EAPC - Gestió de la informació: transparència i protecció de dades

La regulació de la intel·ligència artificial a la Unió Europea

03-07-2023

Els experts ens avisen que la intel·ligència artificial (IA) ha de tenir un paper crucial en la transformació digital de la societat i que el seu ràpid avenç comportarà, i de fet ja està comportant, una revolució en el món tal com el coneixem.

Què està fent la Unió Europea per regular aquesta realitat?

La necessitat d'una regulació en aquest àmbit ve de lluny i en les directrius polítiques de la Comissió Europea per al 2019-2024 ja s'anunciava que la Comissió proposaria una legislació per disposar d'un enfocament europeu coordinat sobre les implicacions humanes i ètiques de la IA. En el marc d'aquest compromís, el febrer de 2020 la Comissió va publicar el *Llibre Blanc sobre la IA*, que recull opcions polítiques sobre com aconseguir el doble objectiu de promoure l'adopció de la IA i abordar els riscos associats a determinats usos d'aquesta tecnologia.

També el Parlament i el Consell Europeu havien posat de manifest la necessitat d'emprendre accions legislatives per garantir un bon funcionament del mercat interior dels sistemes d'IA que abordés tant els beneficis com els riscos de la IA.

En aquest context, l'abril de 2021 la Comissió va impulsar una proposta de Reglament del Parlament Europeu i del Consell per "establir normes harmonitzades sobre intel·ligència artificial (Llei d'intel·ligència artificial) i modificar determinades lleis de la Unió".

En els últims mesos han entrat al mercat sistemes d'IA com ChatGPT, desenvolupat per OpenAI, empresa fundada el 2015 als Estats Units, però també altres sistemes com BERT, de Google. ChatGPT o BERT es basen en els anomenats models generatius de base o bots, que s'entrenen amb grans quantitats de dades com text, imatges, música, veu o codi, amb l'objectiu de complir una sèrie de tasques àmplia i canviant. Les veus crítiques alarmen sobre els riscos de la parcialitat, la incitació a l'odi, les notícies falses, les violacions de la propietat intel·lectual, els ciberatacs i la línia cada cop més difusa entre el que és humà i el que és artificial que poden comportar aquests sistemes.

En el cas de ChatGPT, en tan sols cinc dies des del seu llançament va arribar a més d'un milió d'usuaris, amb un model d'ús gratuït i obert. Per fer-nos una idea de l'impacte que ha representat només cal analitzar les estadístiques d'implantació d'altres serveis digitals: TikTok, per exemple, va trigar nou mesos a arribar als 100 milions d'usuaris, mentre que Instagram va trigar gairebé dos anys i mig i Facebook més de quatre anys. En canvi, ChatGPT els va assolir en només dos mesos. De fet, segons diuen, s'ha convertit en la plataforma d'internet que ha crescut més ràpid de la història.

Aquest ràpid avenç ha generat preocupació a les autoritats de protecció de dades per l'impacte que pot tenir l'ús de la IA, i més singularment del servei ChatGPT, en els drets i les llibertats de les persones físiques afectades. Aquests dubtes provenen, d'una banda, de la manca d'informació sobre les dades de caràcter personal que tracta, tant les de les persones que les utilitzen directament com les de tercers que s'hi poden estar bolcant, que podria incloure el tractament d'informació sensible o relativa a col·lectius vulnerables (com ara menors d'edat) sense control i sense una base jurídica adequada. De l'altra, també genera dubtes la manca de control que tenen les persones afectades sobre aquest tractament, així com la impossibilitat pràctica d'exercir els drets d'autodeterminació informativa (accés, rectificació, supressió, etc.), a més del desconeixement de les condicions de seguretat, els terminis de conservació de les dades o la comunicació de dades a tercers, entre altres qüestions.

Davant dels dubtes respecte a la privacitat dels usuaris, el garant per a la protecció de dades personals d'Itàlia va prohibir-ne l'ús durant un temps, tot i que recentment ha aixecat aquest bloqueig.

Així mateix, l'APDCAT ha publicat una recomanació (Recomanació 1/2023) per informar les entitats i els organismes del seu àmbit d'actuació sobre els riscos d'utilitzar l'eina d'IA ChatGPT. L'APDCAT adverteix que no sembla oportú incorporar el ChatGPT en l'exercici de funcions públiques i en la prestació de serveis públics quan es tractin dades personals fins que el Comitè Europeu de Protecció de Dades s'hi pronunciï.

