

# HOE GA JE AAN DE SLAG?

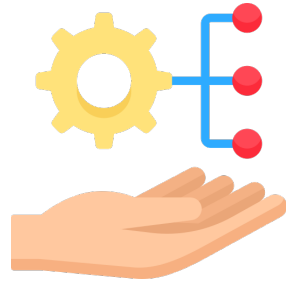
*Responsible AI guide*

# Hoofdpijnen - Voorbereiding op de AI Act



1

Breng het gebruik van algoritmen in kaart



2

Categoriseer uw algoritmen: risico hoog, midden, laag



3

Voer de nodige assessments uit, zoals het IAMA



4

Stel een beleid op tbv omgang met algoritmen



# Hoe stel ik dit beleid op en implementeer ik het in mijn organisatie?

- Stappenplan voor verantwoorde inzet AI
- Voldoen aan de eisen van de AI Act, mensenrechten, AVG en (sector)specifieke wetgeving

 **CONSIDERATI**

**curiosity**

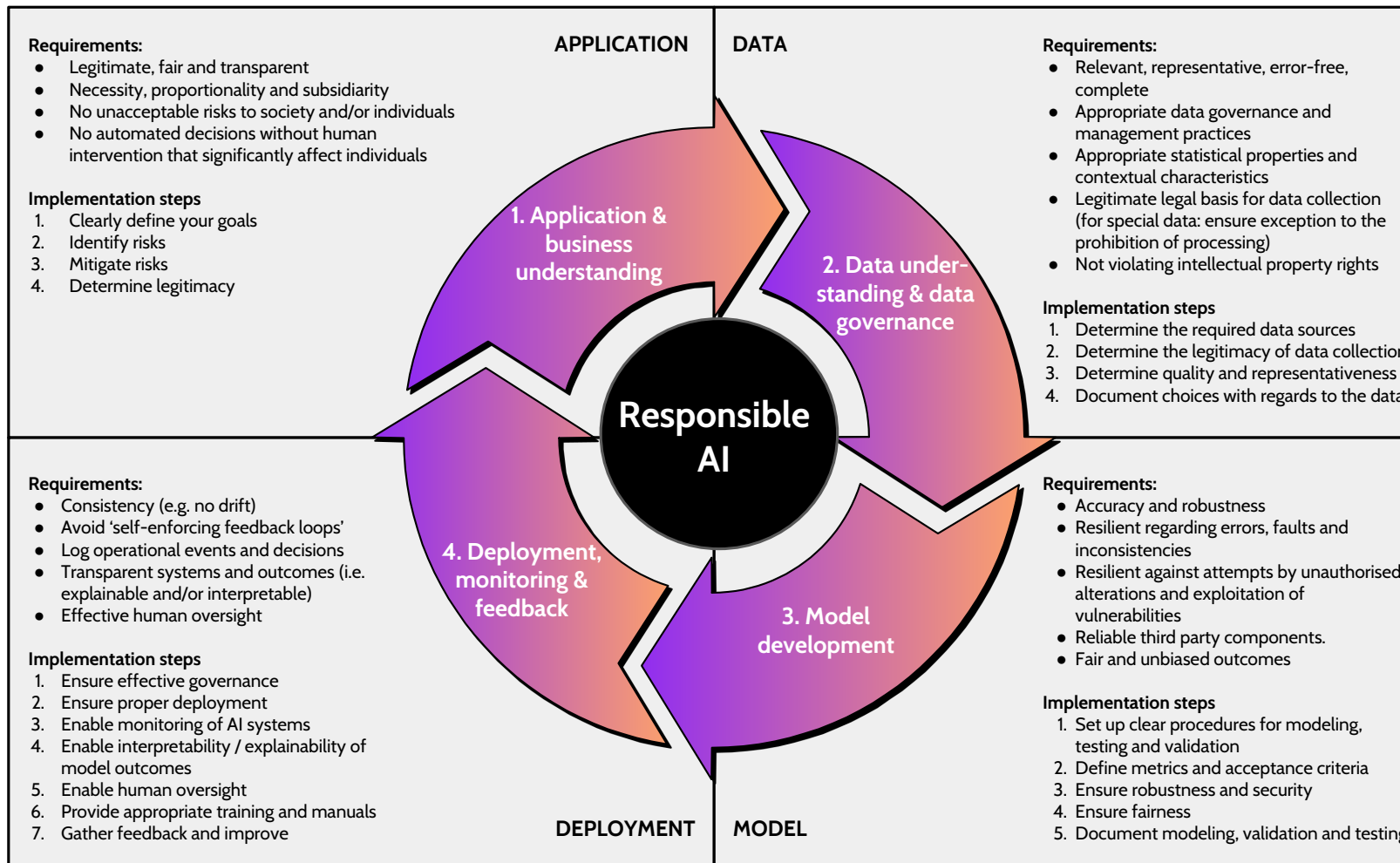


 **CONSIDERATI**



# Responsible AI Guide - 4 fasen





CC BY-NC-ND Licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International



