o se'n queixin, se'ls pot tallar el subministrament pel simple fet que el proveïdor ho pot demanar al gestor de la xarxa de distribució. És això el que volem? O volem tenir la mateixa protecció que ara? Per tant, hem d'adaptar la nostra legislació i també la nostra normativa.

També cal destacar que sense uns límits clars la intel·ligència artificial pot explotar les vulnerabilitats dels consumidors i generar publicitat segmentada nociva i pràctiques comercials deslleials. Ja tenim molts problemes amb les pràctiques comercials d'algunes companyies energètiques, no de totes però sí d'algunes, de manera que crec que hem d'evitar que la gent es vegi obligada a canviar de companyia perquè l'ordinador els diu que han de canviar. Perquè ja hi ha proveïdors de serveis que ofereixen la possibilitat a les persones a les qual ofereixen el servei de buscar-los dues o tres vegades a l'any el millor preu. No el preu més barat, sinó el millor preu en funció del comportament de consum del client. Però cal provar el servei, és clar, cal donar-li una oportunitat abans que aquest proveïdor de serveis pugui canviar aquest client de l'antic proveïdor al nou. Però ja és un servei que existeix i que fins i tot està sent desenvolupat per una organització de consumidors, una organització de consumidors privada de Bèlgica. Per tant, compte amb això. I, per descomptat, pel que fa a l'accés als mètodes alternatius de resolució de conflictes, a la justícia i a la regulació, especialment la regulació ha de ser més àgil i dinàmica en el sector energètic. Parlem de regulació dinàmica perquè hem d'afrontar tots aquests reptes sobre nous serveis, nous productes, nous actors, agregadors i altres proveïdors de serveis.

Crec que sempre hi hauria d'haver una supervisió humana i després, com a ombudsman, hauríem de ser capaços de controlar i verificar la intel·ligència artificial per corregir les mancances i exigir ajustos immediats si fos necessari. Però, per descomptat, estic d'acord amb vosaltres que necessitem els coneixements adequats, els coneixements tècnics necessaris per fer-ho, perquè no és tan fàcil tenir una visió i fer una supervisió reals de tots aquests mecanismes que intervenen en aquests sistemes.

I també és important, crec, perquè ara en la directiva de l'electricitat s'esmenta que tots els estats membres necessiten un sistema de compensació financera. No només com a compensació financera o reparació individual per a les víctimes de problemes relacionats amb l'energia, l'electricitat i el gas natural, sinó també per a les víctimes d'aquests nous productes, serveis o proveïdors de serveis. Com a ombudsman, crec que la intel·ligència artificial crea un nou ventall d'oportunitats per identificar patrons i deficiències, que podem transmetre al regulador i als responsables polítics per ajustar el mercat. Crec que aquest és un paper molt important per a l'ombudsman.

I ara passaré a les conclusions. Penso que és necessari un marc jurídic robust per tractar qüestions relacionades amb la intel·ligència artificial. Els ombudsmen també han de reivindicar el seu paper de vigilants imparcials en les transformacions ecològiques i digitals de l'economia de la UE.

I, finalment, però no menys important, tothom mereix privadesa, equitat, dignitat i el dret a rebre un tracte humà i la resposta d'un ésser humà.

Gràcies.

Intel·ligència artificial i drets humans

reptes, rols i eines de les institucions d'ombudsman

Dimarts, 3 de març de 2020

Conclusions i eines

Moderadora: Catherine De Bruecker, vicepresidenta europea de l'IIO i ombudsman federal de Bèlgica

Ulises Cortés, professor i investigador de la Universitat Politècnica de Catalunya (UPC), Barcelona Supercomputing Center (BSC)

Daniel Marco, director general d'Innovació i Economia Digital

Rafael Ribó, president europeu de l'IIO i síndic de greuges de Catalunya

Catherine De Bruecker

Vicepresidenta europea de l'IIO i ombudsman federal de Bèlgica

En primer lloc, vull agrair a Rafael Ribó i al seu personal l'organització d'aquest taller, ja que és un tema molt important i real que hem d'abordar. La intel·ligència artificial no és ciència ficció. Com ja es va dir, ja és present en la nostra vida quotidiana. El que és absolutament segur és que ahir potser érem ignorants, però avui ja no podem pretendre seguir sent-ho, després del debat tan intens i important que hem tingut durant aquests dos dies. Ara tots dos tenim la difícil tasca d'extreure conclusions. I com que aquests dos dies han estat tan intensos, he pensat que potser hauríem necessitat algun sistema d'intel·ligència artificial que ens ajudés a resumir tota la informació que hem recollit. Però, afortunadament, tinc una cosa molt millor que la intel·ligència artificial per ajudar-me en aquesta tasca, i estic molt content de donar la paraula al professor Cortés, que ens ajudarà a treure aquestes conclusions.

