

AI-policy – Lexicon Interactive

1. Syfte

Syftet med vår AI-policy är att vara vägledande för våra medarbetare kring användningen av AI i yrkesrollen och säkerställa att Lexicon Interactive använder AI på ett ansvarsfullt, transparent och lagenligt sätt, i linje med våra värderingar samt gällande och kommande regelverk – inklusive EU:s AI Act.

Målet är att medarbetarna, genom att förstå policyns innehåll, ska kunna göra medvetna val kring användningen av AI och minska de risker som kan uppstå för sig själva och företaget genom en felaktig eller oansvarig användning av AI.

Policyn syftar också till att erbjuda ett stöd i det egna förhållningssättet av AI.

2. Omfattning

Den här policyn gäller för samtliga medarbetare vid Lexicon Interactive, samt för alla fysiska eller juridiska personer som genom avtal eller annan överenskommelse utför uppdrag för företagets räkning.

Policyn omfattar all användning av AI-system inom företaget, med särskilt fokus på generativa AI-tjänster, exempelvis text-, bild- eller kodgenererande verktyg.

3. Definitioner

- **Generativ AI**

AI-system som kan skapa nytt innehåll, exempelvis text, bild, ljud eller kod baserat på skrivna eller på annat sätt uttryckta instruktioner.

- **AI Act**

EU:s lagstiftning som reglerar utveckling, distribution och användning av AI-system inom EU, med syfte att främja säker och etisk AI.

4. Överensstämmelse med AI Act

Lexicon Interactive använder i dagsläget AI-system som bedöms tillhöra kategorin *låg till begränsad risk* enligt AI Act. Vi bedriver ingen AI-utveckling och använder inte AI inom områden som klassas som högrisk, exempelvis rättsväsende, rekrytering eller känslig dataanalys. Trots detta följer Lexicon Interactive relevanta krav enligt AI Act, inklusive:

Transparens – märkning av AI-genererat innehåll

Vi försöker att inte luras och om det påverkar trovärdigheten är vi tydliga med att berätta att något är AI-genererat. För att säkerställa transparens i vår kommunikation ska AI-genererat eller AI-behandlat material märkas tydligt i följande fall:

Text

- Märkning krävs när texten är helt eller till största del AI-genererad och inte granskad.
- Märkning krävs inte när texten är granskad eller används internt utan koppling till extern kommunikation eller kundupdrag.

Bilder och grafik

- Märkning krävs när bilden är helt AI-genererad och en bärande del av kommunikationen, särskilt om den föreställer människor.
- Märkning krävs inte om:
 - AI endast använts för mindre bearbetning (t.ex. friläggning, färgjustering).
 - Bilden är dekorativ och inte innehåller människor.
 - Innehållet används internt.

Ljud och tal

- Märkning krävs när ljudet är AI-genererat och inte granskats, eller om det ger intryck av att vara en verklig människa.
- Märkning krävs inte om:
 - Talet är anonymt och granskats.
 - Innehållet är för internt bruk.

Film och avatarer

- Märkning krävs när filmen eller avataren är AI-genererad och:
 - utgör en bärande del av kommunikationen.
 - föreställer människor eller avatarer baserade på människor.

- Märkning krävs inte om:
 - Innehållet inte visar människor och inte är kommunikativt bärande. Till exempel klippbilder.
 - Materialet används internt.

Generella metoder

- Vi använder ikonen ✨ ("sparkles") och/eller förklarande text för att signalera AI-användning.
- I alt-texter kan tillägg göras, exempelvis: *"Behandlad eller genererad av AI."*
- Märkningen kan även samlas på en separat plats om det förbättrar användarupplevelsen, men ska alltid vara tydlig.

Mänsklig kontroll

- Om möjligt ska allt AI-genererat material granskas av en människa innan det används externt. Där det inte är möjligt, exempelvis i generativt lärande, ska vi i hög utsträckning försöka förhindra olämpligt eller felaktigt innehåll genom instruktioner och promptning till AI-tjänsten.

Datasäkerhet

- Vi matar inte in känslig affärsinformation, känsliga, sekretessbelagda eller personliga uppgifter i AI-system, såvida inte systemet är särskilt godkänt för det ändamålet.
- Personuppgifter ska behandlas i enlighet med gällande lagstiftning.
- API-nycklar, användaruppgifter och likvärdig information ska skyddas från obehörig åtkomst och användning.
- Misstänkt aktivitet, incidenter och identifierade risker ska rapporteras till närmaste chef.

5. Etiska riktlinjer för användning av generativ AI

För att främja en trygg och ansvarsfull AI-kultur gäller följande principer:

- AI får inte användas för att skapa vilseledande, diskriminerande eller oetiskt innehåll.
- Den som framställer AI-genererat eller AI-bearbetat material ansvarar för att granska och anpassa det.
- AI får inte ersätta mänskliga beslut i känsliga ärenden.

- Användningen av AI ska alltid bidra till ökad kreativitet, effektivitet eller kvalitet – inte minska det mänskliga omdömet värde.

6. Riktlinjer för promptning och granskning

- Formuleringar till AI ska vara tydliga, sakliga och fria från känsliga uppgifter, såväl egna som andras.
- AI-genererat material som ska publiceras externt ska alltid granskas ur aspekterna fakta, tonalitet, tillgänglighet och risk för bias, av den som framställer materialet.
- För utbildningsmaterial, texter eller presentationer som skapats med hjälp av AI rekommenderas att vi i viss mån redovisar AI:s roll i processen.

7. Kompetens och stöd

Alla medarbetare som använder AI ska ha grundläggande kunskap om AI Act och företagets policy. Intern utbildning och stödmaterial tillhandahålls via aktuella kommunikationskanaler.

8. Ansvar och efterlevnad

- Utvecklingschef ansvarar för att uppdatera policyn i takt med förändrad lagstiftning och den tekniska utvecklingen. En generell översyn sker årligen.
- Alla medarbetare ansvarar för att följa riktlinjerna.

9. Framtidssäkring

Vi följer utvecklingen av AI Act och andra internationella regelverk. Policyn uppdateras löpande för att säkerställa fortsatt överensstämmelse med lagkrav, branschpraxis och företagets egna etiska riktlinjer.