Download the full Responsible AI Guide via this QR code:



# 1. Toepassing en bedrijfskennis

## Vereisten

- Legitiem, rechtvaardig en transparant
- Noodzakelijkheid, proportionaliteit, subsidiariteit
- Het AI-systeem mag geen onaanvaardbare risico's opleveren voor de samenleving en/of individuen.
- Geen geautomatiseerde beslissing met impact zonder menselijke tussenkomst

## Aanpak

1 Definieer je doelen duidelijk

2 Identificeer risico's

3 Beperk risico's

4 Bepaal legitimiteit

## *Article 29a*

### *Fundamental rights impact assessment for high-risk AI systems*

1. *Prior to putting a high-risk AI system as defined in Article 6(2) into use, with the exception of AI systems intended to be used in area 2 of Annex III, deployers shall conduct an assessment of the systems' impact in the specific context of use. This assessment shall include, at a minimum, the following elements:*
  - (a) a clear outline of the intended purpose for which the system will be used;*
  - (b) a clear outline of the intended geographic and temporal scope of the system's use;*
  - (c) categories of natural persons and groups likely to be affected by the use of the system;*
  - (d) verification that the use of the system is compliant with relevant Union and national law on fundamental rights;*
  - (e) the reasonably foreseeable impact on fundamental rights of putting the high-risk AI system into use;*
  - (f) specific risks of harm likely to impact marginalised persons or vulnerable groups;*
  - (g) the reasonably foreseeable adverse impact of the use of the system on the environment;*
  - (h) a detailed plan as to how the harms and the negative impact on fundamental rights identified will be mitigated.*
  - (j) the governance system the deployer will put in place, including human oversight, complaint-handling and redress.*

## 2. Inzicht in data en databeheer

### Vereisten

- Gegevens die worden gebruikt voor training, validatie en testen moeten relevant, representatief, foutloos en volledig zijn.
- Op trainings-, validerings- en testgegevensverzamelingen zijn passende praktijken voor databeheer en -governance van toepassing.
- Data wordt verzameld met een legitieme rechtsgrondslag.
- Intellectuele eigendomsrechten.

### Aanpak

- 1 Bepaal de vereiste databronnen
- 2 Stel de legitimiteit van de dataverzameling vast
- 3 Bepaal de kwaliteit en representativiteit
- 4 Documenteer keuzes met betrekking tot de data



# 3. Modelleren

## Vereisten

- AI-systemen moeten een passend niveau van nauwkeurigheid en robuustheid bereiken.
- AI-systemen moeten bestand zijn tegen fouten, gebreken en inconsistenties.
- AI-systemen moeten veerkrachtig zijn
- De gebruikte componenten van derden zijn betrouwbaar (bijv. bibliotheken, AI-modellen).
- Modeluitkomsten moeten eerlijk zijn en vooringenomenheid vermijden.

## Aanpak

- 1 • Stel duidelijke procedures op voor modellering, testen en valideren
- 2 • Definieer metriek en acceptatiecriteria
- 3 • Robuustheid en veiligheid waarborgen
- 4 • Rechtvaardigheid waarborgen
- 5 • Document modeling, validation and testing

## 4. Implementatie en monitoring

### Vereisten

- AI-systemen moeten gedurende hun hele levenscyclus consistent presteren (bijv. voorkomen van drift, feedbackloops).
- AI-systemen moeten gebeurtenissen documenteren die relevant zijn voor hun werking en besluitvorming (bijhouden van gegevens).
- AI-systemen en -resultaten moeten verklaarbaar en/of interpreteerbaar zijn.
- AI-systemen moeten geschikt zijn voor effectief menselijk toezicht.

### Aanpak

- 1 Effectief toezicht waarborgen
- 2 Een juiste inzet waarborgen
- 3 Monitoring van AI-systemen mogelijk maken
- 4 Interpreteerbaarheid / verklaarbaarheid van uitkomsten mogelijk maken
- 5 Menselijk toezicht mogelijk maken
- 6 Voorzie gebruikers van de juiste training en handleidingen
- 7 Feedback verzamelen en verbeteren

# Vragen?

[goudsmit@considerati.com](mailto:goudsmit@considerati.com)