En la mateixa línia, el Comitè Europeu de Protecció de Dades (CEPD) ha creat un grup de treball per cooperar i intercanviar informació sobre accions que poden emprendre les autoritats de protecció de dades en relació amb aquesta qüestió, d'acord amb el principi de coherència recollit al Reglament general de protecció de dades.

També l'Agència Espanyola de Protecció de Dades, integrada en el grup de treball del CEPD, va anunciar que havia iniciat d'ofici actuacions prèvies d'investigació a l'empresa OpenAI, propietària del servei ChatGPT, per un possible incompliment de la normativa.

Aquest context ha propiciat que es reactivi la tramitació de la proposta de Reglament d'IA de la Unió Europea mitjançant la presentació d'un esborrany d'esmenes al text original amb l'objecte, segons van anunciar els eurodiputats encarregats de tramitar-la, de garantir, entre altres qüestions, que els bots siguin transparents, no produeixin continguts il·legals segons la legislació de la Unió Europea, compleixin les normes sobre els drets d'autor i respectin els drets fonamentals. En el nou text s'incorpora la necessitat que "Los modelos de base generativa deben garantizar la transparencia sobre el hecho de que el contenido es generado por un sistema de IA, no por humanos".

Com s'articula la regulació de la IA en el projecte de Reglament d'intel·ligència artificial?

El títol I del projecte de Reglament d'intel·ligència artificial defineix l'objecte del Reglament i el seu àmbit d'aplicació, que abasta la introducció en el mercat, la posada en marxa i la utilització de sistemes d'IA. En aquest títol també es recullen les definicions que s'utilitzen en el text normatiu. Es pretén que la definició d'IA sigui neutra tecnològicament i capaç de resistir el pas del temps. La definició del que s'hagi d'entendre per sistema d'IA és crucial, ja que en funció d'aquesta definició el Reglament serà aplicable o no a cada cas concret.

El projecte de Reglament segueix un enfocament basat en els riscos, i per això estableix obligacions per a proveïdors i usuaris en funció del nivell de risc que pugui generar la IA. Per tant, és fonamental per al sistema que es vol implantar identificar i classificar convenientment els sistemes en funció del risc. A partir d'aquí la proposta de Reglament distingeix entre els usos d'IA que generen un risc inacceptable (pràctiques d'IA prohibides), un risc alt (respecte de les quals es defineixen els requisits que han de complir) i un risc baix o mínim (que han de complir determinades obligacions de transparència).

1. Llista de pràctiques prohibides

El títol II del projecte de Reglament estableix la llista de pràctiques prohibides que inclouen tots els sistemes d'IA i quin ús es considera inacceptable pel fet de ser contrari als valors de la Unió, perquè, per exemple, violen drets fonamentals. En aquest llistat s'inclouen les pràctiques que tenen un gran potencial per manipular les persones mitjançant tècniques subliminars que transcendeixen la seva consciència o que aprofiten les vulnerabilitats de grups concrets (com els menors o persones amb discapacitat) per alterar-ne el comportament de manera substancial. La proposta prohibeix que les autoritats públiques duguin a terme qualificació social basada en IA amb finalitats generals i també prohibeix, tret d'excepcions limitades, l'ús de sistemes d'identificació biomètrica remota en temps real en espais d'accés públic amb finalitats d'aplicar la llei.

En les esmenes al projecte de Reglament es pretén ampliar substancialment la llista del títol I per incloure-hi prohibicions d'usos intrusius i discriminatoris dels sistemes d'IA, com:

- Sistemes d'identificació biomètrica remota en temps real en espais d'accés públic.
- Sistemes d'identificació biomètrica remota a posteriori, amb l'única excepció de les forces de l'ordre per perseguir delictes greus i amb la prèvia autorització judicial.
- Sistemes de categorització biomètrica que utilitzin característiques sensibles o categories especials de dades en els termes del RGPD (sexe, raça, ètnia, estatus, religió, orientació política).
- Sistemes policials predictius (basats en perfils, localització o comportaments delictius anteriors).
- Sistemes de reconeixement d'emocions en les forces de l'ordre, en la gestió de fronteres, en el lloc de treball i en les institucions educatives.
- Extracció indiscriminada de dades biomètriques de xarxes socials o gravacions de circuits tancats de televisió (CCTV) per crear bases de dades de reconeixement facial.

