Richtlijn gebruik van Artificiële Intelligentie Mondriaan Fonds

Doel
Artificiële Intelligentie (AI) biedt kansen voor het efficiënter uitvoeren van werkzaamheden en het verbeteren van dienstverlening, besluitvorming en informatieverwerking. AI kan ondersteunen bij onderzoek, analyse, schrijven, visualiseren, vertalen, structureren van informatie, procesondersteuning en andere werkzaamheden. Tegelijkertijd brengt het gebruik van AI risico's met zich mee op het gebied van betrouwbaarheid, privacy, vertrouwelijkheid, auteursrecht, intellectueel eigendom, transparantie, bias, authenticiteit en zorgvuldige besluitvorming.

Het doel van deze richtlijn is om ervoor te zorgen dat AI op een verantwoorde, veilige, transparante en zorgvuldige manier wordt gebruikt. De richtlijn beschrijft welke uitgangspunten gelden, welke risico's moeten worden meegewogen en welke richtlijnen van toepassing zijn voor verschillende doelgroepen.

Deze richtlijn is van toepassing op:
- aanvragers
- medewerkers van het Mondriaan Fonds;
- directeur-bestuurder;
- zzp'ers en leden van de Raad van Toezicht;
- adviseurs en voorzitters van adviescommissies;

Voor medewerkers, zzp'ers, leden van de Raad van Toezicht, hierna interne gebruikers genoemd, adviseurs en voorzitters, bevat deze richtlijn interne regels en uitgangspunten. Voor aanvragers biedt de richtlijn uitsluitend richtinggevende kaders voor verantwoord AI-gebruik. Deze kaders vormen geen afzonderlijke beoordelingscriteria of voorwaarden voor subsidieverlening.

Vormen van AI:
AI is een verzamelnaam voor technologieën die taken kunnen uitvoeren waarbij normaal gesproken menselijke analyse, interpretatie, creativiteit of besluitvorming een rol speelt. Het Mondriaan Fonds onderscheidt de volgende vormen van AI:

- **Generatieve AI**: genereert nieuwe content, zoals tekst, beeld, audio, video of code.
- **Classificerende en herkennende AI**: deelt gegevens of documenten in categorieën in, herkent kenmerken of signaleert patronen en afwijkingen.
- **Voorspellende AI**: doet voorspellingen op basis van historische gegevens en patronen.
- **Aanbevelende AI**: doet suggesties of rangschikt informatie, opties of resultaten.
- **Procesondersteunende AI**: ondersteunt of automatiseert terugkerende administratieve of organisatorische werkzaamheden.
- **Beslissingsondersteunende AI**: levert analyses, inzichten of aanbevelingen die invloed kunnen hebben op adviezen, beoordelingen of besluiten.

AI-toepassingen combineren in de praktijk vaak meerdere van deze vormen. Zo kan één toepassing tegelijkertijd generatieve, classificerende en procesondersteunende functies bevatten.

Belangrijkste risico's van AI
Hoewel AI veel mogelijkheden biedt, kent het gebruik ervan ook risico's. Het Mondriaan Fonds onderkent de volgende risico's omtrent de inzet van AI:

1. Onjuiste of misleidende informatie
   AI-systemen kunnen onjuiste, onvolledige of misleidende resultaten opleveren. Bij generatieve AI kan dit zich uiten in zogenoemde hallucinaties, waarbij feiten, bronnen of verwijzingen onjuist of verzonnen zijn.

2. Vooroordelen en discriminatie
   AI-modellen worden getraind op bestaande gegevens en bronnen. Hierdoor kunnen vooroordelen en stereotyperingen in de output voorkomen.

3. Privacy- en beveiligingsrisico's
   Afhankelijk van de gebruikte AI-toepassing kan ingevoerde informatie worden opgeslagen, verwerkt of gebruikt voor verdere ontwikkeling van het systeem. Het invoeren van vertrouwelijke informatie, persoonsgegevens, contracten, zakelijke gegevens of informatie van derden, kan daarmee leiden tot ongewenste verspreiding van gegevens.

4. Auteursrecht en intellectueel eigendom
   Het is niet altijd duidelijk op welke bronnen AI gegenereerde output is gebaseerd. Hierdoor kunnen risico's ontstaan op het gebied van auteursrecht, intellectueel eigendom, onjuiste bronvermelding en plagiaat.

5. Afhankelijkheid van AI
   Overmatig vertrouwen op AI kan leiden tot verlies van kritisch denkvermogen, expertise en professionele oordeelsvorming.

6. Transparantie en controleerbaarheid
   Het is niet altijd inzichtelijk hoe AI-systemen tot een uitkomst of aanbeveling komen. Hierdoor kan het moeilijk zijn om resultaten te verklaren, te controleren of te verantwoorden.

