Begeleidende leidraad bij het "Intern beleid voor het gebruik van AI"

1. Achtergrond

Onder deze hoofding willen we op een heel algemene manier toelichten welk het ruimere wettelijk kader is waarmee rekening moet gehouden worden bij het gebruik van AI-systemen binnen de onderneming. Het is tegen de achtergrond van de onder deze titel vermelde regels, dat de eigenlijke beleidsregels werden opgesteld.

2. Toepassing

Het is van belang te onderlijnen dat de toepassing van dit beleid niet beperkt is tot een specifieke groep bestemmeling (bv enkel personeelsleden). Dat maakt dat rekening moet gehouden worden met het gegeven dat het beleid gecommuniceerd wordt op een manier dat alle mogelijke betrokkenen er kennis van kunnen nemen.

3. Contactgegevens

AI is een domein waarin vandaag onmogelijk alles reeds in detail kan uitgewerkt worden. De evoluties volgen elkaar snel op en de meeste onderneming zijn ook nog volop zoekende voor de implementatie van de gewenste AI-systemen. Dat impliceert dat er rekening mee gehouden moet worden dat het omgaan met AI nog verdere vragen kan oproepen bij de gebruikers, aan wie een gecentraliseerd contactpunt dient te worden verstrekt.

Voor een aantal vragen over het omgaan met AI-Systemen zal een opportuniteitstoets (moeten) doorgevoerd worden vanuit de onderneming. Het is dan ook wenselijk dat de onderneming iemand aanduidt op bestuurs- of managementniveau die de vragen niet enkel vanuit technisch, maar ook vanuit bedrijfsstrategisch of opportuniteitsoogpunt kan beoordelen.

Daarnaast kan, optioneel, ook geopteerd worden om voor vragen van technische aard een technisch aanspreekpunt te voorzien, bij voorbeeld in de persoon van de interne of externe netwerkbeheerder of IT-verantwoordelijke.

4. Beleidsregels

4.1. Welke generieke AI-systemen mag de gebruiker aanwenden voor de uitvoering van zijn taken?

In de eerste plaats is van belang voor ogen te houden dat de meeste AI-Systemen die gratis aangeboden worden, in hun gebruiksvoorwaarden stellen dat het systeem de ingevoerde gegevens ook zelf mag gebruiken, bijvoorbeeld om het systeem verder te trainen. Dat creëert mogelijk gegevenslekken of het publiek worden van gevoelige informatie. Het is dan ook aan de onderneming om een beoordeling te maken welke systemen gebruikt worden, onder meer op basis van de afweging welke systemen voldoende garanties bieden op de vertrouwelijkheid van de eigen gegevens. Hieruit volgt dat enkel de onderneming beslist welke AI-systemen binnen het kader van haar activiteiten gebruikt worden, waar de gebruikers zich naar te schikken hebben.

4.2. Waarvoor kunnen de Toegelaten AI-Systemen gebruikt worden?

Onder deze titel is het vooral van belang aan te geven dat een AI-systeem in principe niet in de plaats komt van de kennis en ervaring van de gebruiker, maar dat het voor deze laatste enkel een hulpmiddel is waar die in principe de output van moet evalueren.

De beleidsregels onderkennen de mogelijkheid van AI-systemen om het werk te optimaliseren en stellen dat dan ook als algemene regel voorop. Een medewerker vervult echter zijn taak niet, louter door deze in te voeren in een AI-systeem en de output van dergelijk systeem vervolgens als de eigen bevindingen te presenteren. Een evaluatie van de gegenereerde AI-resultaten is cruciaal.

Het voorgaande kan anders zijn, in bepaalde specifieke contexten. Richtlijnen daarbij zijn dat het gaat om gebruik van een AI-tool voor een dienstverlening die niet tot de kerntaken van de onderneming behoort en waarover transparant gecommuniceerd is met de klant. In dergelijk geval kan de verantwoordelijkheid wel beperkt worden.