2. Intel·ligència artificial d'alt risc

El títol III conté normes específiques per als sistemes d'IA que comporten un alt risc per a la salut i per a la seguretat o per als drets fonamentals de les persones físiques tenint en compte no únicament la funció que desenvolupa el sistema d'IA sinó també la finalitat específica i les modalitats per a les quals s'utilitzi aquell sistema. Aquests sistemes es permeten en la mesura que compleixin determinats requisits i siguin sotmesos a una avaluació de conformitat abans

d'implementar-los i s'inscriguin en una base de dades creada i gestionada per la Comissió Europea (que ha de ser d'accés lliure i públic i fàcilment comprensible). Cada operador de la cadena de valor estaria sotmès a una sèrie d'obligacions específiques. Per exemple:

- Governança de dades: és a dir, que les dades emprades revesteixin uns certs estàndards de qualitat, supervisió, examinació de biaixos, etc.
- Seguretat i supervisió humana: en última instància sempre hi ha d'haver una persona amb capacitat de control per mitigar riscos eventuals.
- Deures de transparència: és a dir, que es descriguin les característiques del funcionament del sistema i la identitat i les dades del proveïdor.
- Inscripció en una base de dades en l'àmbit europeu: s'ha de dur a terme prèviament a la posada a disposició en el mercat.
- Superació del test de conformitat i obtenció de la certificació corresponent: s'han d'aprovar especificacions tècniques que cal complir.

3. Intel·ligència artificial de baix risc o de risc limitat

Sistemes que no comporten un alt risc per als drets i les llibertats. Inclouen determinades tecnologies de menys sofisticació o capacitat d'intrusió com ara assistents virtuals com bots. Aquests sistemes únicament estan sotmesos a un conjunt de normes de transparència adreçades a garantir que els usuaris coneixen el funcionament i les característiques, així com les implicacions inherents a l'ús d'aquests sistemes.

Respecte d'aquests sistemes, les modificacions introduïdes al text original imposen el compliment de requisits concrets de transparència, com revelar que els continguts han estat generats per una IA, dissenyar el model per evitar que generi continguts il·legals i publicar resums de les dades protegides per drets d'autor que s'hagin emprat per a l'entrenament del model en qüestió.

4. Resta de sistemes d'intel·ligència artificial

Aquests últims, en principi, no estan subjectes a cap obligació en particular, i els agents de la cadena poden triar si desitgen adherir-se a sistemes voluntaris de compliment com l'adhesió a codis de conducta voluntaris. Per consegüent, aquests sistemes queden, en principi, fora de l'àmbit d'aplicació del reglament.

Pel que fa a la coherència de la seva aplicació en relació amb la normativa en matèria de protecció de dades existent, el projecte de Reglament d'intel·ligència artificial estableix que la seva aplicació és sense perjudici de l'aplicació del RGPD i de la Directiva sobre protecció de dades en l'àmbit penal, als quals pretén complementar "amb un conjunt de normes harmonitzades aplicables al disseny, al desenvolupament i a la utilització de determinats sistemes d'IA d'alt risc i amb restriccions de determinats usos dels sistemes d'identificació biomètrica remota". En aquest sentit, el concepte de *dada biomètrica* a l'efecte de la proposta de Reglament coincideix amb la definició recollida al RGPD i a la Directiva sobre protecció de dades en l'àmbit penal.

Així mateix, un dels principis generals del projecte de Reglament aplicable a tots els sistemes d'IA és el principi de "governança de dades", segons el qual els sistemes d'IA s'han de desenvolupar i utilitzar de conformitat amb les normes existents de privacitat i protecció de dades.

El text amb les esmenes va ser ratificat el passat 14 de juny pel Ple del Parlament, però encara pot ser modificat durant els diàlegs a tres o negociacions entre les institucions europees, la Comissió, el Parlament i la Presidència del Consell (que correspon a Espanya durant el segon semestre de 2023). S'espera que el projecte de Reglament d'IA sigui aprovat al final d'any.

Dolors Priego Garcia

Lletrada de l'Assessoria Jurídica