

Rregullimi ose Ndalimi? Inteligjenca Artificiale do të çojë në zgjedhje të vështira

Article by Alexandre Erler
April 20, 2022

Akti i Inteligjencës Artificiale i Bashkimit Evropian synon të rregullojë aplikimet e reja të AI në përputhje me “vlerat e BE-së” dhe të parandalojë aplikimet abuzive dhe të dëmshme të kësaj teknologjie në zhvillim ndërsa ajo bëhet më e përhapur. Por një vështrim më i afërt i përdorimeve të mundshme që mund të bien në kundërshtim me këtë legjislacion tregon se linja midis rregullimit dhe ndalimit mund të jetë e ndërlikuar për t’u tërhequr, siç shpjegon edhe Alexandre Erler.

Akti i Inteligjencës Artificiale (AIA), i prezantuar në prill të këtij viti nga Komisioni Evropian dhe ende po kalon rrugën e tij në Parlamentin Evropian, propozon një qasje “të bazuar në rrezik” për rregullimin e sistemeve të AI, që synon ta përfaqëzojë atë me “vlerat evropiane” dhe e bëjnë Evropën një lider në “Inteligjencën Artificiale të besueshme”. Rregullorja e propozuar i ndan aplikimet e AI në katër kategori të ndryshme bazuar në vlerësimin e shkallës së rrezikut: ato që paraqesin rrezik të papranueshëm, rrezik të lartë, rrezik të kufizuar dhe rrezik të ulët ose minimal. Ndërsa aplikimet në kategorinë e parë duhet të ndalohen krejtësisht, ato në tre kategoritë e mëpasshme duhet të rregullohen me nivele në rënie graduale rreptësie.

AIA ka qenë tashmë objekt i shumë debateve publike. Ndoshta çuditërisht, ajo ka marrë kritika nga palët kundërshtarë: disa sugjerojnë se kufizimet që ajo vendos mund të jenë shumë të rrepta, ndërsa të tjerë shqetësohen se është ajo tepër lejuese dhe e hapur për abuzime. Inspektimi më i afërt i disa prej aplikacioneve të AI të kategorizuara si “rrezik i papranueshëm”, dhe për këtë arsye të veçuara si të ndaluara, mund të na ndihmojë të kuptojmë më mirë dilemat që mund të lindin. Tre aplikacione të tilla shqyrtohen më në detaje në paragrafët e mëposhtëm: sistemet që do t’u shkaktonin dëm përdoruesve duke manipuluar sjelljen e tyre, “vlerësimi social” i një natyre të përgjithshme të drejtuar nga shteti dhe përdorimi i sistemeve biometrike në distancë “në kohë reale” për zbatimin e ligjit.

Our latest edition - Moving Targets: Geopolitics in a Warming World - is out now.

It is available to read online & order straight to your door.

[READ & Order](#)

Përcaktimi i fushës së aplikimeve “të dëmshme” të AI

Shqetësimet rreth kufizimeve tepër të rrepta mund të lindin si përgjigje ndaj ndalimit të

propozuar nga AIA të aplikacioneve të AI “që synojnë të shtrembërojnë sjelljen njerëzore, ku ka të ngjarë të ndodhin dëme fizike ose psikologjike” dhe sistemeve që do tu ” vendosin komponentë që ngacmojnë subkoshencën që individët nuk mund të perceptojnë ose të shfrytëzojnë dobësitë e fëmijëve dhe njerëzve për shkak të moshës, paaftësisë së tyre fizike ose mendore” (neni 16). Një gjuhë e tillë mund të shihet si e padëshirueshme e paqartë, duke mbuluar potencialisht praktika që një shoqëri liberale nuk duhet të dëshirojë t’i ndalojë.