Ulises Cortés

Professor i investigador de la Universitat Politècnica de Catalunya (UPC), Barcelona Supercomputing Center (BSC)

Agraeixo a l'organització confiar en mi per presentar les conclusions de l'esdeveniment. Les sessions han estat molt interessants i he après molt dels diferents ponents i dels seus punts de vista sobre el tema que ens ocupa: l'impacte de la IA en els drets humans. Aquests dos dies aquí han estat molt fructífers i emocionants per a mi perquè, com vaig dir ahir, soc un professor d'intel·ligència artificial en una escola d'informàtica, tinc els meus alumnes, els meus equips d'investigació i gairebé no tinc contacte amb institucions legals.

Primer alguns agraïments: M'agradaria agrair a tots els ponents la seva participació. Alguns d'ells m'han ajudat a extreure aquestes conclusions, permeteu-me esmentar per ordre alfabètic Atia Cortés (BSC), Catherine De Bruecker (IOI), Joan Manel del Pozo (UdG), Geraldine Mattioli-zeltner (ECHR) i Martha Stickings (FRA). Access Now i FRA van presentar documents fascinants, i he de confessar que no vaig poder llegir el document que avui ens ha donat Geraldine Mattioli-Zeltner.

IA & Societat

Som-hi amb les conclusions. Fa dos dies de treball que intentem explicar-nos, des de diferents punts de vista, les qüestions següents: Què és el que preocupa quan parlem dels possibles usos de la IA i com aquests usos poden afectar els drets humans?

No hi ha una resposta única, però crec que ara ha quedat clar que la IA no és una entitat independent i aïllada. Molts participants s'adrecen a la IA com una cosa que existeix,

pensa, té intencions i, potser, voluntat política. Però la IA és només una creació intel·lectual que té moltes aplicacions d'èxit que ens afecten cada dia en molts aspectes de la vida, des dels més trivials fins als més complexos i confidencials. El que sí que és evident és que la IA té un valor econòmic important i, per tant, impacta en la societat. La disrupció de la IA està transformant els nostres països, organitzacions, mercats i vides. Hi ha una àmplia evidència d'això. De fet, si comparem les deu principals empreses de fa només quinze anys i avui dia, veurem la profunditat del canvi, només queda en una: Microsoft, i entre les deu principals que hi ha actualment només una no és de tecnologia: Saudi Aramco.

El que no és tan obvi per al públic en general és que aquesta tecnologia gairebé no està regulada, principalment perquè s'ha construït sobre les arrels de la tecnologia de guerra, que al seu torn no està regulada a causa de la seva pròpia naturalesa. A més, la IA s'utilitza amb un propòsit, per a això els humans assignen intencions a les màquines que utilitzen aquesta tecnologia per interactuar amb nosaltres. Una d'aquestes intencions és, per descomptat, guanyar diners, i les empreses intenten aplicar la IA en qualsevol camp i amb qualsevol intenció perquè gairebé no hi ha una regulació, no tenen límits. A més, en aquest mercat no hi ha autoregulació, és una terra lliure per fer-hi tota mena de coses que la imaginació i la tecnologia permetin. Els individus, les empreses i els governs fan el mateix: hi ha governs que utilitzen la tecnologia de la intel·ligència artificial per als seus propis fins, gairebé sense restriccions. Al món, hi ha una bona col·lecció d'exemples de països que utilitzen IA.

És imprescindible que hi hagi supervisió humana en els sistemes basats en IA. És un principi de la UE, si l'acció d'un sistema intel·ligent afecta els éssers vius i/o el medi ambient, aquesta acció l'han de controlar humans. Aquest concepte de control hauria d'estar clarament regulat, no podem posar l'home fora del cercle de la presa de decisions. Això seria simplement una mala idea i podria tenir moltes implicacions indesitjables. Altres artefactes i tecnologies sempre són supervisats. Aquesta idea de control comporta algunes preguntes: Com de prop s'ha d'exercir aquest control? Quantes agències hi han d'intervenir? Amb quina freqüència controleu el sistema? La resposta a aquestes preguntes té implicacions legals i econòmiques.

