

***Kader voor waardevolle, verantwoorde en ethische inzet  
van generatieve AI***

***binnen het toepassingsgebied wetenschappelijk onderzoek***

***door rijkskennisinstituten en planbureaus***

***“RKI-kader”***

**Versie 1.0**

Maandag 26 mei 2025

*Vastgesteld door de RKI-directeuren*

## Managementsamenvatting

Staatssecretaris Szabó (Digitalisering) presenteerde op 22 april 2025 het nieuwe overheidsbrede standpunt over generatieve AI. Daarmee is het voorlopige standpunt van Staatssecretaris Van Huffelen (Digitalisering) uit december 2023 herijkt. Het nieuwe standpunt is het resultaat van een interbestuurlijk traject, waarbij zijn ook (vertegenwoordigers van) rijkskennisinstituten en planbureaus (RKI's) betrokken geweest.

Het nieuwe overheidsbrede standpunt biedt ruimte voor innovatie en stelt kansen van generatieve AI centraal. Concreet en samengevat, stelt het overheidsbrede standpunt de volgende voorwaarden aan ontwikkeling, inkoop en het gebruik van generatieve AI:

- De inzet van generatieve AI is in lijn met geldende wet- en regelgeving;
- Randvoorwaarden – zoals AI-geletterdheid, governance en informatiehuishouding – zijn op orde;
- Voorafgaand aan de inzet van generatieve AI wordt een gedegen risicoanalyse uitgevoerd;
- Ambtenaren gebruiken in beginsel geen generatieve AI 'onder consumentenvoorwaarden', en;
- Bij de inzet van generatieve AI hebben open source oplossingen en samenwerkingen met partijen uit de EU de voorkeur.

Bij het standpunt hoort een aanvullende, ondersteunende handreiking. Daarin staan voorbeelden van toepassingsgebieden waar de inzet van generatieve AI een verlaagd risico inhoudt, zoals ook wetenschappelijk onderzoek.

Over de inzet van generatieve AI in wetenschappelijk onderzoek zegt de overheidsbrede handreiking het volgende: *“Wetenschappers kunnen bijvoorbeeld generatieve AI inzetten om onderzoek efficiënter en creatiever te maken, om groeiende hoeveelheden data te verwerken, maar hebben ook de mogelijkheid om zelf generatieve AI te bestuderen of te ontwikkelen. Omdat wetenschappelijk onderzoek specifieke doelen nastreeft en zich in een specifieke context voltrekt, zijn generieke kaders niet toereikend. Het verantwoord ontwikkelen en toepassen van generatieve AI zal in lijn moeten gebeuren met uitgangspunten van onderzoeksintegriteit en het overheidsbrede beleid. Om dat gericht te adresseren, ontwikkelen rijkskennisinstituten en planbureaus (verenigd in het RKI-netwerk) gezamenlijk een kader voor de inzet en (door-) ontwikkeling van generatieve AI binnen de contexten van het wetenschappelijk onderzoek.”*

Een van de centrale uitgangspunten van het RKI-kader is dat generatieve AI door overheidsorganisaties **kan en mag** worden gebruikt binnen wetenschappelijk onderzoek, met inachtneming van het overheidsbrede standpunt en de uitgangspunten van wetenschappelijke integriteit (*toegestaan, mits*).

Door de randvoorwaarden en het model voor risicoanalyse, waar passend, te specificeren voor RKI's en hun onderzoekers, binnen de context van wetenschappelijk onderzoek, is het RKI-kader een belangrijke stap om waardevolle, verantwoorde en ethische inzet van generatieve AI binnen wetenschappelijk onderzoek beter te borgen, alsmede de aantoonbaarheid daarvan te vergroten. Zo biedt het RKI-kader gewenste duidelijkheid ter voorbereiding op de inzet van generatieve AI binnen wetenschappelijk onderzoek.

Bij de risicoanalyse wordt vervolgens ingegaan op de concrete situatie waarin een overheidsorganisatie generatieve AI wil gebruiken, binnen het toepassingsgebied wetenschappelijk onderzoek. Een belangrijk deel van de relevante aspecten volgt uit het overheidsbrede standpunt. Vervolgens adresseert het model voor risicoanalyse, met inachtneming van de uitgangspunten van wetenschappelijke integriteit, acht thema's die relevant zijn bij het borgen van waardevolle, verantwoorde en ethische inzet van generatieve AI. Deze thema's zijn door de RKI's gezamenlijk vastgesteld. Deze acht thema's zijn:

1. Menselijke controle en menselijk toezicht;
2. Technische robuustheid;
3. Privacy en datagovernance;
4. Transparantie;
5. Reproduceerbaarheid;
6. Diversiteit, non-discriminatie en rechtvaardigheid;
7. Milieu- en maatschappelijke waarden, en;
8. Verantwoording.

Het RKI-kader is tot stand gekomen in samenwerking tussen het RKI-netwerk en het ministerie van Binnenlandse Zaken (CIO Rijk), van januari tot en met mei 2025. De RKI-directeuren stelden het RKI-kader vast op 26 mei 2025.

## Inhoudsopgave

<b>Managementsamenvatting</b> .....	<b>2</b>
<b>1. Inleiding</b> .....	<b>5</b>
1.1. Uitgangspunten.....	6
1.2. Status en beheer .....	10
<b>2. Beginselen voor het gebruik van generatieve AI binnen wetenschappelijk onderzoek</b> .....	<b>11</b>
<b>3. Risicoanalyse voor het gebruik van generatieve AI binnen wetenschappelijk onderzoek</b> .....	<b>16</b>
<b>4. Achtergrond bij het RKI-kader</b> .....	<b>21</b>
<b>Bijlage 1: Model voor risicoanalyse voor het gebruik van generatieve AI binnen wetenschappelijk onderzoek</b> .....	<b>24</b>
<b>Bijlage 2: Voorbeelden ter illustratie voor Onderdeel B van het model voor risicoanalyse</b> .....	<b>25</b>

## Versiebeheer:

Versie	Datum	Omschrijving
0.1	31/1/25	H.1 en opzet ter bespreking met RKI WG en CIO Rijk
0.2	20/2/25	Concept (H.1, 2, 4 en opzet H.3) ter bespreking met RKI WG en CIO Rijk
0.3	28/2/25	Concept (H.1, 2, 3 en 4) ter bespreking met RKI WG en CIO Rijk
0.4	7/4/25	Concept (H. 2 en 3) ter bespreking met RKI WG en CIO Rijk
0.4.1.	9/4/25	Concept ter consultatie (geschoonde versie)
0.5	6/5/25	Concept n.a.v. consultatie (voorlopig)
0.6	12/5/25	Concept n.a.v. consultatie en definitieve publicatie nieuw overheidsbreed standpunt en (aanvullende) handreiking generatieve AI
0.7	20/5/25	Vastgesteld door RKI WG
0.99	20/5/25	Versie t.b.v. RKI-directeurenoverleg
1.0	26/5/25	Vastgesteld door RKI-directeurenoverleg

## 1. Inleiding

De ontwikkeling van kunstmatige intelligentie (hierna: “AI”) heeft in de afgelopen decennia een grote vlucht genomen. Van eenvoudige algoritmen, via machine learning en generatieve AI, tot toekomstbeelden over agents, quantum en artificial general intelligence (AGI). Vanaf 2022 zorgen vrij toegankelijke *large language models* (LLM’s) voor toenemende bekendheid en aandacht voor generatieve AI. Het bekendste voorbeeld hiervan is ChatGPT. Generatieve AI lijkt potentie te hebben om zich te ontwikkelen tot een veelzijdig, krachtig hulpmiddel bij diverse werkzaamheden. Zoals onder andere het creëren van teksten, code, synthetische data(sets), afbeeldingen, muziek, video’s.<sup>1</sup>

Voor wetenschappelijk onderzoek biedt generatieve AI grote kansen, maar kent het ook risico’s. Het kan onderzoekers helpen, bijvoorbeeld bij het simuleren van complexe systemen, het genereren van nieuwe hypothesen, het automatiseren van routinetaken en literatuuronderzoek. Het benutten van kansen van generatieve AI vergroot efficiëntie en creativiteit van onderzoek. Er bestaan echter ook risico’s. Deze dienen te worden geanalyseerd en beheerst. Relevante risico’s gaan bijvoorbeeld over *black box effects*,<sup>2</sup> misleidende of onnauwkeurige uitvoer, ethische aspecten, veiligheid en juridische kwesties, over onderwerpen als privacy en intellectuele eigendomsrechten.

Voor overheidsorganisaties gelden er regels voor het gebruik van generatieve AI. Deze volgen uit relevante wetten en regels, het overheidsbrede standpunt over generatieve AI (hierna: “Standpunt”)<sup>3</sup> en de aanvullende handreiking over verantwoorde inzet van generatieve AI (hierna: “Handreiking”).<sup>4</sup> Een van de centrale uitgangspunten is dat generatieve AI binnen wetenschappelijk onderzoek kan en mag worden gebruikt door overheidsorganisaties, met inachtneming van het Standpunt en de uitgangspunten van wetenschappelijke integriteit (*toegestaan, mits*).

In Nederland zijn er veertien rijkskennisinstellingen en planbureaus (hierna gezamenlijk aangeduid als: “RKI’s”). RKI’s zijn allen onderdeel van de rijksoverheid. Samen vormen zij het RKI-netwerk.<sup>5</sup> RKI’s ondersteunen beleid met hun wetenschappelijke kennis en innovatie. Zij werken aan onderwerpen zoals veiligheid, gezondheid, economie, onderwijs, leefomgeving, mobiliteit en biodiversiteit.<sup>6</sup>

Ontwikkelingen op het gebied van generatieve AI bieden RKI’s en hun onderzoekers grote kansen. In lijn met het Standpunt, beogen RKI’s deze technologie op waardevolle, verantwoorde en ethische wijze te gebruiken, (door) te ontwikkelen en/of in te kopen. Zij hebben daarbij behoefte aan duidelijkheid.

Bij leden van het RKI-netwerk spelen soortgelijke vraagstukken en dilemma’s rondom waardevolle, verantwoorde en ethische inzet van generatieve AI. Vaak spelen deze binnen het toepassingsgebied wetenschappelijk onderzoek. Daarom heeft het RKI-netwerk gezamenlijk dit ‘Kader voor waardevolle, verantwoorde en ethische inzet van generatieve AI binnen het toepassingsgebied

---

<sup>1</sup> Zie ook Gartner, *Hype Cycle for Artificial Intelligence 2024*, via: <https://www.gartner.com/en/documents/5505695>.

<sup>2</sup> Hiermee wordt bedoeld dat uitkomsten niet kunnen worden verklaard of uitgelegd, waardoor transparantie, reproduceerbaarheid en verantwoording van het onderzoek en de resultaten in het geding komen.

<sup>3</sup> Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (CIO Rijk), Overheidsbreed standpunt voor de inzet van generatieve AI, 22 april 2025 (hierna aangeduid als ‘Standpunt’).

<sup>4</sup> Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (CIO Rijk), Overheidsbrede handreiking voor de verantwoorde inzet van generatieve AI, 22 april 2025 (hierna aangeduid als ‘Handreiking’).

<sup>5</sup> Leden van het RKI-netwerk zijn het Centraal Planbureau (CPB), het Centraal Bureau voor de Statistiek (CBS), het Kennisinstituut voor Mobiliteitsbeleid (KiM), het Koninklijk Nederlands Meteorologisch Instituut (KNMI), het Naturalis Biodiversity Center, het Nederlands Forensisch Instituut (NFI), het Nederlands Instituut Publieke Veiligheid (NIVP), het Planbureau voor de Leefomgeving (PBL), de Rijksdienst voor het Cultureel Erfgoed (RCE), het Rijksinstituut voor Volksgezondheid en Milieu (RIVM), de Rijks Kunsthistorische Dienst (RKD), Rijkswaterstaat (RWS), het Sociaal en Cultureel Planbureau (SCP) en het Wetenschappelijk Onderzoek en Data Centrum (WODC).

<sup>6</sup> Zie via: <https://www.rijksoverheid.nl/contact/contactgids/rijkskennisinstellingen-rki>.

wetenschappelijk onderzoek door rijkskennisinstituten en planbureaus' (hierna: **"RKI-kader"**) ontwikkeld.

Met het RKI-kader voorzien RKI's in hun behoefte aan een gezamenlijk, gedragen en functioneel kader om waardevolle, verantwoorde en ethische inzet van generatieve AI binnen de het toepassingsgebied wetenschappelijk onderzoek te borgen. Het RKI-kader bestaat in feite uit twee onderdelen:

1. Hoofdstuk 2 biedt uitgangspunten voor overheidsorganisaties en hun onderzoekers die generatieve AI willen inzetten;
2. Hoofdstuk 3 biedt een functioneel instrument om de kansen en risico's rondom het gebruiken, (door-) ontwikkelen en inkopen van generatieve AI te onderbouwen, te analyseren en te verantwoorden.