Algemene uitgangspunten
Bij de inzet van AI binnen het Mondriaan Fonds zijn zorgvuldigheid, transparantie, menselijke controle en naleving van wet- en regelgeving altijd het uitgangspunt. Het Mondriaan Fonds zet AI momenteel niet in voor subsidiebeoordelingen, advisering of besluitvorming.

Het Mondriaan Fonds verwacht dat AI op een verantwoorde, transparante en zorgvuldige wijze wordt ingezet. De onderstaande uitgangspunten vormen het algemene kader voor verantwoord AI-gebruik. Voor interne gebruikers worden deze uitgewerkt in concrete regels. Voor aanvragers hebben zij een richtinggevend karakter.

1. Menselijke verantwoordelijkheid
   AI kan werkzaamheden ondersteunen, maar neemt geen verantwoordelijkheden over. De gebruiker blijft altijd eindverantwoordelijk voor de inhoud, juistheid van informatie, kwaliteit, bronvermelding en naleving van wet- en regelgeving.

2. Kritische controle
   AI-systemen kunnen onjuiste, verouderde, onvolledige of misleidende resultaten opleveren. AI gegenereerde output wordt daarom door de gebruiker gecontroleerd op juistheid, volledigheid, actualiteit, integriteit en bruikbaarheid.

3. Transparantie
   Wees transparant over wanneer je AI-gegenereerde content (beeld, tekst of audio) hebt gebruikt of AI toepast als bron of methode in een project of onderzoek.

4. Privacy en vertrouwelijkheid
   Vertrouwelijke informatie, persoonsgegevens en informatie uit subsidieaanvragen worden uitsluitend verwerkt binnen daarvoor goedgekeurde systemen, binnen de grenzen van deze richtlijn en volgens het privacy- en informatiebeveiligingsbeleid.

Richtlijnen voor aanvragers
Het Mondriaan Fonds beoordeelt de kwaliteit van het plan, de artistieke visie, de onderbouwing en de professionaliteit van de aanvrager. Het is daarom belangrijk dat de aanvraag de eigen inhoudelijke inbreng, visie, artistieke praktijk en reflectie van de aanvrager weerspiegelt.

AI kan als hulpmiddel worden ingezet voor het ontwikkelen, uitwerken en presenteren van plannen, ideeën en artistieke concepten. Daarnaast kan AI deel uitmaken van het artistieke maakproces, bijvoorbeeld als medium, techniek, onderwerp of onderzoeksinstrument. Bij het gebruik van AI zijn bewust gebruik, transparantie, zorgvuldigheid en eigen verantwoordelijkheid belangrijke uitgangspunten.

Voor een verantwoord gebruik van AI worden aanvragers de volgende richtlijnen meegegeven.
1. Controleer door AI gegenereerde teksten, beelden en andere output zorgvuldig voordat deze onderdeel worden van de aanvraag. Wees daarnaast transparant over het gebruik.
2. Ga zorgvuldig en terughoudend om met vertrouwelijke informatie, persoonsgegevens en informatie van derden. Voer deze alleen in wanneer dit is toegestaan en de AI-toepassing daarvoor passend en veilig is.
3. Gebruik AI bewust en kritisch, met aandacht voor privacy, auteursrecht, intellectueel eigendom en andere toepasselijke wet- en regelgeving.
4. Kies bij voorkeur voor betrouwbare AI-tools die transparant zijn over privacy, gegevensopslag en hergebruik van informatie en rekening houden met toepasselijke Europese regelgeving.

Richtlijnen voor interne gebruikers
Onder interne gebruikers verstaan we medewerkers, directeur-bestuurder, zzp'ers en leden van de Raad van Toezicht. AI kan interne gebruikers ondersteunen bij schrijven, samenvatten, analyseren, organiseren, communiceren en andere werkzaamheden. Het gebruik van AI
is erop gericht interne gebruikers te ondersteunen bij hun werk en niet om professionele verantwoordelijkheid of deskundigheid te vervangen. Interne gebruikers blijven verantwoordelijk voor hun professionele oordeel en de kwaliteit van hun werkzaamheden.

Interne gebruikers mogen AI inzetten ter ondersteuning van hun werkzaamheden, mits dit gebeurt binnen de uitgangspunten van menselijke verantwoordelijkheid, kritische controle, transparantie, privacy en regelgeving.

