



**Institutet för
framtidsstudier**

Transparens och reglering av AI på digitala plattformar

Policy brief

Sammanfattning

Artificiell intelligens (AI) påverkar hur information sprids på digitala plattformar, ofta utan att användarna är medvetna om det. Tekniken har både positiva och negativa effekter. Den kan underlätta användarupplevelsen men riskerar att förstärka desinformation och påverka demokratiska värden.

Denna policy brief bygger på resultat från två vetenskapliga studier som visar att AI ofta sprids osynligt genom uppdateringar av befintliga teknologier. Dessutom analyserar denna skrift hur EU:s Digital Services Act (DSA) försöker skapa transparens och ansvar kring hög-räckvidds AI. För att säkerställa AI-teknikens samhällsnytta krävs tydligare reglering, rätt sorts transparens för tillsynsmyndigheterna och att användardata är tillgängliga för forskare och reglerare.

Bakgrund

AI:s inflytande på samhället och dess spridning genom digitala plattformar är komplex och ofta osynlig för användaren. Tekniken används i de rekommendationssystem som styr vad vi ser på sociala medier, vilka annonser som presenteras på webbsidorna vi besöker och vilka nyheter som medieplattformarna prioriterar för oss.

Högräckvidds-AI är system som tillämpas på en mycket omfattande skala, där sociala mediers rekommendationssystem för användarna är ett exempel.

AI integreras ofta genom "infusion", att nya AI-funktioner läggs till genom uppdateringar av redan etablerade plattformar och tjänster, utan att användarna informeras. Exempelvis har algoritmer som rekommenderar innehåll i videoappar som YouTube och Netflix blivit centrala för användarupplevelsen, med algoritmiska rekommendationer som styr upp till 80 % av det visade innehållet.

Denna spridning skapar flera utmaningar:

€ **Autonomi och insyn**

Användare har begränsad insyn i hur deras beteende påverkas av AI. Algoritmer optimerar ofta för att maximera engagemang för att stödja annonsering snarare än användarnas välmående, vilket kan leda till oönskade konsekvenser som överdriven skärmtid eller desinformation.

€ **Etiska och sociala dilemman**

AI kan förstärka existerande ojämlikheter genom partiskhet i de data som insamlas eller i algoritmer som favoriserar vissa grupper. Dessutom har forskningsstudier pekat på att vissa algoritmer visat sig bidra till spridning av skadligt innehåll, inklusive extremism och desinformation.

Digital Services Act (DSA) är en EU-förordning (EU 2022/2064) för att skapa en inre marknad för digitala tjänster, genom att harmonisera reglerna och skapa en tillsynsprocess. Det svenska namnet för DSA är förordningen om digitala tjänster.

DSA syftar till att skapa mer transparens och ansvarighet för mycket stora onlineplattformar, så kallade Very Large Online Platforms (VLOP:er). Förordningen introducerar krav på transparens för hur plattformarna använder rekommendationssystem.

Den ålägger VLOP:er att identifiera risker med sin hög-räckvidds AI, inklusive algoritmisk snedvridning och spridning av skadligt innehåll. Europeiska kommissionen utövar tillsyn av VLOP:er.

Trots dessa framsteg kvarstår frågor kring myndigheters kapacitet att följa upp dessa krav och säkerställa att plattformarna efterlever reglerna. Forskare och allmänheten får begränsad tillgång till plattformarnas data, vilket försvårar utvärderingen av AI:s samhällseffekter.

AI:s potential att bidra till samhället är stor, men dess inflytande kräver också en reglering som balanserar innovation med skydd för individens rättigheter och samhällsstrukturen. Att förstå och hantera AI:s effekter är avgörande för att säkra dess roll som en kraft för det gemensamma bästa.

Rekommendationer

Hur kan AI regleras på ett sätt som säkerställer att den gynnar samhällets bästa? AI-tillämpningar med hög räckvidd behöver en tydligare reglering. Den bör göra användarna mer uppmärksamma på skillnaderna i integritet mellan olika plattformar och tjänster, ge tillsynsmyndigheten möjlighet att fullt ut utföra den riskbedömning som DSA innehåller samt ge forskare bättre datatillgång. Företagen kan i nuvarande lagstiftning hänvisa till att algoritmer och data är affärshemligheter, och inte ge åtkomst till dem.

Öka transparensen

DSA ger en modell för att offentliggöra strukturerad information om de rekommendationssystem som används på de stora plattformarna. Inför krav på att plattformar tydligt informerar tillsynsmyndigheter och användare om hur AI används i rekommendationssystem och möjliggör användardrivna val av algoritmisk påverkan. VLOP:er bör ansöka om att hålla viss information konfidentiell om det ligger i allmänhetens intresse (exempelvis för att skydda algoritmer för spamfilter).

Stärk tillsynen

Ge Europeiska kommissionen samt de nationella organen för samordning för digitala tjänster (Digital Services Coordinators, DSC) möjlighet att kunna implementera och följa upp regleringen i DSA.

Främja forskning och dataåtkomst

Stärk kraven på att säkerställa att betrodda forskare i sin forskning får insyn i plattformarnas algoritmer och dess påverkan på samhällsnivå.

Utbilda användarna

Det är svårt att ge användare insyn i hur algoritmerna bakom rekommendationssystem fungerar och hur dessa påverkar användarnas beteenden. Om skillnader i integritetsskydd och datainsamling görs tydligare från utvärderingarna av systemrisker kan det bli lättare för dem att bestämma sin egen användning.

Referenser

Engström, E. & Strimling, P. (2020). *Deep learning diffusion by infusion into preexisting technologies - Implications for users and society at large*. *Technology in Society*. 63, 101396. <https://doi.org/10.1016/j.techsoc.2020.101396>

Söderlund, K. Engström, E. Haresamudram, K. Larsson, S. Strimling, P. (2024). *Regulating high-reach AI: On transparency directions in the Digital Services Act*. *Internet Policy Review*, 13(1). <https://doi.org/10.14763/2024.1.1746>

För mer information

- € Forskningsprojektet **Ramverk för att förutsäga hur artificiell intelligens sprids** <https://www.iffs.se/forskning/forskningsprojekt/ramverk-for-att-forutsaga-hur-ai-teknik-sprids/>
- € **Pontus Strimling**, Projektledare Institutet för framtidsstudier. Email: pontus.strimling@iffs.se