### **1.1. Uitgangspunten**

#### **Het RKI-kader sluit aan op het Standpunt en de Handreiking**

Om generatieve AI waardevol, verantwoord en ethisch te gebruiken, (door) te ontwikkelen en/of in te kopen, stelt het Standpunt voorwaarden.<sup>7</sup> Het Standpunt wordt nader aangevuld en ondersteund met de Handreiking. Concreet en samengevat houdt het Standpunt het volgende in:

- Er is ruimte voor innovatie en de kansen die generatieve AI biedt;
- De inzet van generatieve AI is in lijn met geldende wet- en regelgeving;<sup>8</sup>
- Randvoorwaarden – zoals AI-geletterdheid, governance en informatiehuishouding – zijn op orde;
- Voorafgaand aan de inzet van generatieve AI wordt een gedegen risicoanalyse uitgevoerd;
- Ambtenaren gebruiken in beginsel geen generatieve AI 'onder consumentenvoorwaarden',<sup>9</sup> en;
- Bij de inzet van generatieve AI hebben open source oplossingen en samenwerkingen met partijen uit de EU de voorkeur.

#### **Het RKI-kader sluit aan op uitgangspunten van wetenschappelijke integriteit**

Uitgangspunten van wetenschappelijke integriteit spelen een belangrijke rol bij het inkopen, gebruiken en ontwikkelen van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek.<sup>10</sup> Het RKI-kader adresseert de leidende principes van wetenschappelijke integriteit – eerlijkheid, zorgvuldigheid, transparantie, onafhankelijkheid en verantwoordelijkheid – op efficiënte, passende en herkenbare wijze.

#### **Het RKI-kader is een ondersteunend en niet bindend instrument**

Overheidsorganisaties die generatieve AI willen inzetten binnen het toepassingsgebied wetenschappelijk onderzoek zijn niet verplicht het RKI-kader te gebruiken.<sup>11</sup> Zij kunnen ervoor kiezen om een ander instrument te gebruiken voor risicoanalyse, teneinde waardevolle, verantwoorde en ethische inzet van generatieve AI te borgen.

---

<sup>7</sup> Standpunt, p. 2 en 3. Gebruik van generatieve AI zal zijn 'toegestaan, mits'.

<sup>8</sup> Voor RKI's kan dit ook eventueel toepasselijke sectorale wet- en regelgeving omvatten.

<sup>9</sup> Dit is niet toegestaan omdat er met leveranciers geen afspraken kunnen worden gemaakt. Echter, verkenningen of privégebruik van nieuwe (generatieve) AI-systemen zijn toegestaan in een gecontroleerde niet-productieomgeving, mits er geen gevoelige of vertrouwelijke data worden gebruikt. Als kan worden onderbouwd dat van een dergelijke situatie sprake is, hoeft het RKI-kader niet te worden doorlopen om het uitgangspunt van waardevolle, verantwoorde en ethische inzet van generatieve AI te kunnen aantonen., zie ook Standpunt, p. 3 en Handreiking, par. 4.5, p. 13.

<sup>10</sup> Voor relevante bronnen zie Hoofdstuk 4, onder Wetenschappelijke integriteit.

<sup>11</sup> RKI's kunnen er te allen tijde zelf voor kiezen om een uitgebreidere risicoanalyse uit te voeren bij (voorgenomen) gebruik van generatieve AI binnen de context van onderzoek.

## **Het RKI-kader richt zich op het toepassingsgebied wetenschappelijk onderzoek**

Voorafgaand aan de totstandkoming van het RKI-kader identificeerden RKI's vijf situaties (doeleinden) waarin zij behoefte hebben aan duidelijkheid over de inzet en (door-)ontwikkeling van generatieve AI. De vijf vooraf geformuleerde situaties (doeleinden) zijn:

1. Generatieve AI als zelfstandig studieobject;
2. Wetenschappelijk onderzoek met generatieve AI voor grotere effectiviteit en efficiëntie;
3. Ondersteuning bij het gebruik van de (Nederlandse) taal;<sup>12</sup>
4. Verhoging efficiëntie van statistici, software engineers en data scientists;
5. Genereren van (statistisch relevante) synthetische datasets.

Gaandeweg constateerden RKI's dat dit twee hoofddoeleinden betreffen en drie voorbeelden die het tweede doeleinde nader invullen. In de toekomst kan blijken dat deze opsomming niet uitputtend is.

Voorbeelden van onderwerpen waar onderzoek naar generatieve AI zich op richt:

### **Begrijpelijkheid en interpretatie van generatieve AI**

Uitleg en begrip van generatieve AI-modellen zijn belangrijk om het vertrouwen in en de acceptatie van generatieve AI in de samenleving te vergroten. Het helpt ervoor te zorgen dat deze technologie meer verantwoord wordt ingezet.

### **Toepassing in specifieke (onderzoeks)domeinen**

Er worden generatieve AI-modellen ontwikkeld in bijvoorbeeld de gezondheidszorg, de financiële wereld en *cybersecurity*. Bij elke specifieke toepassing is zorg vereist voor domeinspecifieke uitdagingen. Bovendien kan onderzoek naar generatieve AI de effectiviteit en efficiëntie vergroten.

### **Robuustheid en veiligheid**

Net als bij elk ander IT-systeem, zijn robuustheid tegen fouten en aanvallen van derden van belang. Onderzoekers moeten in staat zijn om aanvallen proberen uit te voeren op generatieve AI-modellen, zodat ze daarna kunnen worden aangepast tegen aanvallen van derden, fouten bij generalisatie en andere fouten die kunnen leiden tot ongewenste en schadelijke uitkomsten.

### **Ethiek en verantwoord gebruik**

Onderzoekers bestuderen ethische implicaties en risico's van generatieve AI, zoals *deepfakes*, desinformatie, *bias* en privacy. Zo neemt de kennis toe om misbruik te voorkomen en te zorgen voor eerlijkere en meer verantwoorde generatieve AI.

### **Trainingstechnieken**

Het verbeteren van methoden om generatieve AI te trainen, zoals het genereren van *biasvrije* teksten, zelfsupervisie, *fine tuning* en *transfer learning*. Dit kan tevens helpen om uitdagingen rond generatieve AI, zoals duurzaamheidsaspecten, beter te adresseren.

Toegang tot (nieuwe) generatieve AI-modellen is essentieel om het als studieobject te kunnen gebruiken. Daarbij gaan de ontwikkelingen zo snel, dat het voor RKI's noodzakelijk is hun kennis over generatieve AI actueel te houden. Het Standpunt biedt hiervoor de benodigde ruimte.

Behalve het toepassingsgebied wetenschappelijk onderzoek, geeft de Handreiking ook voorbeelden van andere toepassingsgebieden met een verlaagd risico. Mogelijk worden daarvoor ook specifieke, lichtere instrumenten voor risicoanalyse ontwikkeld, vergelijkbaar met het RKI-kader.

<sup>12</sup> Daarbij kan het bijvoorbeeld gaan om ondersteuning bij het maken van samenvattingen, vertalingen en persberichten, herformuleringen in eenvoudiger taalgebruik en communicatie, zoals beeld en geluid, etc.

**De inzet van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek heeft een relatief laag risico**

Bij het Standpunt hoort een aanvullende, ondersteunende Handreiking. Daarin staan voorbeelden van toepassingsgebieden waar de inzet van generatieve AI een verlaagd risico inhoudt, zoals wetenschappelijk onderzoek.<sup>13</sup> Het RKI-kader gaat niet over hoog risico AI van de AI-verordening.

**Het RKI-kader biedt een gedragen en functioneel instrument voor inzet van generatieve AI**

Een van de uitgangspunten van het Standpunt is dat er voorafgaand aan de inzet van generatieve AI een passende risicoanalyse dient te zijn uitgevoerd.<sup>14</sup> Er bestaan diverse risicoanalyses voor de inzet van AI.<sup>15</sup> RKI's constateerden dat deze vaak generiek van aard en zeer omvangrijk zijn. Daarom kunnen ze, met name binnen het toepassingsgebied van wetenschappelijk onderzoek, buitenproportionele drempels opwerpen. Dit was mede aanleiding om te komen tot het RKI-kader.

Omdat wetenschappelijk onderzoek specifieke doeleinden nastreeft en zich in een specifieke context voltrekt, zijn generieke kaders in de regel onvoldoende toereikend.<sup>16</sup> Een van de doelen van het RKI-kader is het bieden van een gezamenlijk, gedragen en functioneel instrument aan overheidsorganisaties, specifiek voor het afwegen van kansen en risico's van de inzet van generatieve AI in het toepassingsgebied wetenschappelijk onderzoek.

**Het RKI-kader gaat niet uitputtend in op alle mogelijk relevante wetten en regels**

Een uitgangspunt van het Standpunt is dat de inzet van generatieve AI moet voldoen aan bestaande wet- en regelgeving.<sup>17</sup> Het RKI-kader gaat niet uitputtend in op alle mogelijk relevante wetten en regels. Het Algoritmekader van de overheid geeft inzicht in bestaande wet- en regelgeving die van toepassing kunnen zijn.<sup>18</sup>

Bij het inkopen, gebruiken en ontwikkelen van generatieve AI zullen diverse wetten en regels van belang zijn. Zoals de AI-verordening, privacywetgeving zoals de Algemene verordening gegevensbescherming (hierna: “AVG”) en regels over auteursrecht, inkoop en bestuursrecht. Het kan dus zo zijn dat er voor de inzet van generatieve AI nog andere (juridische) vereisten zijn waarop het RKI-kader niet ziet. Een voorbeeld is de AVG, die van toepassing is als de inzet van generatieve AI een verwerking van persoonsgegevens inhoudt.

Omdat het RKI-kader invulling geeft aan de inzet van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek, zal in voorkomende gevallen de wetenschappelijke exceptie van de AI-verordening relevant zijn. Deze houdt in dat de AI-verordening niet van toepassing is op “AI-systemen of AI-modellen, met inbegrip van hun output, die specifiek zijn ontwikkeld en in gebruik gesteld met wetenschappelijk onderzoek en wetenschappelijke ontwikkeling als enig doel”.<sup>19</sup> Ten aanzien van de AVG is relevant dat deze flexibiliteit biedt ten aanzien van verwerking van persoonsgegevens ten behoeve van wetenschappelijke onderzoeksdoeleinden.<sup>20</sup>

---

<sup>13</sup> Handreiking, par. 3.2, p. 10.

<sup>14</sup> Standpunt, p. 2 en 3.

<sup>15</sup> Voorbeelden zijn de Aanpak Begeleidingsethiek (ABE), De Ethische Data Assistent (DEDA), het Impact Assessment Mensenrechten en Algoritmen (IAMA), het Algoritme Impact Assessment (AIA) en de Datadialoog.

<sup>16</sup> Handreiking, par. 3.2, p. 10.

<sup>17</sup> Standpunt, p. 2. Voor RKI's kan dit ook eventueel toepasselijke sectorale wet- en regelgeving omvatten.

<sup>18</sup> Handreiking, par. 1.2, p. 5 en H. 5, p. 15 en 16.