Konsideroni për shembull që kompanitë e mediave sociale si Meta (ish-Facebook) janë raportuar se po i projektojnë qëllimisht produktet e tyre për t’i bërë përdoruesit të varuar prej tyre. Për aq sa varësia ndaj një aplikacioni të mediave sociale mund të interpretohet në mënyrë të besueshme si një formë e dëmtimit psikologjik, a duhet të nxjerrim përfundimin se AIA do të kërkonte ndalimin e një platforme të tillë si Facebook, përveç nëse kompania do të ishte e gatshme të rishikonte veçoritë ofenduese? Apo, siç kanë pyetur disa, a të kërkonte ndalimin e Instagramit, bazuar në ndikimin e tij negativ në shëndetin mendor dhe imazhin e trupit të disa adoleshentëve? Në dritën e shkurtimit të lirive personale që ato do të sillnin, implikime të tilla do të dukeshin problematike (megjithëse kufizimet shtesë të bazuara në moshë mund të jenë një mbrojtje e mirë). Megjithatë, nuk ka gjasa që ato të pasqyrojnë me saktësi qëllimet që qëndrojnë pas kësaj dispozite të AIA-së, e cila në vend të kësaj ne mund të supozojmë se synon praktika më të dyshimta si “modelet e errëta” (d.m.th., metoda mashtruese të përdorura nga disa faqe interneti dhe aplikacione për të mashtruar përdoruesit të bëjnë gjëra që mund të mos kenë dashur t’i bëjnë). Megjithatë, do të ishte e dëshirueshme një qartësi më e madhe për sa i përket fushës së zbatimit të këtij ndalimi të veçantë.

Përderisa varësia ndaj një aplikacioni të mediave sociale mund të interpretohet si një formë e dëmtimit psikologjik, a duhet të arrijmë në përfundimin se AIA duhet të kërkonte ndalimin e një platforme të tillë si Facebook?

A duhet të kemi frikë nga “rezultati social” i aktivizuar nga AI?

Ndalimi i AIA-s për “pikëzimin social të personave fizikë për qëllime të përgjithshme nga autoritetet publike ose në emër të tyre” (neni 17), është një tjetër zhvillim i mundshëm që supozohet se përfaqëson një “rrezik të papranueshëm”, dhe ka shkaktuar reagime krejt të ndryshme deri më tani. Disa e kanë karakterizuar atë – krahas deklaratës së fundit të përbashkët SHBA-BE për AI – si një përpjekje për të pretenduar nivelin e lartë moral ndaj Kinës, ndërkohë që synon të targetojë potencialisht një fantazmë, bazuar në një keqinterpretim të sistemit të kredisë sociale të Kinës. Megjithatë, ky ndalim i propozuar është i kritikuar gjithashtu si i pamjaftueshëm: grupet e shoqërisë civile kanë argumentuar kështu se ai duhet të shtrihet në subjektet private.

Është e vërtetë që keqkuptimet rreth sistemit të kredisë sociale të Kinës janë relativisht të zakonshme në Perëndim: për shembull, në formën e tij aktuale, sistemi është shumë më i

copëzuar sesa supozohet të jetë zakonisht dhe nuk shpërblen apo ndëshkon çdo qytetar kinez për sjelljen e tyre të përditshme bazuar në një pikë të vetme të përgjithshme shoqërore. Megjithatë, kjo nuk do të thotë se nuk ka justifikime më të mira për të dashur për të parandaluar futjen e vlerësimit social të “qëllimit të përgjithshëm” sesa një demonizim të sistemit kinez. Ndarja e shpërblimeve dhe e ndëshkimeve të bazuara në një masë të përgjithshme të “virtytit” ose “besueshmërisë” së një qytetari duket të jetë një pikë përfundimtare e natyrshme e eksperimentit aktual kinez dhe nuk është e paimagjinueshme që ai të mund të shfaqet edhe në Perëndim. Pesë vjet më parë, një aplikacion i diskutuar shumë i quajtur People u lëshua nga një kompani kanadeze, duke i lejuar përdoruesit të vlerësojnë të tjerët (pa pëlqimin e tyre paraprak) në tre kategori kryesore: personale, profesionale dhe romantike. Ndërsa kjo ende nuk është aq e unifikuar sa një rezultat i vetëm i përgjithshëm social, ishte ende mjaft afër për të lindur zemërim, duke bërë që aplikacioni të dobësohej.