El rol de GAFAMA

Els governs, les societats industrials, les organitzacions econòmiques i fins i tot GAFAMA,¹ que significa Google, Amazon, Facebook, Apple, Microsoft i Alibaba, s'afanyen a influir en les bones pràctiques pel que fa al disseny i a l'ús dels sistemes basats en IA. No cal oblidar que la Xina representa, almenys, el 30% de tot el mercat, per la qual cosa cal tenir-ho en compte, ha de sortir en la foto. La Xina està invertint sumes astronòmiques en investigació d'IA amb la intenció de liderar-la i aplicar-la en el seu desenvolupament econòmic. Com vam dir ahir, fins i tot aquestes empreses gegants estan intentant establir polítiques i pautes ètiques, i hem d'elogiar-les per això. Tot i així, l'impacte de les aplicacions de la IA en els drets humans ha d'avaluar-se amb molta precisió. Qui li posarà el cascavell al gat? En aquest cas són tigres, no gats petits. Algú s'ha d'oferir com a voluntari per posar aquest cascavell perquè necessitem sentir-lo, sentir que s'estan acostant a nosaltres.

El joc de la regulació no es tracta només d'ètica i intel·ligència artificial. També hem de parlar d'aspectes legals, socioeconòmics i culturals –els elements ELSEC– és a dir, els que impulsa la Unió Europea. Entre aquests elements, hi hem d'incloure el gènere. No podem seguir endavant en aquesta tasca, sense considerar la meitat de la població, la meitat del talent que ha d'estar en primera línia i que moltes vegades queda oblidat en el nostre discurs. Una vegada que es tenen en compte aquests aspectes, hem de

² GAFAM ara és FAANG. FAANG és un acrònim que es refereix a les accions de les cinc empreses de tecnologia estatunidenques més populars i amb millor acompliment: Facebook, Amazon, Apple, Netflix i Alphabet.

preguntar-nos per què aquestes empreses intenten modificar els valors. Si Europa ha acceptat aquests valors, hem de lluitar per aquests valors. Per què permetem que les empreses modifiquin la nostra forma de comportar-nos?

Dilemes, valors i cintes de vídeo

Aquest títol prové d'una famosa pel·lícula. Pot ser que l'hàgiu vist. En el número 1 mostrem parells d'opcions que es presenten quan pensem en el disseny d'una aplicació, en aquest cas quan intentem dissenyar una aplicació basada en IA. En la nostra societat, gaudim de llibertat i independència. Tenim llibertat d'elecció: per exemple, la llibertat de no utilitzar els serveis d'Internet o dependre'n. Tenim llibertat d'elecció davant les recomanacions basades en màquines. És la vostra opció. Tanmateix, potser les recomanacions basades en les màquines s'haurien de regular per protegir els ciutadans de ser influïts.

Què succeeix si desitgeu triar entre seguretat i privacitat? Si desitgeu que un servei sigui més privat, llavors la seguretat ha de ser més forta, i això implica tenir interfícies menys flexibles. Si parlem d'explicabilitat versus acompliment. Amb la potència computacional disponible, el rendiment és fàcil d'aconseguir, per exemple, utilitzant *Deep Learning* (tècnica d'aprenentatge profund). Avui dia, el *Deep Learning* no pot explicar com arriba a una solució.

Hi ha altres parells com ara responsabilitat i confidencialitat; prosperitat i sostenibilitat. I després tot això crea una sobrecàrrega ètica: no es pot tenir-ho tot. A més, si fingeixes tenir-ho tot, has de pagar-ne el preu.

La navalla d'Ockham

Aquí teniu algunes idees per a la vostra tasca com a ombudsman. Si us plau, no reinventeu tot el marc legal. Com vam veure en la sessió d'abans de dinar, hem dit que Europa té un llegat important que s'ha d'explorar i utilitzar. Tanmateix, s'ha de tenir en compte que la tecnologia que compleix amb ELSEC i es fabrica a casa serà més costosa que la que prové de la Xina o dels Estats Units. A més, necessitem que els ciutadans en siguin conscients i han d'estar disposats que cal pagar el preu. Tenir tecnologia local basat en IA significa, per exemple, que les dades estaran dins Europa, controlarem els servidors, la persistència de les dades, la seguretat i privacitat de les dades i, quan sigui necessari, les dades s'esborraran. A més, els ciutadans poden, si és necessari, exercir el dret a l'oblit a Internet. Crec que totes aquestes característiques són fonamentals, però poden resultar cares.

Obtingueu l'avaluació tecnològica adequada per comprendre l'ús de cada servei basat en IA. No hi ha dos programes que siguin idèntics. Poden ser implementats de manera diferent per dos equips diferents que utilitzen la mateixa especificació, però cal assegurar que compleixin l'especificació. Dins Europa, tenim el talent per fer això. Les universitats tenen molts investigadors, la Unió Europea està finançant la investigació i l'educació del talent, però no podem retenir aquest talent.

L'educació tècnica per a la nostra societat és imprescindible. Per tenir l'educació tècnica i crítica adequades per a les generacions futures pot ser que sigui necessari incloure-les en el programa d'estudis, i potser complementar l'educació tecnològica de la ciutadania d'avui.