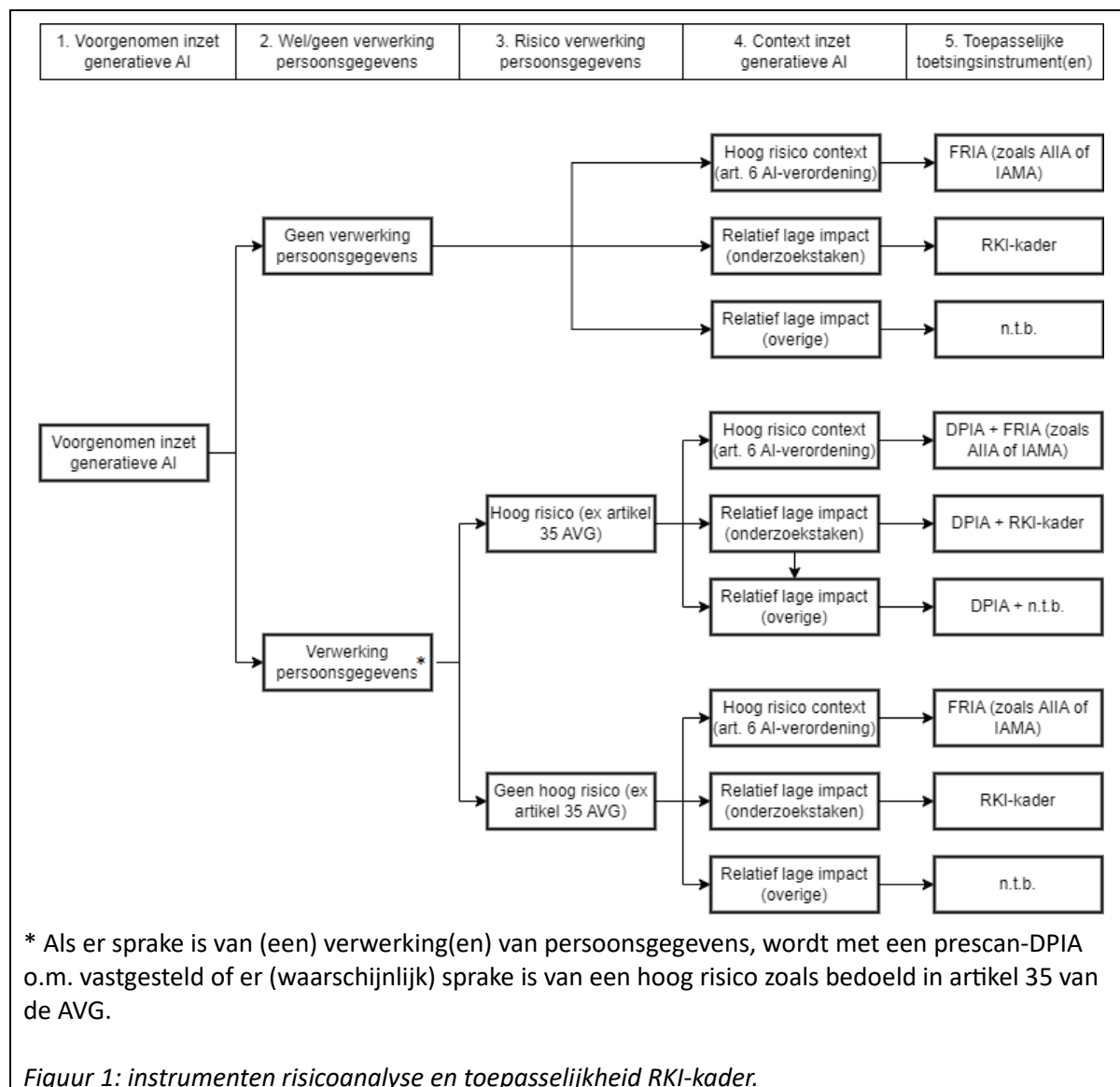
<sup>19</sup> Artikel 2, zesde lid, van de AI-verordening.

<sup>20</sup> Zie bijvoorbeeld artikel 5, eerste lid, aanhef en onderdeel b, artikel 9, tweede lid onderdeel j, artikel 89, en de overwegingen 50 en 156 tot en met 163 van de AVG.



## Wanneer is gebruik van het RKI-kader passend en wanneer andere toetsingsinstrumenten?

Voorafgaand aan de inzet van generatieve AI kan het noodzakelijk of wenselijk zijn om ook andere toetsingsinstrumenten te gebruiken. Figuur 1 helpt bij het bepalen welk(e) toetsingsinstrument(en) van toepassing is/zijn.



De definities van de afkortingen uit Figuur 1 zijn als volgt:

- De figuur definieert enkele gevallen waarin het toepasselijke toetsingsinstrument nog niet is bepaald. Denkbaar is dat er voor dergelijke toepassingsgebieden – waar de inzet van generatieve AI een relatief lage impact heeft – specifieke, lichtere instrument voor risicoanalyse worden ontwikkeld, vergelijkbaar met het RKI-kader.<sup>21</sup> Het kan raadzaam zijn om deze te gebruiken.
- AIIA: het Artificial Intelligence Impact Assessment, een hulpmiddel voor het maken van afwegingen bij het inzetten van AI in een project;<sup>22</sup>

<sup>21</sup> Handreiking, par. 3.2, p. 10.

<sup>22</sup> Ministerie Infrastructuur en Waterstaat, *AI Impact Assessment*, versie 2.0, december 2024, via: <https://www.rijksoverheid.nl/documenten/rapporten/2022/11/30/ai-impact-assessment-ministerie-van-infrastructuur-en-waterstaat>.

- *DPIA*: het Data Protection Impact Assessment, een toetsingsinstrument dat een verwerkingsverantwoordelijke<sup>23</sup> uitvoert bij verwerkingen van persoonsgegevens die waarschijnlijk een hoog risico inhouden voor de rechten en vrijheden van natuurlijke personen.<sup>24</sup>
- *FRIA*: het Fundamental Rights Impact Assessment, een voor gebruikersverantwoordelijke verplicht toetsingsinstrument dat zij uitvoeren als zij (generatieve) AI willen inzetten in een hoog-risico context;<sup>25</sup>
- *IAMA*: het Impact Assessment Mensenrechten en Algoritmes, een instrument dat helpt om een afgewogen discussie te voeren bij de afweging om wel of niet een algoritmische toepassing te gaan ontwikkelen;<sup>26</sup>

## 1.2. Status en beheer

Op 14 november 2024 deelden RKI-directeuren hun zorgen en aandachtspunten met betrekking tot proportionaliteit van de randvoorwaarden voor het gebruik van generatieve AI met het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (CIO Rijk). De grootste last leek te volgen uit het verplicht uitvoeren van generieke, uitgebreide risicoanalyses bij voorgenomen inzet van generatieve AI bij wetenschappelijk onderzoek. Daarom stelde het RKI-netwerk aan CIO Rijk voor om in samenwerking te komen tot een passend kader voor die context.

In zijn brief van 26 november 2024 reageerde de CIO Rijk positief op het voorstel tot samenwerking tussen het RKI-netwerk en CIO Rijk. Daarop werd binnen het RKI-netwerk de Ad hoc werkgroep Generatieve AI opgericht. Deze kwam van januari tot en met mei 2025 acht keer bijeen. Het doel van deze werkgroep was om in gezamenlijkheid te komen tot een passend, gedragen en functioneel kader voor waardevolle, verantwoorde en ethische inzet van generatieve AI binnen wetenschappelijk onderzoek. Het RKI-kader kwam in consensus tot stand binnen deze werkgroep, samen met CIO Rijk.

De RKI-directeuren stelden het RKI-kader vast op 26 mei 2025. Het RKI-netwerk moedigt het gebruik van het RKI-kader door de leden aan. Door hiermee ervaring op te doen, vergroten zij hun kennis en vaardigheden over waardevolle, verantwoorde en ethische inzet van generatieve AI binnen wetenschappelijk onderzoek. RKI's kunnen het RKI-kader zelf doorontwikkelen, zodat zij zelf op passende wijze rekening kunnen houden met specifieke omstandigheden en risico's. Het RKI-netwerk is eigenaar van dit document.<sup>27</sup> Het document wordt opgenomen in het Algoritmekader van de overheid.<sup>28</sup>

Het RKI-netwerk zal aan het einde van 2025 de ervaringen die RKI's opdeden met dit kader inventariseren. Indien het gewenst is, kan dit leiden tot verbetering of aanpassing van het RKI-kader. In dat geval informeert en/of consulteert het RKI-netwerk CIO Rijk. Ook zal dan stil worden gestaan bij de cyclus om het RKI-kader te blijven verbeteren.

<sup>23</sup> Artikel 4, onderdeel 7, van de AVG.

<sup>24</sup> Artikel 35 van de AVG, RKI's gebruiken hiervoor de Model DPIA Rijksdienst, welke is te raadplegen via: <https://www.cip-overheid.nl/producten-en-diensten/producten?product=de-8216-rijkspia-8217->. Het is mogelijk het IAMA te integreren in de Model DPIA Rijksdienst, zie via: <https://www.cip-overheid.nl/producten-en-diensten/producten?product=Handreiking+gebruik+IAMA+samen+met+Model+DPIA+Rijksdienst>. De risicoanalyse van het RKI-kader kan – desgewenst – als bijlage worden toegevoegd bij een DPIA.

<sup>25</sup> Vgl. artikel 27 van de AI-verordening. In het verlengde van de FRIA wordt opgemerkt dat bij hoog risico AI-toepassingen meer maatregelen nodig kunnen zijn. Voorbeelden zijn conformiteitsbeoordelingen en aanvullende aandacht voor aspecten als datakwaliteit, risicomanagementsystemen, klachtregelingen en monitoring.

<sup>26</sup> Prof. mr. Janneke Gerards, Dr. Mirko Tobias Schäfer, Arthur Vankan en Iris Muis (Universiteit Utrecht), *Impact Assessment Mensenrechten en Algoritmes*, via: <https://www.rijksoverheid.nl/documenten/rapporten/2021/02/25/impact-assessment-mensenrechten-en-algoritmes>.

<sup>27</sup> Binnen het RKI-netwerk is een eerste aanspreekpunt aangewezen, verantwoordelijk voor het beheer van het RKI-kader.

<sup>28</sup> Het Algoritmekader voorziet in een (niet-uitputtende) lijst met instrumenten, zie Standpunt, p. 2.

## 2. Beginselen voor het gebruik van generatieve AI binnen wetenschappelijk onderzoek

**Overheidsorganisaties die generatieve AI (willen) gebruiken binnen het toepassingsgebied wetenschappelijk onderzoek:**

### 1. Innoveren en benutten de kansen van generatieve AI

Wetenschappers kunnen generatieve AI inzetten om onderzoek efficiënter en creatiever te maken. Daarnaast hebben zij de mogelijkheid om zelf generatieve AI te bestuderen of te ontwikkelen.<sup>29</sup>

Als overheidsorganisaties generatieve AI (willen) gebruiken binnen het toepassingsgebied wetenschappelijk onderzoek, maken zij gebruik van toepassingen die op verschillende manieren worden aangeboden. Elke methode heeft unieke voordelen en veiligheidsrisico's, die invloed hebben op hoe je met deze technologie omgaat. De meest voorkomende toepassingen van generatieve AI zijn:

- Openbare generatieve AI-toepassingen en webdiensten;
- Ingebedde generatieve AI-toepassingen;
- Openbare generatieve AI-API's;
- Lokale ontwikkeling, en;
- Cloudoplossingen.<sup>30</sup>

Bij inkoop van generatieve AI kunnen overheidsorganisaties aanhaken op overheidsbrede, rijksbrede of departementale overeenkomsten. Daarnaast kunnen zij gezamenlijk of zelf contracteren. Afspraken met leveranciers van generatieve AI adresseren relevante onderwerpen,<sup>31</sup> overeenkomstig met de Algemene Rijksvoorwaarden bij IT-overeenkomsten (ARBIT). Naast juridische vereisten is het goed om stil te staan bij aspecten als ethiek, duurzaamheid, en soevereiniteit.<sup>32</sup> Overeenkomstig met het Standpunt hebben open source oplossingen en Europese dienstverleners de voorkeur en dienen mogelijkheden van *offline* en *on premise* generatieve AI te worden verkend.<sup>33</sup>

Het is randvoorwaardelijk dat overheidsorganisaties hun informatiehuishouding op orde hebben. Overheidsorganisaties die generatieve AI (willen) inzetten binnen het toepassingsgebied wetenschappelijk onderzoek, gebruiken doorgaans generatieve AI-toepassingen die lokaal gehost of cloudgebaseerd zijn. Ze kunnen de generatieve AI zelf beheren, of een gecontracteerde partner doet dat. Hierbij is relevant dat overheidsorganisaties de Baseline Informatiebeveiliging Overheid (hierna: “**BIO**”) hanteren als basis voor hun informatiebeveiliging.<sup>34</sup>

### 2. Begeleiden en ondersteunen onderzoekers om generatieve AI verantwoord in te zetten

Op grond van de AI-verordening dienen organisaties die AI-systemen gebruiken ervoor te zorgen dat hun medewerkers daarover voldoende kennis en vaardigheden hebben. Het organiseren van deze AI-

---

<sup>29</sup> Handreiking, par. 3.2, p. 10.

<sup>30</sup> Handreiking, par. 4.5, p. 13 en 14.

<sup>31</sup> Standpunt, p. 3. Belangrijke onderwerpen voor deze afspraken (en/of bedrijfsvoorwaarden) zijn datadeling, dataretentie en het al dan niet toestaan van het hertrainen van modellen met de input van gebruikers (zoals prompts).

<sup>32</sup> Handreiking, par. 5.4, p. 16. De Nederlandse Digitaliseringsstrategie, die naar waarschijnlijkheid in juni 2025 verschijnt, adresseert gezamenlijke aanpak van inkoop van digitaliseringsoplossingen, inclusief generatieve AI. Ook Hoofdstuk 3 van het RKI-kader gaat in op deze onderwerpen.

<sup>33</sup> Standpunt, p. 4.