Çfarë do të ishte saktësisht e kundërshtueshme për një sistem të tillë vlerësimi social? Në fund të fundit, disa mund të argumentojnë se përdorimi i pikëve si bazë për të trajtuar individët pak a shumë në mënyrë të favorshme është tashmë një praktikë e zakonshme edhe në demokracitë perëndimore. Shembujt përfshijnë vlerësimet e kredive, programet e besnikërisë komerciale, rishikimet e performancës së punonjësve dhe sistemet e pikëve – të manaxhuara nga shteti – për shkeljet e drejtimit të automjeteve. Nëse praktika të tilla janë të pranueshme, pse vlerësimi social i “qëllimit të përgjithshëm” të qytetarëve të një vendi është përtej zbehjes? Një arsye është potenciali i tij i madh për abuzime. Nëse qëllimi i një sistemi është të identifikojë qytetarët përgjithësisht “të mirë” ose “të besueshëm”, në vend të sjelljeve pozitive ose negative të një lloji shumë më specifik, ai duket se u ofron liderëve me prirje autoritare një justifikim të përshtatshëm për të fshehur elementë të papërshtatshëm në skemën e vlerësimit. Për shembull, zbritje të pikëve për ndarjen e “lajmeve të rreme” – duke e trajtuar kategorinë si, të gjithë fjalimet që autoritetet nuk pëlqejnë. Vlerësimet e përgjithshme të “virtytit” të qytetarëve rrezikojnë gjithashtu të inkurajojnë nivele gjithnjë e më ndërhyrëse të mbikëqyrjes së qeverisë, duke pasur parasysh gamën e gjerë të sjelljeve private që mund të konsiderohen të rëndësishme për këtë qëllim.

Edhe nëse masat mbrojtëse mund të imponohen disi për të parandaluar tejkalime të tilla, një problem më themelor me konceptin e një rezultati të përgjithshëm shoqëror është padyshim natyra e tij me mënyrën e vjetër dhe shënjuese. Kjo është më e qartë në rastet kur rezultati i ulët i dikujt do të bëhej publik, me synimin për ta turpëruar ata në sjellje “më të mirë”. Megjithatë, shqetësimi do të zbatohet ende nën një sistem që nuk ua zbulonte rezultatin e një personi palëve të treta, përveç furnizuesve zyrtarë të shpërblimeve dhe ndëshkimeve. Të paktën nga një perspektivë liberale, ekzistenca e plotë e një rezultati të tillë të përgjithshëm do të thotë se shteti po u dërgon një mesazh patronizuese në mënyrë të padrejtë individëve për statusin e tyre si qytetarë “të mirë” ose “të këqij”, diçka që nuk kërkohet për të nxitur sjelljen të dobishme shoqërore. Thënë kjo, theksi mbi respektimin e lirisë individuale vë gjithashtu në pikëpyetje përshtatshmërinë e shtrirjes së ndalimit të vlerësimit të përgjithshëm shoqëror tek aktorët privatë. Me sa duket, njerëzit duhet të jenë të lirë të përdorin aplikacione si People; çështja kryesore është nëse mund të sigurohemi që ata ta bëjnë këtë në mënyrë vullnetare dhe të informuar, pa pasur nevojë të vuajnë kosto të papranueshme personale nëse ata refuzojnë t’i nënshtrohen një vlerësimi të tillë shoqëror (le të themi, jo sepse ishte bërë normë).

Disa mund të argumentojnë se përdorimi i pikëve si bazë për të trajtuar individët pak a shumë në mënyrë të favorshme është tashmë një praktikë e zakonshme edhe në demokracitë perëndimore.

Sistemet biometrike në distancë për zbatimin e ligjit: ndalojnë apo rregullojnë?