> Voorbeeld: Op verzoek van een klant wordt een doorlichting gemaakt van een bestaande verzekeringsportefeuille. Dat rapport behoort tot de kerntaak van de onderneming en het gebruik van een AI-systeem bij de opmaak ervan kan ondersteunend zijn, maar ontslaat de onderneming – en bij uitbreiding de gebruiker van het systeem – niet van de verantwoordelijkheid om de inhoud van het rapport te verifiëren en te valideren.
> De klant ontvangt het rapport en vraagt vervolgens of hiervan ook een vertaling kan bezorgd worden. Indien transparant met de klant afgesproken wordt dat een AI-gegenereerde vertaling kan bezorgd worden, kan hier wel en beperking van de verantwoordelijkheid optreden.

4.3 HOE KAN DE GEBRUIKER DE TOEGELATEN I-SYSTEMEN OP EEN VEILIGE MANIER GEBRUIKEN?

Niet enkel het doel van het inzetten van AI is van belang, maar ook het verstrekken van richtlijnen over de vraag hoe met dergelijke systemen moet worden omgegaan.

Het eerste artikel dat onder deze noemer wordt aangehaald, heeft betrekking op de zogenaamde AI-geletterdheid. AI-toepassingen zijn maar zinvol in de mate dat men hiermee kan omgaan. Dit kan betrekking hebben op het attest maken op de risico's, maar even goed op technische elementen zoals het opmaken van correcte prompts. Er kan eventueel voor geopteerd worden de toegang tot AI-systemen voor te behouden aan gebruikers die bereid zijn daarvoor de nodige (aanvangs- of periodieke) opleiding te volgen.

Een eerste aandachtspunt wat, vervolgens, het effectieve gebruik betreft, heeft betrekking op de vertrouwelijkheid van gegevens. Er moet over gewaakt worden dat vertrouwelijke gegevens ook enkel binnen een bewaakte context blijven circuleren. Daarnaast blijft het nog altijd best practice om het gebruik van bedrijfsgevoelige, vertrouwelijke gegevens of persoonsgegevens zoveel als mogelijk te vermijden dan wel minstens dergelijke soort gegevens te anonimiseren of te pseudonimiseren. Anonimiseren betekent dat alle persoonlijke gegevens of informatie uit een tekst worden gehaald. Pseudonimiseren betekent dat een persoonsgegeven vervangen wordt door een fictieve naam, letter, nummer of anderszins.

Bijzondere aandacht dient er te zijn voor lichamelijke of mentale gezondheidsgegevens, genetische gegevens (persoonsgegevens die verband houden met overgeërfde of verworven genetische kenmerken van een natuurlijke persoon), ras of etnische afkomst, politieke opvattingen, religieuze of levensbeschouwelijke overtuigingen, seksueel gedrag of geaardheid van een natuurlijke persoon. Deze soort persoonsgegevens kwalificeren als bijzonder categorieën van persoonsgegevens en de verwerking ervan is in principe verboden volgens artikel 9.1 GPDR.

Een volgend aandachtspunt betreft de verificatie van de door AI gerealiseerde resultaten. Reeds onder punt 4.2 werd aangegeven dat het gebruik van AI-systemen de verantwoordelijkheid van de Gebruiker niet beperkt of wegneemt. Onder punt 4.3.2 wordt uitgewerkt dat dit impliceert dat de gebruiker de plicht heeft om de gegenereerde gegevens na te kijken. Dat kan in de eerste plaats gebeuren door – indien toepasselijk – de bronvermelding te checken. Daarnaast dient de gebruiker zich een aantal vragen te stellen zoals:

- Heeft de gegenereerde output een relevante en concrete inhoud?
- Bevat de gegenereerde output objectieve, correcte en onderbouwde informatie of is deze inconsistent, onjuist of subjectief?
- Bevat de gegenereerde output één of meerdere bronvermeldingen?
- Bevat de gegenereerde output informatie die beschermd is door enig intellectueel eigendomsrecht (bv. auteursrechten bij afbeeldingen, merkrechten, tekening- en modelrechten)?

Het is ook belangrijk dat binnen de organisatie open gecommuniceerd wordt over het gebruik van AI. Niet enkel is het uitgesloten dat een gebruiker de resultaten van een AI-interventie voorstelt als het resultaat van eigen, menselijke arbeid. Daarnaast moet ook de mogelijkheid bestaan om, bijvoorbeeld na het stellen van de hiervoor weergegeven vragen, in geval van twijfel de gegenereerde AI-resultaten te bespreken met een leidinggevende.