<sup>34</sup> De Baseline Informatiebeveiliging Overheid is te raadplegen via: <https://zoek.officielebekendmakingen.nl/stcrt-2020-7857.html>.

geletterdheid is onder meer afhankelijk van de context waarin organisaties AI-systemen gebruiken en wat de doelgroepen zijn. Aspecten als technische kennis, ervaring en het opleidingsniveau van medewerkers zijn hierbij ook relevant.<sup>35</sup>

Voor overheidsorganisaties die generatieve AI (willen) gebruiken binnen het toepassingsgebied wetenschappelijk onderzoek, betekent dit dat zij aandacht hebben voor de ontwikkeling van kennis en vaardigheden over de inzet en (door-)ontwikkeling van generatieve AI binnen die context. Daarbij zullen zij zich richten tot diverse groepen relevante medewerkers,<sup>36</sup> zoals bijvoorbeeld onderzoekers, data scientists en andere wetenschappelijk medewerkers. Overheidsorganisaties besteden aandacht aan het vergroten van hun kennis over waardevolle, verantwoorde en ethische inzet van generatieve AI. Dat kan bijvoorbeeld met interne of externe opleidingen en met behulp van handreikingen en instrumenten, zoals het RKI-kader.<sup>37</sup> Relevante onderwerpen zijn bijvoorbeeld het verifiëren van input en uitkomsten, bescherming van privacyrechten en intellectuele eigendom en het tegengaan van vooroordelen.

### **3. Houden actief toezicht op het gebruik van generatieve AI**

Overheidsorganisaties dienen de juiste rollen te betrekken bij besluitvorming over de inzet van generatieve AI. Hiervoor is het opzetten van een governance structuur rondom (generatieve) AI, welke past binnen de beleidsmatige kaders, randvoorwaardelijk.<sup>38</sup>

Om de inzet van generatieve AI te verantwoorden, beschrijven overheidsorganisaties die generatieve AI (willen) gebruiken binnen het toepassingsgebied wetenschappelijk onderzoek welk doel de inzet ervan dient. Het instrument van Hoofdstuk 3 van het RKI-kader faciliteert daarin. Naar aanleiding van de uitkomsten van de risicoanalyse nemen overheidsorganisaties op een voor hen passende wijze een besluit over de inzet van generatieve AI.<sup>39</sup>

Transparantie en verantwoording worden vergroot door een actueel overzicht bij te houden van alle generatieve AI-toepassingen die worden gebruikt. Daarnaast kunnen overheidsorganisaties ervoor kiezen om individuele projecten waarbij zij generatieve AI gebruiken opnemen in het landelijk Algoritmeregister.<sup>40</sup>

Bij het doen van wetenschappelijk onderzoek kunnen overheidsorganisaties mogelijkheden en beperkingen van generatieve AI analyseren. Het uitgangspunt van *Open Science* houdt in dat zij hierover hun bevindingen en aanbevelingen in beginsel zullen delen met collega's, elkaar en de academische gemeenschap. Dat zorgt er onder meer voor dat:

- generatieve AI beter wordt gebruikt;
- kennis en kunde over het gebruik van generatieve AI beter en gericht ontwikkelt, en;
- negatieve uitkomsten door het gebruik van generatieve AI beter worden geadresseerd.

### **4. Kunnen dit kader gebruiken voor de verantwoorde inzet van generatieve AI**

Overheidsorganisaties die generatieve AI (willen) gebruiken binnen het toepassingsgebied wetenschappelijk onderzoek kunnen, waar zij dat wenselijk achten, het RKI-kader toepassen en/of ernaar verwijzen in bijvoorbeeld eigen beleid, kaders en werkinstructies. Bovendien bestaat de

---

<sup>35</sup> De verplichting voor het organiseren van AI-geletterdheid volgt uit artikel 4 van de AI-verordening.

<sup>36</sup> Voor een niet-uitputtende lijst, zie Handreiking, par. 4.2, p. 11 en 12.

<sup>37</sup> Handreiking, par. 4.2, p. 11 en 12.

<sup>38</sup> Standpunt, p. 3 en Handreiking, par. 4.3, p. 12.

<sup>39</sup> Standpunt, p. 3.

<sup>40</sup> Handreiking, par. 4.3, p. 12 en 13, het Algoritmeregister is te raadplegen via <https://algoritmes.overheid.nl/nl>.

mogelijkheid om het RKI-kader verder te specificeren en/of aan te vullen voor de eigen organisatie, bijvoorbeeld als er specifieke risico's zijn vanwege het type onderzoek dat zij doen.

Als overheidsorganisaties generatieve AI inkopen, geldt het Standpunt. Het RKI-kader kan helpen bij het maken van betere afspraken bij inkoop, in aanvulling op de Algemene Rijksvoorwaarden bij IT-overeenkomsten (ARBIT). Het Standpunt geldt ook als overheidsorganisaties onderzoeken uitbesteden aan externe onderzoeksbureaus. Het RKI-kader kan ook helpen bij het vormgeven van dat type samenwerkingen.

### **Onderzoekers van overheidsorganisaties die generatieve AI (willen) gebruiken binnen het toepassingsgebied wetenschappelijk onderzoek:**

#### **1. *Zijn verantwoordelijk voor de borging van wetenschappelijke integriteit.***

Voordat onderzoekers ervoor kiezen generatieve AI in te zetten binnen het toepassingsgebied wetenschappelijk onderzoek, beschrijven zij het doel dat hiermee wordt gediend en analyseren zij risico's om na te gaan of de inzet verantwoord en veilig is.<sup>41</sup> Het RKI-kader, en meer specifiek het model voor risicoanalyse van Hoofdstuk 3, helpt om systematisch te kijken naar de effecten van de inzet van generatieve AI.<sup>42</sup>

Onderzoekers kunnen onderzoek doen naar generatieve AI. Als zij generatieve AI gebruiken bij de uitvoering van onderzoek voor grotere effectiviteit en/of efficiëntie, is belangrijk uitgangspunt dat zij het gebruiken als hulpmiddel. Zij gebruiken het dan dus niet voor vaardigheden die zij zelf niet bezitten of begrijpen. Het kan bijvoorbeeld worden ingezet om taken die veel tijd kosten (deels) te automatiseren. Onderzoekers blijven verantwoordelijk voor onderzoeksresultaten, en daarmee ook de borging van onderzoeksintegriteit en de integriteit van de uitkomsten die wordt gegenereerd door of met ondersteuning van generatieve AI.<sup>43</sup> Zij benaderen de uitkomsten kritisch. Onderzoekers kennen eventuele beperkingen van de toepassingen, zoals vooringenomenheid, niet relevante resultaten en onnauwkeurigheden (ook: "hallucinaties"), of zij doen daar onderzoek naar.<sup>44</sup>

Als het binnen de academische gemeenschap de norm is om in een bepaalde situatie geen generatieve AI in te zetten, gebruiken onderzoekers het niet. Of ze motiveren waarom ze hier bewust van afwijken en communiceren ze dat actief.<sup>45</sup> Wetenschappelijke integriteit houdt ook in dat onderzoekers generatieve AI niet gebruiken om te manipuleren, bijvoorbeeld door originele onderzoeksgegevens moedwillig te vervalsen of te veranderen.

#### **2. *Ontwikkelen en onderhouden hun kennis over generatieve AI.***

Overheidsorganisaties besteden actief aandacht aan kennisontwikkeling over de inzet en (door-) ontwikkeling van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek.<sup>46</sup> Het

---

<sup>41</sup> Standpunt, p. 2.

<sup>42</sup> Er kan voor gekozen worden om hierbij ook een (multidisciplinaire en diverse) ethische commissie te betrekken, zie Handreiking, par. 4.2, p. 11.

<sup>43</sup> Om deze verantwoordelijkheid te kunnen dragen, is het belangrijk dat overheidsorganisaties onderzoekers begeleiden en ondersteunen onderzoekers om generatieve AI verantwoord in te zetten (zie p. 11).

<sup>44</sup> Europese Commissie, ERA Forum Stakeholders' document, *Living guidelines on the responsible use of generative AI in research*, maart 2024, par. 2.1, onder 1, p. 6.

<sup>45</sup> Momenteel is het gebruik van generatieve AI-modellen in het hele NWO-beoordelingsproces van aanvragen bijvoorbeeld ontoelaatbaar, zie via: <https://www.nwo.nl/nieuws/nwo-komt-met-voorlopige-richtlijnen-gebruik-ai>. Verdedigbaar is dat dit risico kan worden gemitigeerd door – conform dit RKI-kader – een lokaal gehoste generatieve AI-toepassing te gebruiken.

<sup>46</sup> Zie p. 10 en 11.

past binnen de taken van onderzoekers om op de hoogte te blijven van ontwikkelingen en *best practices*. Zij delen hun opgedane kennis actief met collega's en andere belanghebbenden.<sup>47</sup>

Van medewerkers die onderzoekers van overheidsorganisaties ondersteunen bij het verantwoord inzetten van generatieve AI, wordt verwacht dat zij voldoende kennis hebben van de juridische, ethische en organisatorische maatregelen die nodig zijn om generatieve AI verantwoord in te zetten, alsmede de context van het toepassingsgebied wetenschappelijk onderzoek.<sup>48</sup>

### **3. Gebruiken generatieve AI op een transparante manier.**

Het beginsel van transparantie is essentieel voor integer wetenschappelijk onderzoek én verantwoorde inzet van generatieve AI. Om transparantie te bieden, zijn onderzoekers in staat om aan te geven welke generatieve AI-toepassingen zij gebruiken binnen het onderzoeksproces. Dat kan door informatie over de gebruikte toepassing, zoals de naam, versie en datum, en te beschrijven hoe deze is gebruikt. Daarbij gaan onderzoekers in op de wijze waarop de inzet van generatieve AI het onderzoeksproces beïnvloedt. Het model voor risicoanalyse (Hoofdstuk 3) helpt om dit onderwerp in een vroeg stadium te adresseren.

Generatieve AI-toepassingen kunnen verschillende uitkomsten produceren uit dezelfde input. Onderzoekers houden rekening met deze kans op willekeur. Onderzoekers zijn – waar nodig – in staat om de input en uitkomsten beschikbaar te maken, volgens principes van Open Science.<sup>49</sup> Dit omvat ook de diverse uitkomsten en onderzoeksresultaten.

Als onderzoekers gegenereerde content gebruiken in publicaties, verduidelijken zij dat deze met generatieve AI zijn gemaakt.<sup>50</sup>

### **4. Houden zich bij het gebruik van generatieve AI aan toepasselijke wet- en regelgeving.**

Verantwoorde inzet van generatieve AI houdt in dat onderzoekers zich houden aan relevante wet- en regelgeving.<sup>51</sup> Onderzoekers begrijpen dus de voor hen relevante risico's met betrekking tot bijvoorbeeld privacy, informatiebeveiliging, vertrouwelijkheid en intellectuele eigendomsrechten.<sup>52</sup> Voor RKI's kan dit ook eventueel toepasselijke sectorale wet- en regelgeving omvatten. Dat betekent dat van hen onder meer het volgende wordt verwacht:

- Onderzoekers betrekken tijdig expertise op het gebied van privacy en informatiebeveiliging om risico's te analyseren die voortvloeien uit bijvoorbeeld de privacyopties van tools, de partij die de tool beheert, waar deze wordt gehost en wat implicaties zijn voor alle geüploade data en informatie.<sup>53</sup> Dit is ook relevant als onderzoekers samenwerken met externe onderzoeksbureaus.
- Onderzoekers beschermen hun ongepubliceerde stukken en gegevenssets.

---

<sup>47</sup> Europese Commissie, ERA Forum Stakeholders' document, *Living guidelines on the responsible use of generative AI in research*, maart 2024, par. 2.1, onder 5, p. 7.

<sup>48</sup> Handreiking, par. 4.2, p. 11 en 12.

<sup>49</sup> Europese Commissie, ERA Forum Stakeholders' document, *Living guidelines on the responsible use of generative AI in research*, maart 2024, par. 2.1, onder 2, p. 6.

<sup>50</sup> Handreiking, par. 9.2, p. 22; artikel 50 van de AI-verordening.

<sup>51</sup> Standpunt, p. 2.

<sup>52</sup> Het is derhalve raadzaam dat overheidsorganisaties dit meenemen bij de begeleiding en ondersteuning van onderzoekers bij verantwoorde inzet van generatieve AI (zie p. 11)

<sup>53</sup> Vgl. Europese Commissie, ERA Forum Stakeholders' document, *Living guidelines on the responsible use of generative AI in research*, maart 2024, par. 2.1, onder 3, p. 6 en 7.