Deures

Tinc alguns deures per als ombudsman. La primera tasca té a veure amb la transparència: s'han de supervisar els certificats de transparència de tots els serveis d'IA en ús. Algú amb autoritat reconeguda ha de certificar que aquests serveis són justos. Per exemple, a l'Administració pública, tots els serveis basats en IA han de ser transparents, fins i tot es pot intentar fer alguna cosa directa i senzilla com utilitzar etiquetes de tres colors:

verd, groc i vermell, que seran suficients per informar la ciutadania. El motiu d'aquesta sol·licitud és el següent. Tot i que puc llegir i comprendre tots els documents legals que acompanyen qualsevol aplicació informàtica, és possible que no sigui capaç de comprendre els drets i les obligacions que contenen. Seria millor que algú autoritzat per l'Estat s'ho llegís, n'avalués la legalitat i els qualifiqués. Aquesta persona ha d'indicar si tant l'aplicació com el contracte són correctes per ser utilitzats, és a dir, si operen dins la nostra legalitat. Si és així, l'aplicació haurà de tenir un segell verd. Els ombudsman, per exemple, han de supervisar que totes les aplicacions en ús tinguin un segell.

Una altra tasca és què fer quan un servei té un segell vermell. Què fer quan les persones utilitzen un servei marcat amb un indicador vermell? Això és admissible? Com afecta això uns altres ciutadans. S'hi ha de reflexionar. A més, per descomptat, cal exigir al Parlament de la UE i/o al Consell d'Europa la legislació adequada per instal·lar aquests certificats de transparència a tot arreu d'Europa.

Una altra afirmació que hem sentit de David Cabo, de Civio, és que tot el programa produït amb fons públics i utilitzat a l'Administració pública ha d'estar disponible per ser inspeccionat. Estic d'acord que no es pot donar accés a aquest programa a qualsevol perquè el supervisi. Potser a través de la mediació dels serveis dels ombudsman, tercers, mitjançant l'ús de certificats de no-divulgació, podrien inspeccionar aquest programa. Crec que això augmentarà la confiança de la ciutadania en aquestes eines. Hem de recordar que si el govern va pagar pel programa, és propietat de la ciutadania.

Respecte a la responsabilitat (de la màquina). Què passa si una màquina falla en una decisió? Qui n'és el responsable?

Si us plau, cal donar suport i certificar només les aplicacions que siguin justes i sense biaixos. Sé que ambdós són termes confusos. Tanmateix, és necessari començar a pensar en aquests termes i com provar-los. És necessari tenir un *sandbox* públic per provar tots aquests programes i una regulació per al cicle de vida del programa. Per respondre preguntes com ara: Quan serà necessari requalificar una peça de programa?

A més, crec que cal supervisar tots els algorismes que s'utilitzen a l'Administració pública, en particular, els que s'utilitzen per agregar dades personals per prendre decisions en, almenys, els camps següents: salut; autonomia humana; llibertat de moviment; llibertat, seguretat, justícia, igualtat davant els tribunals (ja hem parlat de Compass); privacitat i protecció de dades; igualtat, no-discriminació i participació política (el desafortunat exemple de Cambridge Analytica als Estats Units i altres llocs d'Europa, per exemple a Espanya, són exemples de com es poden utilitzar aquests algorismes per modificar l'opinió de les persones a l'hora de votar, és a dir, per triar els seus representants polítics); educació i medi ambient.

Permeteu-me cridar la vostra atenció sobre la idea de permetre que Google presti serveis a les escoles, instal·lant Google Suite, penso que és una aberració. Els infants no necessiten això. Hem arribat a aquest punt de la civilització i tots, la majoria de vosaltres, mai no va utilitzar Google Suite quan va anar a escola. El màxim que teníem era una petita calculadora. Per què els infants ara necessiten tenir tots aquests serveis i eines? Els infants no ho necessiten. Necessiten desenvolupar el pensament crític i després utilitzar la tecnologia. Som en aquesta reunió perquè estem utilitzant eines basades en intel·ligència artificial sense entendre realment com funcionen, i temem que sigui així. Com a societat, no necessitem que la tecnologia s'utilitzi simplement per defecte, sense comprendre-la.