Bij het gebruik van AI door interne gebruikers gelden de volgende regels:

1. Alleen door het Mondriaan Fonds goedgekeurde en beschikbaar gestelde AI-toepassingen mogen voor het werk worden gebruikt.
2. Informatie uit subsidieaanvragen, preadviezen en commissiestukken wordt uitsluitend verwerkt binnen daarvoor goedgekeurde toepassingen, overeenkomstig deze richtlijn en de geldende werkprocessen.
3. Persoonsgegevens mogen alleen bewust en beschermd worden gebruikt in goedgekeurde, veilige AI-toepassingen. Bijzondere persoonsgegevens worden niet ingevoerd in AI-toepassingen.
4. AI mag niet worden gebruikt voor inhoudelijke beoordeling van subsidieaanvragen of voor het vervangen van menselijke oordeelsvorming in besluitvormingsprocessen.
5. Interne gebruikers blijven verantwoordelijk voor de inhoud van hun werk en controleren alle AI gegenereerde output voordat deze wordt gebruikt of gedeeld.
6. Interne gebruikers handelen bewust ten aanzien van mogelijke bias, discriminatie, plagiaat en auteursrechtelijke risico's.
7. Wanneer AI een substantiële bijdrage heeft geleverd aan een publicatie, presentatie of andere uiting, wordt dit waar relevant vermeld overeenkomstig de uitgangspunten over transparantie.
8. Interne gebruikers passen AI bewust toe, proportioneel en met oog voor duurzaamheid. AI wordt alleen ingezet wanneer dit daadwerkelijk bijdraagt aan de kwaliteit, efficiëntie, dienstverlening, toegankelijkheid of uitvoering van werkzaamheden.
9. Interne gebruikers moeten kunnen uitleggen hoe AI is ingezet. Wanneer AI een substantiële bijdrage heeft geleverd aan een analyse, advies of document, moet de interne gebruiker kunnen toelichten welke AI-toepassing is gebruikt en welke rol deze heeft gespeeld.
10. Interne gebruikers melden incidenten. Wanneer vertrouwelijke informatie onbedoeld wordt gedeeld met een AI-systeem of wanneer sprake is van een mogelijk datalek, wordt dit direct gemeld volgens de geldende procedures voor informatiebeveiliging en privacy.
11. Interne gebruikers houden hun kennis over verantwoord AI-gebruik actueel door beschikbare trainingen en voorlichting te volgen. Interne gebruikers die AI gebruiken, beschikken over voldoende kennis van de mogelijkheden, beperkingen en risico's van de betreffende toepassing. Dit kennisniveau sluit aan bij de eisen voor AI-geletterdheid uit de Europese AI-verordening.

Richtlijnen voor adviseurs en voorzitters
Subsidieaanvragen worden beoordeeld door onafhankelijke adviescommissies. Van adviseurs en voorzitters wordt verwacht dat zij hun oordeel zelfstandig, zorgvuldig en zonder tussenkomst van AI vormen. Daarom gelden voor deze groep strengere regels dan voor medewerkers en aanvragers.

Bij het gebruik van AI door adviseurs en voorzitters gelden de volgende regels:

Voor adviseurs en voorzitters geldt dat AI niet mag worden gebruikt bij:

- het analyseren, samenvatten of beoordelen van aanvragen;
- het opstellen van preadviezen of beoordelingsrapporten;
- de voorbereiding van inhoudelijke oordelen;
- beraadslagingen binnen adviescommissies.

Deze uitsluiting is gebaseerd op verschillende uitgangspunten. Aanvragen, bijlagen, preadviezen en beoordelingsrapporten bevatten vertrouwelijke informatie. Het delen van deze informatie met AI-systemen is niet toegestaan en kan strijdig zijn met de geheimhoudingsplicht. Daarnaast moet het Mondriaan Fonds kunnen nagaan hoe een advies tot stand is gekomen en op welke gronden een oordeel is gevormd. Het gebruik van AI kan de controleerbaarheid en navolgbaarheid van adviezen verminderen en daarmee afbreuk doen aan de zorgvuldigheid van het beoordelingsproces.

Daarnaast hecht het Mondriaan Fonds belang aan een onafhankelijke, zorgvuldige en transparante beoordeling van aanvragen. AI-toepassingen kunnen bias bevatten, onjuiste uitkomsten geven of onvoldoende transparant zijn over de wijze waarop resultaten tot stand komen. Door AI uit te sluiten van het beoordelingsproces wordt bevorderd dat adviezen tot stand komen op basis van menselijke deskundigheid, professionele oordeelsvorming en onafhankelijke afwegingen.

Beheer en verantwoordelijkheden
De afdeling IT ondersteunt medewerkers binnen het Mondriaan Fonds bij het gebruik van goedgekeurde AI-toepassingen. IT werkt daarbij samen met de betrokken proceseigenaren en andere relevante collega's. Dit houdt in:

- het beschikbaar stellen en beheren van goedgekeurde AI-toepassingen;
- het volgen van ontwikkelingen in wet- en regelgeving;
- het bevorderen van AI-geletterdheid binnen de organisatie;
- het adviseren over verantwoord AI-gebruik binnen het Mondriaan Fonds;
- het actualiseren van de richtlijn op basis van wijzigingen in wet- en regelgeving, technologische ontwikkelingen, praktijkervaringen en risico's.

Het Mondriaan Fonds evalueert deze richtlijn periodiek en past deze indien nodig aan.