- Onderzoekers gebruiken diensten die hun eigen organisatie aanbiedt.<sup>54</sup>
- Onderzoekers verwerken geen persoonsgegevens in een generatief AI-systeem als daarvoor geen grondslag is. Als er met de inzet van generatieve AI persoonsgegevens worden verwerkt, dient te worden voldaan aan relevante vereisten uit toepasselijke (privacy-) wetgeving. Het is raadzaam om in dit geval een privacy professional van binnen of buiten de organisatie te consulteren.
- De kans is groot dat generatieve AI is getraind met auteursrechtelijk beschermd materiaal. Doordat artikel 53 van de AI-verordening transparantievereisten aan aanbieders van generatieve AI stelt, komt er over dit onderwerp meer duidelijkheid en zekerheid.<sup>55</sup> Onderzoekers dienen zich er derhalve van bewust te zijn dat uitkomsten van generatieve AI kunnen zijn gebaseerd op auteursrechtelijk beschermde werken. Als onderzoekers twijfels hebben of er met het gebruik van generatieve AI inbreuk wordt gemaakt op het auteursrecht, vragen zij advies aan een auteursrechtsspecialist van binnen of buiten de organisatie en/of gebruiken zij de uitkomsten niet.<sup>56</sup> Onderzoekers respecteren het auteurschap van anderen en citeren waar nodig. Als er onderzoek wordt gedaan naar generatieve AI, kan het trainen ervan generatieve AI met auteursrechtelijk beschermde werken zijn toegestaan op basis van de uitzondering voor tekst- en datamining.<sup>57</sup>
- Onderzoekers die generatieve AI (willen) gebruiken binnen het toepassingsgebied wetenschappelijk onderzoek gebruiken in beginsel geen generatieve AI-toepassingen onder consumentenvoorwaarden, zoals ChatGPT of andere publiek toegankelijke versies. Daarover kunnen overheidsorganisaties immers geen goede afspraken maken met leveranciers. Echter, verkenningen of privégebruik van nieuwe (generatieve) AI-systemen zijn toegestaan in een gecontroleerde niet-productieomgeving, mits er geen gevoelige of vertrouwelijke data worden gebruikt.<sup>58</sup> Als kan worden onderbouwd dat van een dergelijke situatie sprake is, hoeft het RKI-kader niet te worden doorlopen om het uitgangspunt van waardevolle, verantwoorde en ethische inzet van generatieve AI te kunnen aantonen. Ook is er ruimte voor onvermijdbaar gebruik van generatieve AI, zoals bij het gebruik van zoekmachines of automatisch vertaalde websites. Het uitgangspunt is dat onderzoekers daar gewetensvol mee omgaan. Zij zijn zich ervan bewust dat uitkomsten gebaseerd zijn op generatieve AI en controleren deze.<sup>59</sup>

---

<sup>54</sup> Zie p. 10. Het is randvoorwaardelijk dat overheidsorganisaties hun informatiehuishouding op orde hebben.

<sup>55</sup> Handreiking, par. 5.3, p. 15.

<sup>56</sup> Handreiking, par. 9.1, p. 22.

<sup>57</sup> Handreiking, par. 5.3, p. 16.

<sup>58</sup> Standpunt, p. 3 en Handreiking, par. 4.4, p. 13.

<sup>59</sup> Handreiking, par. 9.2, p. 22.

### 3. Risicoanalyse voor het gebruik van generatieve AI binnen wetenschappelijk onderzoek

Hoofdstuk 2 van het RKI-kader ging in op randvoorwaarden van de inzet van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek. In het verlengde daarvan, biedt dit hoofdstuk een passend instrument voor risicoanalyse. Deze dient als basis voor een *interdisciplinaire discussie* tussen verschillende rollen, met de juiste kennis en vaardigheden, over de kansen en risico's van generatieve AI.<sup>60</sup> Bijlage 1 bevat hiervoor een model. De uitkomsten kunnen worden uitgewerkt in een beslisdocument waarmee een besluit kan worden genomen over de inzet van generatieve AI.

#### Onderdeel A: beschrijving van de concrete situatie

In Onderdeel A omschrijven overheidsorganisaties die generatieve AI willen gebruiken binnen het toepassingsgebied wetenschappelijk onderzoek de concrete situatie. Wat is het doel van de inzet van generatieve AI? Wat is de impact van het wel of juist niet inzetten van generatieve AI? En voor welk probleem gaat het een oplossing bieden?

De volgende vragen zijn relevant. De antwoorden worden bondig vastgelegd.

**1. Welk doel wordt gediend met de inzet van generatieve AI?**

Leg uit wat het probleem is waarvoor de beoogde inzet van generatieve AI een oplossing moet bieden. Definieer vervolgens het doel van de inzet van generatieve AI zo concreet en specifiek mogelijk. Het zo expliciet mogelijk maken van de doelstelling is belangrijk, omdat de prestaties van de toepassing in een later stadium moeten kunnen worden getoetst aan de doelstelling.

**2. Is er sprake van een voorgenomen inzet van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek? Wordt het niet gebruikt in een andere context?**

Leg uit of generatieve AI gaat worden gebruikt of (door-)ontwikkeld. Bespreek en beschrijf dat het wordt ingezet binnen het toepassingsgebied wetenschappelijk onderzoek. Als het wordt ingezet in een andere context, kunnen risico's van een andere aard zijn.

**3. Wat zijn de kansen en mogelijkheden die de inzet van generatieve AI biedt?**

Bespreek en beschrijf de kansen van generatieve AI in de concrete situatie. Voorbeelden van kansen zijn het vergroten van de kennisbasis, effectiviteit en efficiëntie, kwaliteitsverbetering of kostenbesparing.

**4. Welke generatieve AI-toepassing(en) gaat gebruikt worden?**

Geef aan welke generatieve AI-toepassing(en) gaat/gaan worden gebruikt. Is er sprake van:

- Openbare generatieve AI-toepassingen en webdiensten;
- Ingebedde generatieve AI-toepassingen;
- Openbare generatieve AI-API's;
- Lokale ontwikkeling, en/of;
- Cloudoplossingen.

**5. Waarom is/zijn deze generatieve AI toepassing(en) geschikt?**

Bespreek en beschrijf waarom te voorzien is dat generatieve AI in het concrete geval een betere oplossing biedt dan de inzet van andere methoden of technieken (noodzaak en proportionaliteit). Zijn er alternatieven onderzocht die minder impact hebben (subsidiariteit)?

---

<sup>60</sup> Handreiking, par. 4.4, p. 13. Om de interdisciplinaire discussie constructiever en scherper te laten verlopen, kan bijvoorbeeld ook een 'kritisch tegenlezer' en/of een 'socratische vragensteller' worden betrokken.



**6. *Waarom zijn data die als input dienen voor de generatieve AI toepassing(en) geschikt?***

Bespreek en beschrijf de geschiktheid van de data die als input worden gebruikt. Waar komen deze vandaan? En zijn de kwaliteit en betrouwbaarheid ervan voldoende voor de beoogde toepassing(en)?

**7. *Welke partijen en personen zijn er bij de ontwikkeling/het gebruik/het onderhoud van het algoritme betrokken?***

Bespreek en beschrijf welke partijen en personen zijn betrokken bij de inzet van generatieve AI. Is te voorzien dat zij voldoen aan bestaande wet- en regelgeving, zoals de AVG, de AI-verordening en de Grondwet? Zijn er acties nodig om dit (beter) te adresseren?

**Onderdeel B: meest relevante thema's bij waardevolle, verantwoorde en ethische inzet van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek**

RKI's definieerden op basis van het Standpunt, de Handreiking en uitgangspunten van wetenschappelijke integriteit de belangrijkste thema's bij waardevolle, verantwoorde en ethische inzet van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek. Aan de hand van Onderdeel B wordt besproken en (bondig) beschreven of er risico's zijn te voorzien met betrekking tot deze acht thema's. Hierbij wordt tevens ingegaan op strategieën of maatregelen om eventuele risico's te verkleinen of weg te nemen.<sup>61</sup>

**1. *Menselijke controle en menselijk toezicht***

Zijn er risico's te voorzien met betrekking tot menselijke controle en menselijk toezicht? Wat zijn strategieën of maatregelen om eventuele risico's te verkleinen of weg te nemen?

Dit uitgangspunt houdt in dat generatieve AI menselijke autonomie en beslissingen moeten ondersteunen. Ook bij onderzoek naar generatieve AI zijn menselijke controle en toezicht van belang, al zal het dan meer zien op beoordeling van juistheid door een onderzoeker.

Elementen die relevant zijn, kunnen bijvoorbeeld zijn dat:

- De wijze waarop menselijke controle en toezicht worden geïmplementeerd;
- Er een menselijke beoordeling achteraf mogelijk is;
- De toegevoegde waarde en proportionaliteit van de inzet van generatieve AI wordt geëvalueerd;
- In hoeverre er sprake is van automatische besluiten, die direct raken aan mensen, en in hoeverre er sprake is van betekenisvolle menselijke tussenkomst.<sup>62</sup>

**2. *Technische robuustheid***

Zijn er risico's te voorzien met betrekking tot technische robuustheid? Wat zijn strategieën of maatregelen om eventuele risico's te verkleinen of weg te nemen?

Voor technische robuustheid moet generatieve AI worden ontwikkeld met een preventieve benadering van risico's en op dusdanige wijze dat het zich betrouwbaar en zoals bedoeld gedraagt. Onbedoelde en onverwachte schade moet zo veel mogelijk worden beperkt, onacceptabele schade moet worden voorkomen. Elementen die hierbij relevant zijn, houden in dat er een plan is hoe om te gaan met de beperkingen van de technologie (zorgvuldigheid) en er

---

<sup>61</sup> Bijlage 2 bij het RKI-kader adresseert per thema van dit Onderdeel B drie voorbeelden van strategieën of maatregelen die potentiële risico's tegengaan of verkleinen.

<sup>62</sup> Zie ook artikel 22 van de AVG; Autoriteit Persoonsgegevens, *Advies artikel 22 AVG en geautomatiseerde selectie-instrumenten*, 10 oktober 2024, via: <https://open.overheid.nl/documenten/277154d5-99db-4f5c-8af9-eff5c5c33b7d/file>.

een passende strategie is die het mogelijk maakt om te stoppen met het gebruik van generatieve AI in het geval dat er (mogelijk) problemen optreden (veiligheid).

#### Zorgvuldigheid

- Er wordt nagegaan of de te gebruiken generatieve AI technisch robuust is;
- Er wordt nagegaan of de te gebruiken generatieve AI kan omgaan met fouten;
- De vereisten van informatiebeveiliging en privacywetgeving, zoals de BIO en de AVG, worden gewaarborgd.

#### Veiligheid

- De te gebruiken generatieve AI wordt gemonitord;
- Als er problemen zijn, kan er worden gestopt. Of er is een andere, meer passende strategie;
- De vereisten van informatiebeveiliging en privacywetgeving, zoals de BIO en de AVG, worden gewaarborgd.

### **3. Privacy en datagovernance**

Zijn er risico's te voorzien met betrekking tot privacy en datagovernance? Wat zijn strategieën of maatregelen om eventuele risico's te verkleinen of weg te nemen?

Als generatieve AI een verwerking van persoonsgegevens inhoudt, worden wetten en regels over privacy en gegevensbescherming nageleefd. Daartoe dient een (prescan-)DPIA.<sup>63</sup>

Als er bijvoorbeeld sprake is van geanonimiseerde gegevens, leg je in dit onderdeel uit waarom dat zo is. Dat geldt ook voor andere relevante aspecten rond privacy en datagovernance.

### **4. Transparantie**

Zijn er risico's te voorzien met betrekking tot transparantie? Wat zijn strategieën of maatregelen om eventuele risico's te verkleinen of weg te nemen?

Transparantie houdt nauw verband met het beginsel van verantwoording en omvat elementen die relevant zijn voor de te gebruiken generatieve AI (o.a. *Open Science* en *Explainable AI*). Wetenschappelijk onderzoek waarbij generatieve AI studieobject is, kan zich bijvoorbeeld richten op het verkrijgen van inzicht onder welke voorwaarden generatieve AI goed functioneert en wanneer deze accuraat werkt. Het thema transparantie wordt dan op andere wijze aangevlogen.

Ten aanzien van transparantie zijn er drie aspecten te onderscheiden die op passende wijze geadresseerd dienen te worden: het onderliggende model, de overkoepelende applicatie en de onderliggende methodologische kennis. Hieronder zijn per aspect voorbeeldvragen opgenomen.

#### Model (training data, gewichten, metadata, licentie)

- Kunnen alle data, en metadata (training, architectuur) openbaar beschikbaar worden gesteld onder een open licentie? En de software onder een open source licentie?<sup>64</sup>
- Kunnen de metadata worden vrijgegeven, terwijl gewichten en data enkel intern of onder (restrictieve) voorwaarden toegankelijk zijn?
- Kunnen de naam, versie en licentievorm van de generatieve AI bekend worden gemaakt, maar de data, gewichten of architectuur niet?