Més sobre responsabilitat. Donar suport a l'evitació de l'ús en l'Administració pública de serveis de vigilància inapropiats i/o recopilar inadvertidament dades personals. Hem d'evitar-ho. Tanmateix, per descomptat, hi ha organismes encarregats de fer complir la llei que poden necessitar aquestes tècniques en un moment donat per prestar a la societat el nivell de seguretat adequat. Sé que aquest és un tema controvertit i

font de discussió. Els estats han d'estar preparats per tenir un debat amb deteniment sobre aquest tema perquè la societat necessita tenir aquests serveis. Pot ser necessari utilitzar-los en determinades circumstàncies per rastrejar una persona considerada un objectiu durant un període acordat dins les forces de seguretat. Això implica tenir-los a mà per quan es necessitin i desconnectar-los quan hagin complert la seva funció. Ha de quedar clar qui supervisa que aquest procés sigui respectuós amb els drets humans. Les societats menys democràtiques ja comencen a utilitzar-los per controlar els seus ciutadans.

Pel que fa a la rendició de comptes, és necessari certificar la presa de decisions automatitzada i evitar la bretxa en la rendició de comptes.

La confiança és un tema central. Ja hem sentit al ponent anterior parlar-ne. Una forma de generar confiança és ser responsable. És responsabilitat humana dissenyar correctament l'ús de dades i la presa de decisions que fa una màquina. Les màquines no ho faran. Hi ha d'haver algú identificable a qui se li pugui preguntar directament "per què aquesta màquina fa això o funciona així?". És necessari crear una metodologia que permeti identificar la responsabilitat en el disseny del sistema basat en IA perquè això implicarà la intencionalitat amb què es va programar en el sistema. Quan una màquina pot respondre com va arribar a una solució i pot ser auditada i és responsable, la utilitzaríeu per respondre davant un jurat? És una pregunta.

La resposta és sí; podeu fer-ho. Ja hi ha màquines que poden presentar arguments i argumentar com un advocat, més i tot, si tenen prou evidències que poden utilitzar. Estem preparats per fer això? Si estem preparats per fer una intervenció quirúrgica mitjançant l'ús d'un robot, per què no? La majoria de nosaltres volem en avions que en la seva majoria volen amb el suport de tres computadores. El pilot pilota la major part del temps, però qui fa la major part del càlcul és el sistema de tres computadores. Per tant, confiem en les màquines que prenen decisions crítiques. Per descomptat, el pilot sempre hi és per supervisar el circuit.

Aquest és l'últim punt i és molt dur. Totes les dades produïdes per ciutadans en una ciutat, regió o país pertanyen als productors. Les administracions públiques han d'assegurar que aquests productors puguin obtenir un benefici directe. A la ciutat de Mèxic, les autoritats lluiten contra Google i altres empreses per obtenir totes les dades produïdes per l'ús de Waze i Google Maps per millorar les seves infraestructures. Per què? Això mostra que aquestes empreses treballen acumulant dades a canvi d'un servei. Les ciutats i els governs els permeten operar; utilitzen la infraestructura nacional de telecomunicacions. Els usos de les dades dels ciutadans han de ser en benefici de la ciutadania.

Crec que hem de regular el cicle de vida dels serveis basats en IA. A més, poso aquí un exemple que és molt fàcil d'entendre: si es vol crear un químic per curar alguna malaltia, aquest procés pot durar anys. El descobriment del fàrmac pot portar diversos anys, entre trobar els objectius i dissenyar la molècula. Després els assajos clínics, que duren sis anys. La societat no necessita que la tecnologia canviï cada any, perquè primer hem d'experimentar-hi, hem d'estar segurs que la tecnologia dona els resultats adequats i que no és perjudicial. És necessari regular el temps i les condicions de lliurament al mercat amb les proves oportunes. Ho fem amb empreses farmacèutiques, per què no podem fer-ho amb empreses de programa o altres telecomunicacions?

No sé si la IA és el problema o la solució. És una àrea científica. La nostra investigació intentem esbrinar quins són els mecanismes que hi ha darrere del que anomenem *intel·ligència animal*. Volem reproduir-ho en una màquina. Un dels objectius que pretenem és aconseguir una mica d'intel·ligència artificial que sigui diferent de la intel·ligència humana. En aquest moment, un diàleg amb aquesta intel·ligència pot ajudar a resoldre problemes des d'un altre punt de vista. La IA té com a objectiu reemplaçar els humans. Aquesta no és la idea. En la meua opinió, és una idea perversa crear màquines, qualsevol màquina, per reemplaçar algun lloc de treball

de qualsevol individu. Si algú vol reemplaçar un ciutadà en una línia de treball, primer ha de proporcionar els mitjans perquè aquestes persones es guanyin la vida i paguin impostos pel seu treball, pel que fan.

"La humanitat era el meu negoci; la caritat, la misericòrdia, la paciència i la benevolència eren, tot, cosa meva. Els tractes del meu comerç no eren més que una gota d'aigua en l'extens oceà del meu negoci!"

