Datum
2021-06-11

Europeiska kommissionens förslag till förordning om harmoniserade regler för artificiell intelligens

(dnr I2021/01304)

Sammanfattning

Wikimedia Sverige tackar för möjligheten att inkomma med synpunkter på Europeiska kommissionens förslag till förordning om harmoniserade regler för artificiell intelligens.

På en övergripande nivå vill vi understryka vikten av krav på transparens för AI-system, i synnerhet när systemen är finansierade med offentliga medel. Utgångspunkten bör vara att sådana system är öppen källkod.

Vi vill också framföra vikten av att följa Europeiska kommissionens princip om bättre lagstiftning, där en av huvudpunkterna är att förenkla och förbättra EU-lagstiftningen. Stora delar av de regler som här föreslås ska harmoniseras hade kunnat regleras utan en Lex specialis, exempelvis genom att uppdatera produktansvarsdirektivet. För oss är det, även om vi sympatiserar med stora delar av förslaget, något otydligt vad man egentligen vill åstadkomma. Vi förstår att det är svårt att ta den typen av kommentarer i beaktande i det här läget, men vill framföra att det för framtiden är viktigt att principen om bättre lagstiftning upprätthålls, och att inte varje ny teknisk trend i samhället blir grund till ny lagstiftning, utan att lagstiftningen hålls effektiv, enkel och teknikneutral.

Övergripande kommentarer

Transparenskrav (artiklarna 13 och 52)

AI-system finansierade med offentliga medel ska vara öppen källkod.
Artificiell intelligens (AI) är och kommer att kunna vara en stor tillgång för många delar av samhället. Samtidigt finns det flera olika områden där AI också medför stora risker, inte minst för den personliga integriteten. Ett antal sådana användningsområden listas i Annex III. Artikel 13, särskilt i första stycket, fastställer i sin tur att AI-system med hög risk måste ha en tillräcklig nivå av transparens för att användare ska kunna förstå systemets output och hur det ska användas. Även för vissa AI-system, det som kallas “certain AI systems” i lagförslaget, uppställs ett antal transparenskrav, bland annat att man ska medvetandegöras om att det är AI som man har att göra med när så är fallet.

Wikimedia Sverige instämmer med dessa skrivelser och vill betona deras betydelse. I synnerhet när det gäller AI-system som har finansierats med offentliga medel bör dock kraven ställas högre. En viktig anledning till det är att flera sådana AI-system, vilket exempelvis även anförs i skäl 37, kan ha en betydande påverkan på människors liv, med kapacitet att påverka människors grundläggande rättigheter, såsom rätten till mänsklig värdighet och socialt skydd. För system baserade på AI finansierade med offentliga medel bör utgångspunkten vara att resultatet blir öppen källkod. För att medborgare ska kunna förstå vad som händer och hur det system som har möjlighet att påverka deras liv är konstruerat, vilket är en grundläggande rättighet, är öppen källkod det enda tillförlitliga sättet. Det innebär rent konkret att algoritmerna som AI-systemet använder sig av också måste vara öppna. Vi vill också framföra vikten av transparens vad gäller datan som används, även om denna data av integritetsskäl kan vara omöjlig att begära ut. Det bör införas mekanismer som kan tillförsäkra tillräcklig granskning av datan som används. Källkoden till programvara som används eller finansieras av myndigheter eller offentlig förvaltning ska dock vara öppen.

Datastyrning och systematiska fel (“bias”) (artiklarna 10 och 14)

Wikimedia Sverige poängterar betydelsen av att aktörer motverkar systematiska fel
Wikimedia Sverige noterar skrivelserna i artiklarna 10.2(f) och 14.4(b), om att AI-system med hög risk ska åläggas att identifiera potentiella systematiska fel eller bias, liksom att vara medvetna om riskerna med att i alltför hög grad förlita sig på resultatet från ett AI-system (“automation bias”). Skäl 44 förtydligar även att hög kvalitet på data är helt central, i synnerhet när det rör sig om tekniker som försöker träna modeller, eftersom systemet annars har potential att utgöra grunden till diskriminering av sådant slag som är förbjuden enligt unionsrätten.