---

<sup>63</sup> Aanbevolen wordt om deze parallel met deze risicoanalyse uit te voeren.

<sup>64</sup> Bij open (meta)data is het gangbaar om Creative Commons-licenties (CC) te gebruiken, zoals CC-BY. Voorbeelden van open source licenties voor software en code zijn MIT, GPL en EUPL, al kan EUPL ook data omvatten.

#### Toepassing (implementatie, versie, prestatie, datastroom)

- Kan de complete technische documentatie van de applicatie, inclusief onderliggende modelnummers, -versies en -prestaties, binnen de context van het onderzoek worden gedeeld?
- Kan een beschrijving van de modelversie en prestaties in hoofdlijnen worden gedeeld? Is er een risico dat de gebruikte modellen in de toekomst niet beschikbaar zijn voor replicatie?
- Kan worden gedeeld waar en hoe generatieve AI wordt ingezet en met welke partijen daarbij wordt samengewerkt?
- Kan naar gebruikers en/of burgers worden gecommuniceerd dat generatieve AI wordt ingezet en welke rol deze techniek speelt?

#### Methodologie (kennis, validatie)

- Kunnen onderzoekers de methoden (op architectuurniveau) en de uitkomsten (technisch) verantwoorden?
- Is er een passend validatiemechanisme geïmplementeerd waaraan de uitkomsten van generatieve AI toepassingen onderworpen zijn?
- Kennen de onderzoekers de hoofdlijnen van de methodologie en voeren zij passende checks uit op de uitkomsten?

### **5. Reproduceerbaarheid**

Zijn er risico's te voorzien met betrekking tot reproduceerbaarheid? Wat zijn strategieën of maatregelen om eventuele risico's te verkleinen of weg te nemen?

Het gebruik van generatieve AI dient te passen binnen de principes van integer wetenschappelijk onderzoek. Dat houdt in dat onderzoek met of naar generatieve AI waarheidsgetrouw, valide, origineel en reproduceerbaar moet zijn. Generatieve AI heeft naar zijn aard een stochastische component die reproduceerbaarheid bemoeilijkt. Bespreek en beschrijf het onderwerp reproduceerbaarheid aan de hand van relevante elementen die de betrouwbaarheid van het onderzoek vergroten.

Elementen die relevant zijn, houden bijvoorbeeld in dat:

- Relevante informatie wordt gedocumenteerd, zoals experimenten, prompts en parameters;
- Er open source modellen worden gebruikt;
- Data worden gescheiden;
- Transparantie wordt vergroot (zie ook onder transparantie en methodologie).

### **6. Diversiteit, non-discriminatie en rechtvaardigheid**

Zijn er risico's te voorzien met betrekking tot diversiteit, non-discriminatie en rechtvaardigheid? Wat zijn strategieën of maatregelen om eventuele risico's te verkleinen of weg te nemen?

Dit vereiste is nauw verbonden met het beginsel van rechtvaardigheid. Gedurende het gehele proces moet met alle belanghebbenden rekening worden gehouden. Er dient te worden gezorgd voor inclusieve ontwerpprocessen en gelijke behandeling. Hierbij is relevant te benadrukken dat het doel van integer wetenschappelijk onderzoek is om een eerlijk, onafhankelijk en correct beeld te bieden van een bepaalde situatie, niet om direct te raken aan mensen.

Elementen die relevant zijn, kunnen bijvoorbeeld zijn dat:

- Er aandacht is voor de uitkomsten van de generatieve AI-toepassing. Gecontroleerd wordt dat deze niet ongewenst discrimineert, hallucineert en/of bevooroordeeld is;

- Er – waar dat mogelijk en passend is – belanghebbenden worden geraadpleegd. Dat kunnen bijvoorbeeld mensen zijn die werken met de generatieve AI. Of afgevaardigden van groepen waaraan het gebruik van generatieve AI raakt.

## 7. Milieu- en maatschappelijke waarden

Zijn er risico's te voorzien met milieu- en maatschappelijke waarden? Wat zijn strategieën of maatregelen om eventuele risico's te verkleinen of weg te nemen?

In het ideale scenario wordt generatieve AI gebruikt ten voordele van alle mensen, inclusief toekomstige generaties. Generatieve AI heeft de potentie om een maatschappelijke zorgen te verhelpen, maar er moet wel worden gezorgd dat dit op passende en duurzame wijze gebeurt.

### Duurzaamheidsaspecten

Het kost energie om generatieve AI-toepassingen te trainen en prompts te beantwoorden. Bovendien is er water nodig om servers te koelen en zitten er (zeldzame) grondstoffen in hardware, die na enkele jaren afgeschreven is. Maatregelen om de milieu-impact van de volledige toeleveringsketen van het AI-systeem te verkleinen, worden aangemoedigd.<sup>65</sup>

Duurzaamheidsaspecten kunnen bijvoorbeeld worden geadresseerd door:

- Ermee rekening te houden bij de keuze voor een specifieke techniek of model;<sup>66</sup>
- Milieueffecten (bijv. energie- en waterverbruik) voor training, *fine tuning* en implementatie van modellen in te schatten of te meten, en/of te documenteren;
- Nadelige milieueffecten te compenseren;
- Fossielvrije energie te gebruiken.

### Maatschappelijke aspecten

Hoewel generatieve AI-toepassingen kunnen worden gebruikt om sociale vaardigheden te vergroten, kunnen ze evenzeer bijdragen aan de verslechtering ervan. Dat zou ook gevolgen kunnen hebben voor het fysieke en geestelijke welzijn van mensen. De effecten van deze systemen moeten daarom zorgvuldig worden gemonitord en afgewogen. Dit aspect zal in andere mate worden geadresseerd in het geval van onderzoek naar (het gebruik van) generatieve AI.

## 8. Verantwoording

Zijn er risico's te voorzien met betrekking tot verantwoording? Wat zijn strategieën of maatregelen om eventuele risico's te verkleinen of weg te nemen?

Het beslisdocument is van belang om verantwoording af te leggen over de inzet van generatieve AI. Elementen die verder relevant zijn, kunnen bijvoorbeeld zijn dat:

- Afwegingen en keuzes worden gedocumenteerd, zodat deze traceerbaar zijn. Dit kan bijvoorbeeld relevant zijn bij het trainen, testen en ontwerpen van het model;
- Het proces hoe data wordt verkregen, geselecteerd en bewerkt wordt gedocumenteerd;
- De karakteristieken van de te gebruiken datasets worden gedocumenteerd;
- Methoden om risico's te identificeren worden gedocumenteerd, alsmede de maatregelen om deze tegen te gaan;
- Er wordt een besluit genomen over de inzet van generatieve AI, waarbij een passend, vastgesteld governance-model is gehanteerd.

<sup>65</sup> Er bestaat op dit moment nog geen vastgestelde systematiek om duurzaamheidsaspecten te adresseren.

<sup>66</sup> Dit omvat bijvoorbeeld ook de vraag of er niet beter een zoekmachine kan worden gebruikt dan generatieve AI.

## 4. Achtergrond bij het RKI-kader

Het RKI-kader is onder meer gebaseerd op de volgende bronnen:

### *Wetenschappelijke integriteit*

- Europese Gedragscode voor Wetenschappelijke Integriteit<sup>67</sup>
  - De Europese Gedragscode voor Wetenschappelijke Integriteit (ook: “ALLEA-code”) biedt onderzoekers beginselen om praktische, ethische en intellectuele uitdagingen te adresseren die inherent zijn aan het onderzoeksproces. De beginselen zijn:
    - ❖ Betrouwbaarheid, bij het waarborgen van kwaliteit van het onderzoek, zoals deze tot uiting komt in het ontwerp, de methodologie, de analyse en het gebruik van middelen;
    - ❖ Eerlijkheid, onderzoek moet op transparante, billijke, volledige en onbevooroordeelde wijze worden ontwikkeld, uitgevoerd, getoetst, gedocumenteerd en met anderen gedeeld;
    - ❖ Respect, voor collega's, deelnemers aan het onderzoek, de maatschappij, ecosystemen, cultureel erfgoed en het milieu, en;
    - ❖ Verantwoordelijkheid, voor het onderzoek van idee tot publicatie, voor het beheer en de organisatie ervan, voor opleiding, toezicht en mentorschap en de bredere impact van onderzoek.
- Nederlandse Gedragscode Wetenschappelijke Integriteit<sup>68</sup>
  - De Nederlandse Gedragscode Wetenschappelijke Integriteit definieert vijf principes van wetenschappelijke integriteit, 61 normen voor goede onderzoekspraktijken en zorgplichten voor instellingen. Met deze gedragscode sluit Nederland aan bij internationale ontwikkelingen op het gebied van wetenschappelijke integriteit. De principes kunnen worden gezien als ‘deugden’ van een goede onderzoeker. Ze brengen de onderzoeker ertoe om in verschillende omstandigheden de juiste keuze(s) te maken. De principes zijn:
    - ❖ Eerlijkheid
    - ❖ Zorgvuldigheid
    - ❖ Transparantie
    - ❖ Onafhankelijkheid
    - ❖ Verantwoordelijkheid
- Voorlopige richtlijn voor transparant en integer gebruik van generatieve AI van de NWO<sup>69</sup>
  - Uitgangspunt van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) is dat het gebruik van generatieve AI in de context van onderzoekswerk niet wordt uitgesloten, gezien de mogelijkheden en het ontwikkelingspotentieel ervan. Transparantie en controleerbaarheid van het onderzoeksproces en -resultaten zijn essentiële basisprincipes van wetenschap en onderzoek. Wetenschappelijke integriteit blijft ook bij het gebruik van generatieve AI het uitgangspunt. Elke wetenschapper die onderzoek doet op basis van een NWO-subsidie dient de Nederlandse Gedragscode Wetenschappelijke Integriteit te volgen.

---

<sup>67</sup> ALLEA - All European Academies, *Europese gedragscode voor wetenschappelijke integriteit*, herziene uitgave 2018, te raadplegen via: [https://www.allea.org/wp-content/uploads/2018/01/DU\\_ALLEA\\_Europese\\_gedragscode\\_voor\\_wetenschappelijke\\_integriteit.pdf](https://www.allea.org/wp-content/uploads/2018/01/DU_ALLEA_Europese_gedragscode_voor_wetenschappelijke_integriteit.pdf).

<sup>68</sup> KNAW, NFU, NWO, TO2-federatie, Vereniging Hogescholen en Vereniging van Universiteiten, *Nederlandse gedragscode wetenschappelijke integriteit*, 1 oktober 2018, te raadplegen via: <https://www.nwo.nl/nederlandse-gedragscode-wetenschappelijke-integriteit>.

<sup>69</sup> NWO, *NWO komt met voorlopige richtlijnen gebruik AI*, nieuwsbericht 16 januari 2024, te raadplegen via: <https://www.nwo.nl/nieuws/nwo-komt-met-voorlopige-richtlijnen-gebruik-ai>.

## **Overheidsbrede standpunt voor de inzet van generatieve AI, aanvullende handreiking, etc.**

- Overheidsbreed standpunt voor de inzet van generatieve AI ('Standpunt')<sup>70</sup>
  - De overheid kiest voor de waardevolle, verantwoorde en ethische inzet van generatieve AI. Het kabinet stimuleert innovatie en ziet het belang in van het gebruik van generatieve AI voor publieke waarden. Om generatieve AI waardevol, verantwoord en ethisch te gebruiken, (door) te ontwikkelen en/of in te kopen, stelt het overheidsbrede standpunt voor de inzet van generatieve AI enkele voorwaarden.
- Overheidsbrede handreiking voor de verantwoorde inzet van generatieve AI ('Handreiking')<sup>71</sup>
  - De Handreiking gaat nader in op de technologische, organisatorische, ethische en juridische (rand)voorwaarden die van meerwaarde zijn om als overheidsorganisatie tot de verantwoorde inzet van generatieve AI te komen. Deze vult het Standpunt aan en bevat voorbeelden van toepassingsgebieden waar de inzet van generatieve AI relatief lage impact heeft, zoals wetenschappelijk onderzoek. Het erkent dat generatieve AI grote kansen biedt voor het wetenschappelijk onderzoek: *"Wetenschappers kunnen bijvoorbeeld generatieve AI inzetten om onderzoek efficiënter en creatiever te maken, om groeiende hoeveelheden data te verwerken, maar hebben ook de mogelijkheid om zelf generatieve AI te bestuderen of te ontwikkelen. Omdat wetenschappelijk onderzoek specifieke doelen nastreeft en zich in een specifieke context voltrekt, zijn generieke kaders niet toereikend. Het verantwoord ontwikkelen en toepassen van generatieve AI zal in lijn moeten gebeuren met uitgangspunten van onderzoeksintegriteit en het overheidsbrede beleid. Om dat gericht te adresseren, ontwikkelen rijkskennisinstituten en planbureaus (verenigd in het RKI-netwerk) gezamenlijk een kader voor de inzet en (door-) ontwikkeling van generatieve AI binnen de contexten van het wetenschappelijk onderzoek."*
- Het Algoritmekader van de overheid<sup>72</sup>
  - De overheid wil graag effectief en verantwoord gebruik kunnen maken van algoritmes en AI. De randvoorwaarden voor verantwoorde inzet van algoritmes en AI ontwikkelt zich in hoog tempo. In relatief korte tijd komen er veel nieuwe mogelijkheden, maar ook nieuwe wetten, regels, normen, kaders en instrumenten. Met het Algoritmekader geeft het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (CIO Rijk) – samen met andere belanghebbenden – antwoorden op vragen als: hoe voldoe je aan minimale normen? Wat houdt verantwoorde inzet in? Welke wetten zijn relevant? En, welk instrument pas je wanneer toe?