Aquesta és una cita d'*Un conte d'hivern*, de Charles Dickens. Sempre parlem d'humanitat i volem millorar la vida de la humanitat. No volem utilitzar la tecnologia per fer-vos viure en una realitat virtual on somrigueu i les persones que no somriuguin siguin les que controlin la societat. Moltes gràcies.

Daniel Marco

Director general d'Innovació i Economia Digital

Bona tarda, hello everyone. Agraïment i felicitacions al Síndic de Greuges i a l'International Ombudsman Institute per l'organització d'aquest seminari i també a tots els experts que hi heu participat com a ponents o com a públic. Per mi és un honor poder participar avui aquí en aquesta cloenda, que miraré que sigui breu. Però la veritat és que he pogut seguir *online* les diferents xerrades i taules rodones i serà difícil afegir alguna cosa que no sigui reiterativa. O pitjor, intrascendent. Però miraré d'aportar algunes conclusions i també des de la meua perspectiva personal i en nom del Govern de la Generalitat les properes passes que hauríem de seguir.

El que ha quedat clar és que la revolució digital està transformant la nostra societat i que la intel·ligència artificial és una de les tecnologies que més efecte tindrà per la seva capacitat transformadora i impacte en el conjunt de la societat. Aquest efecte transformador de la revolució digital també afectarà en canvis, com hem vist, en el que seria com entenem les relacions humanes i els drets de la ciutadania amb la mateixa societat. De la mateixa manera que la revolució industrial va canviar, va modernitzar la societat i també va permetre l'assoliment de certs drets socials, ens trobem en un moment en què les tecnologies digitals canviaran, voldria dir radicalment, però com a mínim segur que canviaran les estructures socials, i això ens fa que haguem de tenir en compte com renovem aquest pacte social o com aconseguim en aquest món digital no perdre els drets socials adquirits i sobretot millorar. Al final, una modernització, una transformació de la societat, ha de ser per a la millora del bé comú. Seran necessaris nous conceptes associats a una reflexió i un debat sobre els límits del desenvolupament ètic i l'impacte positiu que aquest desenvolupament ha de tenir en la nostra societat.

Aquesta és una problemàtica global, i avui també s'ha estat aquí, aquests dos dies s'ha estat veient aquest impacte en l'àmbit internacional. Però com també s'ha comentat, la lluita pels drets, les llibertats... ha tingut una necessitat d'efecte d'acció local, i els guanys adquirits localment s'han pogut estendre de manera global.

Però tant, Barcelona com Catalunya, en aquest sentit, crec que han de tenir de manera local i amb impacte global un paper rellevant en aquest nou repte que s'ha estat comentant aquí. Jo personalment ho crec, per molts aspectes. N'anomenaré alguns, a banda de l'experiència a Catalunya en el camp de la intel·ligència artificial i de l'excel·lent comunitat científica que tenim, ja parlant de l'àmbit de l'ètica o els drets aplicats, el 2017 la comunitat científica va ser pionera a Barcelona, en la Declaració de Barcelona sobre la intel·ligència artificial, que de fet va marcar les bases amb experts d'arreu d'Europa sobre el que després ha estat la guia ètica, les directrius ètiques sobre la intel·ligència artificial de la Comissió Europea i altres països que ho han assolit. El 2018 vam posar en marxa el Digital Future Society, que també busca un fòrum de debat i reflexió sobre els àmbits de què avui i ahir es van estar parlant aquí. El 2019 el Govern va presentar la Carta de drets i responsabilitats digitals, que el que mira és d'incorporar aquests nous drets digitals en l'ideari del bé comú i de les administracions. I fa molt poquet, fa tres setmanes,

vam aprovar l'estratègia Catalonia.AI, l'estratègia d'intel·ligència artificial de Catalunya, que té un eix específic en l'àmbit de l'ètica i la societat, en el desenvolupament d'una intel·ligència artificial per al bé comú, transparent, que estigui centrada i pensant en les persones i que el desenvolupament i l'impacte siguin de consens.

Així, tenim un històric que ens pot posicionar bé en aquest entorn, però també tenim entitats d'activisme digital fortes, tenim un associacionisme i una societat civil organitzada en molts entorns, tenim experts en recerca en l'àmbit de l'ètica de la intel·ligència artificial. Per tant, tenim els ingredients de fer propostes i articular un debat polític i democràtic. El que sí que falta, potser, i hi hem d'incidir des de les administracions públiques i la societat en general, és portar aquest debat al primer nivell, tant de l'àmbit polític com de les places cíviques, de les organitzacions de barris i en la societat en general.