Anders dan ten aanzien van de interne communicatie die steeds open en transparant dient te gebeuren, is het niet opportuun dat elke gebruiker zelf ten aanzien van derden communiceert over het gebruik van AI. Aangezien dergelijke communicaties een impact kunnen hebben op de wijze waarop het cliënteel kijkt naar de dienstverlening van de onderneming, is het aan deze laatste om de regie van de communicatie daarover met derden, in eigen hand te houden. In bepaalde gevallen is het wenselijk of zelfs verplicht om ten aanzien van derden aan te geven dat gebruik gemaakt wordt van AI. Een verplichte situatie is bijvoorbeeld deze waarbij gebruik gemaakt wordt van een chatbot die communicatie met derden verzorgt. Deze derden dienen te weten dat zij niet communiceren met een fysieke persoon. Een situatie waarin de communicatie wenselijk is, is het eerder aangehaalde voorbeeld van de machinevertaling. Immers, door hierover transparant te communiceren is duidelijk dat het engagement van de onderneming – en dus haar aansprakelijkheid – beperkt is tot het maken van de machinevertaling.

Tot slot wijzen we op artikel 4.3.7 waarmee kan gesensibiliseerd worden over het feit dat AI een energie- en (dus) milieu-impact heeft. Dergelijke clausule is optioneel maar kan kaderen in een ESG-beleid van de onderneming.

4.4 WELK GEBRUIK IS VERBODEN?

Het toepasselijk wettelijk kader voorziet een aantal gevallen waarin het gebruik van AI uitgesloten is.

In de eerste plaats wordt hierbij verwezen naar de zogenaamde “subliminale technieken”, wat technieken zijn waarvan personen zich niet bewust zijn of die doelbewust manipulatief of misleidend zijn, waarvan het doel of het effect is het gedrag van personen of een groep personen te verstoren door hun vermogen om een geïnformeerd besluit te nemen merkbaar te belemmeren. Dergelijke technieken kunnen onder geen beding aangewend worden om personen te beïnvloeden.

In de tweede plaats wordt gewezen op het gegeven dat het verboden is om door middel van AI op geautomatiseerde wijze beslissingen te genereren over fysieke personen. Dergelijke beslissingen kunnen weliswaar een voorbereiding kennen waarin AI een rol kan spelen, maar eigenlijke beslissingen dienen door concrete personen te worden genomen.

Naast deze specifieke verboden, die veeleer zijn opgenomen ten titel van sensibilisering, is een algemene clausule die wijst op de wettelijke, sectorale of deontologische regels die ook bij het gebruik van AI dienen te worden gerespecteerd.

Bovendien dient het gebruik ook te gebeuren in overeenstemming met de instructies van de onderneming, wat eveneens expliciet is opgenomen.

4.5 REGELS VOOR HET WAARDIG EN GEDISCIPLINEERD GEBRUIK VAN IT-INFRASTRUCTUUR

Aangezien AI niet op zichzelf bestaat, maar bediend wordt vanuit de IT-infrastructuur, is het aangeraden om ook een korte passage op te nemen die vooral sensibiliserend werkt op het vlak van het gebruik van de infrastructuur. Mogelijk zit dergelijke passage reeds in een algemene IT-policy in welk geval deze hier eventueel geschrapt kan worden.

4.6 REGELS IN GEVAL VAN VASTGESTELDE OF VERMOEDELIJKE MISBRUIKEN

Onder de titel wordt een tweevoudige passage opgenomen. Vooreerst is de gebruiker, als effectief gebruiker van het systeem, goed geplaatst om vast te stellen of mogelijk een gebruikt systeem gehackt is of anderszins misbruikt wordt. Hierbij wordt geen doorgedreven technische kennis verwacht, maar enkel de alertheid om anomalieën te melden.

Anderzijds wordt er in dit artikel op gewezen dat ook de gebruiker zelf geen misbruik mag plegen. Mocht dat toch het geval zijn dan wordt enerzijds gewezen op de mogelijkheid van sancties, die afhankelijk kunnen zijn van de aard van de relatie tussen de gebruiker en de onderneming, waarnaast ook gewezen wordt op de mogelijkheid dat dit kan leiden tot het afnemen of opschorten van de toegang tot AI-systemen. Dit laatste wordt niet als een sanctie, maar als een veiligheidsmaatregel beschouwd, reden waarom dit vermeld wordt “onverminderd eventuele sancties”.