## **Ethische richtsnoeren over de inzet van AI**

- Ethische richtsnoeren voor betrouwbare AI<sup>73</sup>
  - De kern van deze richtlijnen, die als fundament dienden voor de AI-verordening, is het uitgangspunt dat AI-toepassingen gedurende hun gehele levenscyclus wettig, ethisch en robuust moeten zijn. Het vertaalt ethische beginselen van respect voor de menselijke autonomie, preventie van schade, rechtvaardigheid en verantwoording naar zeven concrete principes om betrouwbare AI te verwezenlijken.

<sup>70</sup> Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (CIO Rijk), *Overheidsbreed standpunt voor de inzet van generatieve AI*, 22 april 2025, te raadplegen via: <https://www.rijksoverheid.nl/documenten/rapporten/2025/04/16/het-overheidsbrede-standpunt-voor-de-inzet-van-generatieve-ai>.

<sup>71</sup> Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (CIO Rijk), *Overheidsbrede handreiking voor de verantwoorde inzet van generatieve AI*, 22 april 2025, te raadplegen via: <https://www.rijksoverheid.nl/documenten/rapporten/2025/04/16/overheidsbrede-handreiking-generatieve-ai>.

<sup>72</sup> Het Algoritmekader is te raadplegen via <https://minbzk.github.io/Algoritmekader/>.

<sup>73</sup> European Commission High-Level Expert Group on AI, *Ethics Guidelines for Trustworthy Artificial Intelligence*, 8 april 2019, te raadplegen via: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

De zeven principes van betrouwbare AI zijn:

1. Menselijke controle en menselijk toezicht;
2. Technische robuustheid;
3. Privacy en datagovernance;
4. Transparantie;
5. Diversiteit, non-discriminatie en rechtvaardigheid;
6. Milieu- en maatschappelijke waarden, en;
7. Verantwoording.

- Living guidelines on the responsible use of generative AI in research<sup>74</sup>
  - Als onderzoekers generatieve AI op de juiste manier gebruiken, kan het een krachtige technologie bieden die wetenschappelijk werk vergemakkelijkt en versnelt. De Europese Commissie ontwikkelde samen met belanghebbenden op het gebied van onderzoek en innovatie een reeks aanbevelingen om verantwoorde integratie van generatieve AI in onderzoek te ondersteunen. De richtlijnen volgen de principes van wetenschappelijke integriteit en pakken de belangrijkste uitdagingen aan waarmee onderzoekers worden geconfronteerd bij het gebruik van generatieve AI.

Deze richtsnoeren vormden een belangrijke inspiratiebron voor RKI's bij de totstandkoming van het RKI-kader.

---

<sup>74</sup> Europese Commissie, ERA Forum Stakeholders' document, *Living guidelines on the responsible use of generative AI in research*, maart 2024, te raadplegen via: [https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc\\_en?filename=ec\\_rtd\\_ai-guidelines.pdf](https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?filename=ec_rtd_ai-guidelines.pdf).

## **Bijlage 1: Model voor risicoanalyse voor het gebruik van generatieve AI binnen wetenschappelijk onderzoek**

### **Onderdeel A: beschrijving van de concrete situatie**

Bespreek aan de hand van de volgende zeven vragen de concrete situatie waarin een overheidsorganisatie generatieve AI wil gebruiken binnen het toepassingsgebied wetenschappelijk onderzoek. Beschrijf de bevindingen bondig.

1. **Welk doel wordt gediend met de inzet van generatieve AI?**
2. **Is er sprake van een voorgenomen inzet van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek? Wordt het niet gebruikt in een andere context?**
3. **Wat zijn de kansen en mogelijkheden die de inzet van generatieve AI biedt?**
4. **Welke generatieve AI-toepassing(en) gaat gebruikt worden?**
5. **Waarom is/zijn deze generatieve AI toepassing(en) geschikt?**
6. **Waarom zijn data die als input dienen voor de generatieve AI toepassing(en) geschikt?**
7. **Welke partijen en personen zijn er bij de ontwikkeling/het gebruik/het onderhoud van het algoritme betrokken?**

### **Onderdeel B: meest relevante thema's bij waardevolle, verantwoorde en ethische inzet van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek**

Bespreek en beschrijf bondig of er risico's te voorzien zijn met betrekking tot deze acht meest relevante thema's bij waardevolle, verantwoorde en ethische inzet van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek. Wat zijn strategieën of maatregelen om eventuele risico's te verkleinen of weg te nemen?<sup>75</sup>

De thema's zijn:

1. **Menselijke controle en menselijk toezicht**
2. **Technische robuustheid**
3. **Privacy en datagovernance**
4. **Transparantie**
5. **Reproduceerbaarheid**
6. **Diversiteit, non-discriminatie en rechtvaardigheid**
7. **Milieu- en maatschappelijke waarden**
8. **Verantwoording**

---

<sup>75</sup> Bijlage 2 bij het RKI-kader adresseert per thema van Onderdeel B drie voorbeelden van strategieën of maatregelen die potentiële risico's tegengaan of verkleinen.



## Bijlage 2: Voorbeelden ter illustratie voor Onderdeel B van het model voor risicoanalyse

Het RKI-kader ziet specifiek op de inzet en (door-)ontwikkeling van generatieve AI binnen het toepassingsgebied wetenschappelijk onderzoek. RKI's identificeerden vijf situaties (doeleinden) waarin zij behoefte hebben aan duidelijkheid over de inzet en (door-)ontwikkeling van generatieve AI.<sup>76</sup> Deze bijlage beschrijft per thema van Onderdeel B van het instrument voor risicoanalyse drie voorbeelden die potentiële risico's tegengaan of verkleinen.

	A. Generatieve AI als zelfstandig studieobject	B. Wetenschappelijk onderzoek met generatieve AI voor grotere efficiëntie	C. Ondersteuning bij het gebruik van de (Nederlandse) taal	D. Verhoging efficiëntie statistici, software engineers en data scientists	E. Genereren van statistisch relevante (synthetische) datasets
<b>1. Menselijke controle en menselijk toezicht</b>	<ul style="list-style-type: none"> <li>De onderzoeker begrijpt hoe de generatieve AI werkt zodat deze proactief kan ingrijpen bij ongewenst gedrag of onbedoelde uitkomsten;</li> <li>De onderzoeker beoordeelt de uitkomsten om mogelijke fouten of risico's vroegtijdig(er) te signaleren;</li> <li>De onderzoeker blijft zelf verantwoordelijkheid voor de werking van het AI-systeem, binnen de onderzoekscontext.</li> </ul>	<ul style="list-style-type: none"> <li>Uiteindelijke beslissingen liggen bij de onderzoeker;</li> <li>De onderzoeker beoordeelt de uitkomsten op bijvoorbeeld fouten, misinterpretaties, of ongewenste patronen;</li> <li>Er is een iteratieve samenwerking tussen de onderzoeker en AI, waarbij de expertise van de onderzoeker doorslaggevend is bij het trainen en aanpassen van de AI-modellen.</li> </ul>	<ul style="list-style-type: none"> <li>De onderzoeker blijft zelf verantwoordelijk voor publicatie; deze controleert gegenereerde teksten om te waarborgen dat deze feitelijk correct, coherent en geschikt zijn;</li> <li>De onderzoeker controleert taaluitvoer op discriminatie en ongepaste inhoud; waar nodig past de onderzoeker het aan;</li> <li>De onderzoeker gebruikt generatieve AI uitdrukkelijk als hulpmiddel.</li> </ul>	<ul style="list-style-type: none"> <li>Zorg ervoor dat generatieve AI duidelijk ondersteunend is en geen autonome beslissingen neemt;</li> <li>Implementeer een proces waarin uitkomsten van generatieve AI worden gecontroleerd, gevalideerd en – indien nodig – gecorrigeerd door een menselijk expert;</li> <li>Houd op passend detailniveau bij welke interventies door mensen worden uitgevoerd.</li> </ul>	<ul style="list-style-type: none"> <li>Laat gegenereerde synthetische datasets beoordelen door experts op aspecten als validiteit, consistentie en mate van bruikbaarheid;</li> <li>Om blind vertrouwen in automatisch gegenereerde data te voorkomen, zijn momenten van passende menselijke controle geïntegreerd in het proces;</li> <li>Creëer feedbackloops om de iteratieve samenwerking onderzoeker en AI continu te verbeteren.</li> </ul>
<b>2. Technische robuustheid</b>	<ul style="list-style-type: none"> <li>Analyseer het generatieve AI-model realtime en achteraf om mogelijke kwetsbaarheden op passende wijze te identificeren en documenteren;</li> </ul>	<ul style="list-style-type: none"> <li>Zorg dat de trainings- en testdatasets zo veel mogelijk representatief, schoon en vrij zijn van ruis en/of biases;</li> </ul>	<ul style="list-style-type: none"> <li>Streef naar de inzet van een AI-model dat naar behoren is getraind met datasets die hoogwaardig, representatief en actueel zijn;</li> <li>Onderzoekers zijn zich bewust van mogelijke</li> </ul>	<ul style="list-style-type: none"> <li>Beoordeel of het AI-model is getraind met datasets die voldoende hoogwaardig, representatief en actueel zijn;</li> <li>Pas technieken toe om overfitting te minimaliseren;</li> </ul>	<ul style="list-style-type: none"> <li>De generatieve AI wordt getraind met (een) vooraf gecontroleerde en getoetste dataset(s). Deze is/zijn representatief zodat de synthetische data robuust is en geen biases of foutpatronen bevat;</li> </ul>

<sup>76</sup> Zie par. 1.1. van dit RKI-kader.