Crec que els debats que hem tingut aquests dos dies d'experts de diferents vessants és necessari portar-los al conjunt de la política democràtica per refermar i desenvolupar aquest consens, i les administracions polítiques ens ho hem de creure de la manera més profunda. I aquí, des del Govern de la Generalitat, hi treballarem, i no des d'una perspectiva únicament i exclusivament discursiva, sinó amb accions. Com us he comentat, l'estratègia dintre de la intel·ligència artificial de Catalunya anirà en aquest sentit. No només dient el que farem i no, sinó després fent accions relacionades els tancaments de webs, el *decretazo* digital o no estar en contra de la provisió d'armes letals autònomes en el seu desenvolupament.

És clar que fa setanta anys, quan es va realitzar la Declaració universal dels drets humans, el món digital, la societat digital, era inimaginable en aquell àmbit i que després de tot aquest desenvolupament el que hem d'aconseguir són dos aspectes fonamentals. En primer terme, que els drets universals dels drets humans també s'apliquin en l'entorn digital de la mateixa manera que ho hem aconseguit en molts àmbits, en la major part del territori i la societat. Per tant, mantenir, assegurar i vetllar per aquests drets. Però també començar a treballar sobre els nous drets digitals, nous drets emergents del segle XXI. Aspectes que potser no eren recollits en aquella Carta i que avui, aquests dies, hem començat a fer la reflexió.

Per part del Govern de la Generalitat, sapiguen que ho mantenim com a prioritat. Els drets, les llibertats, la igualtat de la nostra societat ha de ser també en un país digital, que perquè sigui també un país social necessitem la revolució digital.

Gràcies a tots per aquests dos dies, al Síndic per l'organització, i també per poder vetllar en els propers mesos i anys per, al cap i a la fi, tenir una societat més lliure, igualitària i equitativa.

Moltes gràcies.

Rafael Ribó

President europeu de l'IIO i síndic de greuges de Catalunya

En primer lloc, el meu agraïment a tots els traductors i al fantàstic treball que heu fet durant aquests dos dies. Però també moltes gràcies a tot l'equip del Síndic que des de fa moltes setmanes es va posar a treballar per preparar aquest esdeveniment. Moltes gràcies a tots els meus companys. Gràcies a tots els membres de l'IIO. L'últim taller de l'IOI que vam fer va ser un molt interessant a Riga. Tenim amb nosaltres en aquest seminari l'ombudsman de Riga. I també gràcies a tots els altres ombudsmen, locals i universitaris, que ens acompanyen durant aquests dies. Gràcies a totes les entitats que treballen en aquests temes a tot el món. El meu agraïment al Museu Hub, que ens ha proporcionat totes les facilitats, als ponents que van aportar diferents punts de vista als nostres debats, i als patrocinadors, especialment a "La Caixa". Vull expressar un agraïment molt especial a dues persones que, des de fa setmanes, mesos, quan comencem a parlar sobre aquest taller, van començar immediatament a parlar-nos d'intel·ligència artificial, encara sense estar familiaritzats amb el tema. El Dr. Ulises Cortés,

un home fantàstic, com heu sentit en l'última intervenció, que en sap molt, i amb un enfocament molt objectiu i crític sobre la intel·ligència artificial. Ulises, gràcies, moltes gràcies. I també a Carina, de Digital Future Society, que va ser l'última ponent de la primera taula rodona, i també ens va ajudar molt amb la presentació i la introducció dels diferents camps que havíem de tocar.

Hem estat retransmetent totes les jornades en reproducció en directe, ho teniu també a YouTube i es penjaran en el nostre web en qüestió de dies les intervencions completes. I considero que és un primer pas, almenys per a nosaltres, com he dit al principi, atès que érem ignorants en aquests temes. Com a primer pas i, com acaba de dir la vicepresidenta de l'IOI Catherine De Bruecker, estic segur que l'IOI continuarà en aquesta mateixa direcció. Almenys per, en la meua humil opinió, seguir aprenent.

Hem de confiar o no? Hem de confiar en la IA o no? Torno a utilitzar la paraula màgica: *confiança*. Jo diria: en sabem o no en sabem? Estem informats o no? Encara estem imaginant, a Catalunya, i potser en termes generals a Espanya, quina va ser la raó principal per suspendre el Mobile World Congress de la setmana passada. Hi ha diverses interpretacions, no en vull triar cap, però hi ha diverses interpretacions que demostren que ens falta la informació completa sobre decisions molt importants. I estic molt orgullós que el Mobile World Congress es faci a Barcelona, per descomptat.

Sabem exactament com validar els algorismes? I com aconseguir, com deia Ulises, informació lliure de biaixos? I això és molt important. Sabem com fer un treball constant d'auditoria, de rendició de comptes, sobre totes aquestes qüestions? Fa molt pocs dies, setmanes, tots sabíem a Espanya que s'aprovaria un impost per a algunes d'aquestes grans empreses. L'impost es va aprovar, però se'n va endarrerir l'aplicació sine die, potser a finals de 2020 o potser l'any vinent. Què va passar? Novament, falta d'informació o transparència dels factors reals i de la confiança.