Bias är dock inte bara något som drabbar användare av hög risk-system. Även i de fall som maskininlärning används på Wikimediaplattformarna blir bias snart tydliga. Exempelvis pågår experiment med maskininlärning för att föreslå rubriker när nya artiklar skrivs på Wikipedia. En slutsats från de försöken är att när någon vill skriva en artikel om en manlig person, föreslår mjukvaran ett stycke om “karriär” först. När någon vill skriva en artikel om en kvinnlig person föreslår mjukvaran istället ett stycke om “familj” först. Det bygger på de bias som redan, dessvärre, finns bland de miljontals biografier som finns på Wikipedia.

Även vad gäller programvaran Wikispeech (se mer nedan) som föreningen håller på att utveckla, där maskininlärning används för att ta fram röster till en talsyntes, märks snart vådan av bias. Om enbart taldata från vissa dialekter, kön eller andra egenskaper samlas in kommer det att återspeglas i rösterna som läses upp. Bias i data medför på så vis en sämre slutprodukt, dels genom att det förstärker och återproducerar det bias som redan finns, och dels för att det kan göra det svårare för folk som är vana att höra ett visst uttal att använda mjukvaran om det inte finns en röst med det uttalet. Liknande problem finns vad gäller taligenkänning. Om mjukvaran bakom taligenkänningen är dåligt tränad, eller tränad på data med systematiska fel, blir själva taligenkänningen obrukbar för de som har ett annorlunda uttal, eller någon form av talfel.

Wikimediaplattformarna använder crowdsourcing för att försöka att få bukt med problemet, medvetna om hur centralt det är. Wikimediarörelsen har även skrivit under Torontodeklarationen, där just motverkandet av bias är en viktig del. Samtidigt inser vi problemen med att gå för långt i lagstiftning på detta område. En möjlighet vore att föreslå någon form av *duty of due diligence* (krav på skälig aktsamhet), för utvecklare av system baserade på artificiell intelligens. Utan att inskränka möjligheterna till innovation skulle det på sikt kunna motverka AI-system som upprätthåller eller förstärker systematiska fel.

Kommentarer specifika för Wikimedia Sveriges verksamhet

EU bör hitta lösningar på låsningen mellan AI, taldata och GDPR
Wikimedia Sverige har under några år, tillsammans med bland annat Kungliga tekniska högskolan och med stöd av Post- och telestyrelsen, utvecklat Wikispeech. Wikispeech är en text-till-tal-lösning för MediaWiki, mjukvaran som bland annat Wikipedia bygger på, men som också används av flera svenska myndigheter och även internationellt. Wikispeech innehåller dessutom en taldatainsamlare, där röstdata crowdsourcas för att skapa nya röster. Den data som samlas in kan dock även användas för andra lösningar som bygger på maskininlärning.

Bättre data möjliggör en taligenkänning som fungerar lika bra på manliga och kvinnliga röster, stark dialekt eller brytning. Det möjliggör också uppläsning på olika dialekter av ett språk, samt en unik möjlighet att utveckla Wikispeech på mindre språk, språk som inte har tillräckligt stor spridning för att kommersiellt intresse för talsynteser ska föreligga. Inledande försök sker exempelvis på baskiska, men med stor potential att relativt snabbt innefatta en stor mängd språk. Taldatainsamlaren, och byggandet av nya språkmoduler, bygger på AI-system.

Här finns dock en stor konflikt gentemot artikel 17 i GDPR, om rätten att bli bortglömd. Relationen mellan taldata och GDPR är komplicerad, vilket även flera forskare och institutioner internationellt har konstaterat. I en mindre vetenskaplig studie torde det vara möjligt att ta bort taldata om en enskild som tidigare har gett medgivande drar tillbaka detta. Men Wikispeech är ett försök att skapa en multimodal infrastrukturlösning. I ett sådant sammanhang skulle hela infrastrukturen kunna undermineras genom att en enskild drar tillbaka sitt medgivande. I dagsläget innebär det rent konkret att försök till crowdsourcing av ny data, till gagn för mänskligheten, är stoppad. Det finns, med GDPR, för stora risker.

Vi vill medvetandegöra regeringen om detta lås, och framföra att det för möjligheten till taldatainsamling av en större omfattning vore angeläget med ett försök till lösning på europeisk nivå. Det är dock självfallet viktigt att det, med en lösning av problemet, inte ska kunna gå att begära ut grunddata, med hänvisning till integritetsskydd. En möjlighet skulle kunna vara att förtydliga i förordningen om harmoniserade regler för artificiell intelligens att taldata som används till AI-system uppfyller GDPR så länge inte enskilda data kan begäras ut.