Një pikë e tretë mosmarrëveshjeje ka të bëjë me ndalimin e AIA-s për “përdorimin e sistemeve të identifikimit biometrik në distancë “në kohë reale” në hapësira me qasje publikisht për qëllime të zbatimit të ligjit” (neni 18). Ndalimi pranon përjashtime në rastet e tre llojeve specifike: 1) kërkimin e fëmijëve të zhdukur dhe viktimave të tjera të mundshme të krimit; 2) parandalimin e kërcënimeve për jetën ose sigurinë fizike, siç janë sulmet terroriste; dhe 3) identifikimin e të dyshuarve për krime që parashikojnë dënime maksimale deri në tre vjet ose më shumë. Disa, duke përfshirë edhe organizatat e shoqërisë civile, kanë bërë thirrje për heqjen e këtyre përjashtimeve dhe për shtrirjen e ndalimit përtej përdorimeve “në kohë reale” dhe për aktorë të ndryshëm nga zbatimi i ligjit.

Çështja nëse duhet të ndalohen plotësisht, ose më saktë të rregullohen me kujdes, sistemet biometrike të largëta si teknologjia e njohjes së fytyrës, është pa dyshim një çështje e vështirë. Dëshira për të mbrojtur privatësinë e qytetarëve dhe për të parandaluar keqpërdorimet dhe arrestimet e gabuara bazuar në algoritme potencialisht të pasakta dhe të njëanshme, mund të bëjë që një ndalim i drejtpërdrejtë të duket si opsioni më i mirë. Megjithatë, duhet pranuar gjithashtu se një politikë e ndalimit të pacaktuar ka të ngjarë të përfshijë kosto të konsiderueshme oportune në drejtim të aftësisë së shoqërisë për të parandaluar krimin. Prandaj, nevojitet një debat publik thelbësor për të përcaktuar se në cilat kushte të sakta, nëse ka, teknologji të tilla mund të përdoren siç duhet, dhe cilat zgjidhje teknike mund të jenë të disponueshme për të ndihmuar në adresimin e problemeve të saktësisë dhe paragjykitimit. Një moratorium i përkohshëm për përdorimin e tyre mund të konsiderohet ndërkohë që po zhvillohet ai debat.

Në përgjithësi, diskutimet rreth AIA ilustrojnë sfidat e përfshira në përpjekjen për të balancuar të drejtat e njeriut dhe “vlerat e tjera evropiane” kundër nevojës për të promovuar inovacionin teknologjik dhe përfitimet ekonomike dhe sociale që shoqërojnë atë në Bashkimin Evropian. Ndërsa shkalla në të cilën bërësit e politikave të BE-së duhet të përballen me shkëmbime ndërmjet këtyre qëllimeve të ndryshme është në vetvete një çështje mosmarrëveshjeje, nevoja e përgjithshme për të balancuar konsideratat potencialisht konkurruese nuk mund të mohohen. Nëse AIA i bën gjërat siç duhet në këtë drejtim, mbetet një çështje për debate të mëtejshëm, megjithatë, të paktën, kjo përpjekje mund të lavdërohet për synimin drejt dhe qëllimin e duhur.

Ky artikull u botua për herë të parë nga zyra e Heinrich-Böll-Stiftung në Hong Kong.



Alexandre Erlar is a philosopher studying the ethical and political implications of emerging technologies. He is currently a Research Assistant Professor in Philosophy and Bioethics at the Chinese University of Hong Kong.

Published April 20, 2022

Article in Albanian

Translation available in English

Published in the *Green European Journal*

Downloaded from <https://www.greeneuropeanjournal.eu/rregullimi-ose-ndalimi-inteligjenca-artificiale-do-te-coje-ne-zgjedhje-te-veshtira/>

The Green European Journal offers analysis on current affairs, political ecology and the struggle for an alternative Europe. In print and online, the journal works to create an inclusive, multilingual and independent media space. Sign up to the newsletter to receive our monthly Editor's Picks.