Realment coneixem els nostres drets? I m'atinc molt a la paraula *drets*. Permeteu-me de dir-ho de manera brutal. Estic fart de sentir "codis d'ètica". Crec que va ser Daniel Innerarity o potser Lasalle que va dir: He comptat 84 codis d'ètica diferents. La tan famosa autoregulació. Per a mi, és molt important quedar-se amb el concepte de *drets humans*. I recordar que els drets van sempre més enllà de les lleis. Mai no es conquereixen els drets només amb l'estricta marc legal; es va sempre més enllà, com ha fet la humanitat durant la nostra història. En la segona taula rodona d'avui hem vist com els drets es definien com a qüestions vinculants, a càrrec dels poders públics: estatals, autonòmics, locals...i de protecció. Són vinculants, no són recomanacions, tenen responsabilitat pública vinculant i de protecció. Hem vist una llarga cobertura d'aquests drets. L'últim presentat per Ulises, sobre el nostre dret a la intimitat, el nostre dret a la igualtat de gènere, etc. I, per descomptat, els drets humans s'encarregaran de l'enorme tasca de l'educació, sobretot quan tots estem d'acord que hi ha un problema d'alfabetització o el que fa anys se'n deia la bretxa digital. Ara parlem d'alfabetització o analfabetisme en totes aquestes matèries.

Per això estic molt content, com a ombudsman, i com veig a través dels meus col·legues ombudsman, que l'IOI estigui ara en condicions d'abordar la intel·ligència artificial i els drets humans continuant amb l'aprenentatge i emprenent accions. Per què no presentar tot això també en la propera Conferència a Dublín? Ho dic sense cap compromís, perquè no en soc l'organitzador, i no sé si hi seré, però és la Conferència Mundial. Potser algú presenta aquestes coses allà.

Estic intentant incloure tota la tasca que hem vist durant aquests dos dies, especialment el que hem vist avui o el que haurien de ser els nostres deures. I recordo les tres crides que feia en la meua primera intervenció. Apel·lació a tots els ombudsman, com deia Catherine, nosaltres, els membres de la junta, demanem a la junta sortint i als que governaran l'IOI a partir de Dublín que continuïn amb aquest treball. Apel·lo a la implicació social. Algú va dir: Per què no obriu vosaltres, els ombudsman, una finestra a la participació social, començant per les queixes de tantes entitats que estan lluny de nosaltres, treballant per aquests drets i aquestes qüestions? I per descomptat

una tercera crida, no només als que heu vingut aquí en representació d'organismes internacionals, com Marta, Geraldine i el Sr. Bonnor, no només a vosaltres, a tots nosaltres, hem de... no diria canviar, sinó millorar o reformar, almenys millorar, els nostres tractats per tenir en compte tots els vessants d'aquest fenomen?

Acabaré amb una simple reflexió: ens trobem en un barri que històricament està a prop del que anomenem Poblenou, la ciutat nova, la nova ciutat petita. Però ara n'hi diem 22@, referint-nos a una concentració d'empreses emergents, d'empreses tecnològiques, etc. Vull recordar amb molt orgull que fa segles Catalunya va ser l'enlairament del capitalisme a Espanya. I aquí va començar la veritable industrialització, i per això som més o menys la màquina de l'economia. També vull recordar això, com va ser fa dos segles. Vull recordar que el 1851 a Barcelona va tenir lloc la primera vaga general en tot l'Estat espanyol, i naturalment va ser aquí, perquè és on hi havia totes les fàbriques i indústries. 1851. Però la bandera de la vaga era: associació o mort. Un dret humà molt profund, el dret d'associació, dels sindicats. I ara, també amb orgull, dic que aquí, al nostre voltant, hi ha la màquina de tota la digitalització, i més que això, de la intel·ligència artificial desenvolupant-se.

Assumeixo el meu compromís, de forma molt humil, de lluitar perquè a més del Mobile World Congress, la Digital Future Society, a més de totes aquestes coses fantàstiques, algun dia a Barcelona hi hagi un gran centre de drets humans per a la intel·ligència artificial.


Moltes gràcies.



SÍNDIC
EL DEFENSOR
DE LES
PERSONES


IOI - Europe
International Ombudsman Institute
European Chapter

Amb la col·laboració de:

 Digital
Future Society

 **BSC**
Barcelona
Supercomputing
Center

Amb el patrocini de:

 **Obra Social**
Fundació 'la Caixa'