	A. Generatieve AI als zelfstandig studieobject	B. Wetenschappelijk onderzoek met generatieve AI voor grotere efficiëntie	C. Ondersteuning bij het gebruik van de (Nederlandse) taal	D. Verhoging efficiëntie statistici, software engineers en data scientists	E. Genereren van statistisch relevante (synthetische) datasets
	<ul style="list-style-type: none"> <li>Onderzoek de consistentie en betrouwbaarheid van de gegenereerde uitkomsten onder verschillende omstandigheden;</li> <li>Onderwerp het model aan tests om te evalueren hoe het systeem presteert en zich herstelt.</li> </ul>	<ul style="list-style-type: none"> <li>Pas technieken toe om overfitting te minimaliseren;<sup>77</sup></li> <li>Voer tests uit naar de technische weerbaarheid van het AI-systeem, zowel realtime als achteraf.</li> </ul>	<ul style="list-style-type: none"> <li>fouten of onverwachte uitkomsten en hanteren kwaliteitscontroles en fallbackmechanismen;</li> <li>Onderzoekers begrijpen de beperkingen en mogelijkheden van de gebruikte generatieve AI.</li> </ul>	<ul style="list-style-type: none"> <li>Voer tests uit naar de technische weerbaarheid van het AI-systeem, zowel realtime als achteraf.</li> </ul>	<ul style="list-style-type: none"> <li>Evalueer en test de werking van de generatieve AI voortdurend op passende wijze;</li> <li>Beveilig het AI-model op proportionele wijze tegen mogelijke manipulatie(s) door kwaadwillenden.</li> </ul>
<b>3. Privacy en datagovernance<sup>78</sup></b>	<ul style="list-style-type: none"> <li>Maak bij het trainen en testen van generatieve AI-modellen gebruik van zo anoniem mogelijke of synthetische datasets;</li> <li>Beperk de toegang tot datasets tot bevoegde onderzoekers;<sup>79</sup></li> <li>Houd een overzichtelijke administratie bij van welke datasets zijn gebruikt, welke bewerkingen zijn uitgevoerd, en voor welke doeleinden.</li> </ul>	<ul style="list-style-type: none"> <li>Gebruik alleen noodzakelijke gegevens, die waar mogelijk anoniem of gepseudonimiseerd zijn;</li> <li>Implementeer voldoende passende, robuuste beveiligingsmaatregelen, zoals encryptie en toegangscontrole;</li> <li>Houd een overzichtelijke administratie bij van welke data wordt gebruikt, hoe deze wordt verwerkt, en door wie.</li> </ul>	<ul style="list-style-type: none"> <li>Onderzoekers voeren zo min mogelijk persoonsgegevens, vertrouwelijke informatie of gevoelige onderzoeksgegevens in;</li> <li>Er zijn passende toegangscontroles en logsystemen om te monitoren wie het AI-systeem gebruikt;</li> <li>Houd een overzichtelijke administratie bij van welke data wordt gebruikt, hoe deze wordt verwerkt, en door wie.</li> </ul>	<ul style="list-style-type: none"> <li>Onderzoekers voeren zo min mogelijk persoonsgegevens, vertrouwelijke informatie of gevoelige onderzoeksgegevens in;</li> <li>Implementeer voldoende passende, robuuste beveiligingsmaatregelen, zoals encryptie en toegangscontrole;</li> <li>Houd een overzichtelijke administratie bij van welke data wordt gebruikt, hoe deze wordt verwerkt, en door wie.</li> </ul>	<ul style="list-style-type: none"> <li>Zorg dat synthetische data geen direct (onbedoeld) herleidbare informatie uit originele datasets bevat;</li> <li>Scheid 'echte' en synthetische data van elkaar;</li> <li>Houd een overzichtelijke administratie bij van welke synthetische data wordt gebruikt en controleer of deze juist wordt gebruikt.</li> </ul>
<b>4. Transparantie</b>	<ul style="list-style-type: none"> <li>Leg voldoende gedetailleerd vast hoe de generatieve AI-toepassing wordt geanalyseerd,</li> </ul>	<ul style="list-style-type: none"> <li>Houd voldoende gedetailleerd bij welke datasets, algoritmes, methodologieën en</li> </ul>	<ul style="list-style-type: none"> <li>Er wordt vastgelegd hoe en waarom generatieve AI wordt ingezet, welke</li> </ul>	<ul style="list-style-type: none"> <li>Er wordt vastgelegd hoe en waarom de generatieve AI wordt ingezet, welke</li> </ul>	<ul style="list-style-type: none"> <li>Documenteer hoe het AI-model is geselecteerd, getraind en gevalideerd, inclusief de gebruikte</li> </ul>

<sup>77</sup> Bij 'overfitting' is het model te nauwkeurig toegespitst en afgestemd op de trainingsgegevens, waardoor het met andere gegevens onnauwkeurige voorspellingen doet.

<sup>78</sup> Als bij de inzet van generatieve AI persoonsgegevens worden verwerkt, is de AVG van toepassing.

<sup>79</sup> Daarentegen wordt opgedane kennis over de werking van generatieve AI gedeeld (Open Science).

	A. Generatieve AI als zelfstandig studieobject	B. Wetenschappelijk onderzoek met generatieve AI voor grotere efficiëntie	C. Ondersteuning bij het gebruik van de (Nederlandse) taal	D. Verhoging efficiëntie statistici, software engineers en data scientists	E. Genereren van statistisch relevante (synthetische) datasets
	<p>getraind en geëvalueerd, inclusief gebruikte datasets, en prompts;</p> <ul style="list-style-type: none"> <li>• Publiceer de gebruikte methoden, frameworks en andere relevante informatie;</li> <li>• Wees voldoende open over eventuele beperkingen, onzekerheden en mogelijke biases van de generatieve AI-toepassing, zowel in het onderzoeksontwerp als in de resultaten.</li> </ul>	<p>parameters zijn gebruikt tijdens de ontwikkeling en training van het generatieve AI-model;</p> <ul style="list-style-type: none"> <li>• Zorg dat de uitkomsten van de generatieve AI begrijpelijk en uitlegbaar zijn;</li> <li>• Communiceer duidelijk de doeleinden en toepassingsmogelijkheden van het AI-systeem, inclusief de beperkingen en potentiële risico's.</li> </ul>	<p>functies worden gebruikt en wat deze bijdragen;</p> <ul style="list-style-type: none"> <li>• Beschrijf welke generatieve AI-tool is gekozen en waarom;</li> <li>• Als generatieve AI is gebruikt voor (bijvoorbeeld) publicaties, wordt dat duidelijk gemeld.</li> </ul>	<p>functies worden gebruikt en wat deze bijdragen;</p> <ul style="list-style-type: none"> <li>• Beschrijf welke generatieve AI-tool is gekozen en waarom;</li> <li>• Communiceer duidelijk de doeleinden en toepassingsmogelijkheden van het AI-systeem, inclusief eventuele beperkingen en potentiële risico's.</li> </ul>	<p>datasets, instellingen en gemaakte keuzes;</p> <ul style="list-style-type: none"> <li>• Zorg dat de werking van het AI-model en de wijze waarop de synthetische data wordt gegenereerd in begrijpelijke termen kan worden uitgelegd;</li> <li>• Vermeld duidelijk dat het gaat om synthetische data en beschrijf de kenmerken en beperkingen ervan. Leg uit hoe de datasets zijn samengesteld en waarvoor ze geschikt zijn.</li> </ul>
<b>5. Reproduceerbaarheid</b>	<ul style="list-style-type: none"> <li>• Gebruik versiebeheertools om wijzigingen in de datasets en programmacode te volgen;</li> <li>• Maak, waar mogelijk, de gebruikte datasets, code en AI-modellen openbaar beschikbaar, rekening houdend met wetten en regels;</li> <li>• Beschrijf niet alleen de resultaten, maar ook de procedures die zijn gevolgd om de resultaten te valideren. Zo kunnen experimenten opnieuw worden uitgevoerd. Door deze een groot aantal keer</li> </ul>	<ul style="list-style-type: none"> <li>• Leg alle stappen van het onderzoek vast, inclusief de selectie van datasets, modelarchitectuur, parameters en gebruikte AI-modellen, zodat anderen het onderzoek kunnen herhalen;</li> <li>• Maak, waar mogelijk datasets, algoritmes en getrainde modellen openbaar toegankelijk, rekening houdend met wetten en regels;</li> <li>• Implementeer versiebeheer voor datasets, code en modellen.</li> </ul>	<ul style="list-style-type: none"> <li>• Zorg voor gestructureerde opslag van de invoer (prompts) die aan de AI-tool is gegeven en de gegenereerde uitvoer;</li> <li>• Documenteer welke specifieke versie van het generatieve AI-model is gebruikt;</li> <li>• Gebruik standaard workflows en methodologieën.</li> </ul>	<ul style="list-style-type: none"> <li>• Log invoer (prompts), instellingen en door de AI gegenereerde uitkomsten, zodat elke stap in het proces achteraf kan worden gereproduceerd;</li> <li>• Leg alle stappen van het onderzoek vast, inclusief de selectie van datasets, modelarchitectuur, parameters en gebruikte AI-modellen, zodat anderen het onderzoek kunnen herhalen;</li> <li>• Documenteer welke specifieke versie van het generatieve AI-model is gebruikt.</li> </ul>	<ul style="list-style-type: none"> <li>• Beschrijf elke stap in het proces, van (dataset-) voorbereiding en modelselectie tot instellingen en methoden, zodat anderen het gehele proces kunnen herhalen;</li> <li>• Houd versies bij van de gebruikte AI-modellen, softwaretools, libraries, datasets, etc.;</li> <li>• Deel modelparameters, configuratiebestanden en scripts als onderdeel van de methodiek, zodat anderen het kunnen herhalen.</li> </ul>

	A. Generatieve AI als zelfstandig studieobject	B. Wetenschappelijk onderzoek met generatieve AI voor grotere efficiëntie	C. Ondersteuning bij het gebruik van de (Nederlandse) taal	D. Verhoging efficiëntie statistici, software engineers en data scientists	E. Genereren van statistisch relevante (synthetische) datasets
	te herhalen, zijn gemiddelden vergelijkbaar.				
<b>6. Diversiteit, non-discriminatie en rechtvaardigheid</b>	<ul style="list-style-type: none"> <li>Onderzoek het generatieve AI-model op passende wijze op biases die kunnen leiden tot discriminerende of oneerlijke uitkomsten;</li> <li>Analyseer de gegenereerde uitkomsten om te waarborgen dat deze geen schadelijke inhoud bevatten, en zorg dat deze rechtvaardig zijn;</li> <li>Betrek een diverse groep belanghebbenden en experts om verschillende perspectieven mee te nemen bij het ontwerp en de evaluatie van het model.</li> </ul>	<ul style="list-style-type: none"> <li>Zorg dat de datasets die worden gebruikt voldoende representatief zijn voor diverse groepen;</li> <li>Blijf op passende wijze monitoren of het systeem eerlijk naar behoren functioneert, vooral als datasets of toepassingsgebieden veranderen;</li> <li>Betrek een diverse groep belanghebbenden en experts gedurende het onderzoeksproces.</li> </ul>	<ul style="list-style-type: none"> <li>Controleer of de gegenereerde tekst geen disproportionele discriminatoire, stereotype of bevooroordeelde uitdrukkingen bevat;</li> <li>Analyseer of het gebruikte AI-model is getraind op een voldoende diverse en representatieve dataset die recht doet aan de variatie in de (Nederlandse) taal;</li> <li>Controleer de resultaten van de AI-tool regelmatig om te beoordelen of uitkomsten voldoende rechtvaardigheid en inclusief zijn.</li> </ul>	<ul style="list-style-type: none"> <li>Zorg dat de datasets die worden gebruikt voldoende representatief zijn voor diverse groepen;</li> <li>Blijf op passende wijze monitoren of het systeem eerlijk blijft functioneren, vooral als datasets of toepassingsgebieden veranderen;</li> <li>Betrek een diverse groep belanghebbenden en experts bij het onderzoeksproces.</li> </ul>	<ul style="list-style-type: none"> <li>Zorg ervoor dat de originele datasets voldoende representatief zijn, zodat de gegenereerde synthetische data geen specifieke groepen onder- of oververtegenwoordigd;</li> <li>Betrek een diverse groep belanghebbenden en experts bij het onderzoeksproces;</li> <li>Beoordeel op passende wijze of het gebruik van de synthetische datasets geen onaanvaardbaar oneerlijke gevolgen heeft voor bepaalde groepen.</li> </ul>
<b>7. Milieu- en maatschappelijke waarden</b>	<ul style="list-style-type: none"> <li>Selecteer en optimaliseer generatieve AI-modellen om energie-efficiënt te zijn en gebruik hardware en infrastructuur met een lage ecologische voetafdruk;</li> <li>Toets of het gebruik van generatieve AI in het onderzoek maatschappelijke meerwaarde heeft;</li> </ul>	<ul style="list-style-type: none"> <li>Selecteer en train generatieve AI-modellen met een focus op duurzaamheid;</li> <li>Evalueer vooraf de potentiële maatschappelijke gevolgen van het generatieve AI-onderzoek, zoals effecten op werkgelegenheid en sociale structuren;</li> <li>Focus op AI-toepassingen die niet alleen efficiëntie</li> </ul>	<ul style="list-style-type: none"> <li>Kies voor duurzame en efficiënte AI-modellen;</li> <li>Beperk onnodige iteraties om de ecologische voetafdruk te verkleinen;</li> <li>Kies AI-modellen van leveranciers die prioriteit geven aan verantwoordelijkheid, duurzaamheid en maatschappelijke waarden.</li> </ul>	<ul style="list-style-type: none"> <li>Kies voor duurzame en efficiënte AI-modellen;</li> <li>Beperk onnodige iteraties om de ecologische voetafdruk te verkleinen;</li> <li>Kies AI-modellen van leveranciers die prioriteit geven aan verantwoordelijkheid, duurzaamheid en maatschappelijke waarden.</li> </ul>	<ul style="list-style-type: none"> <li>Stem af met een diverse groep belanghebbenden en experts of het gebruik van synthetische datasets aansluit bij maatschappelijke waarden;</li> <li>Kies voor duurzame en efficiënte AI-modellen;</li> <li>Kies AI-modellen van leveranciers die prioriteit geven aan verantwoordelijkheid,</li> </ul>

	A. Generatieve AI als zelfstandig studieobject	B. Wetenschappelijk onderzoek met generatieve AI voor grotere efficiëntie	C. Ondersteuning bij het gebruik van de (Nederlandse) taal	D. Verhoging efficiëntie statistici, software engineers en data scientists	E. Genereren van statistisch relevante (synthetische) datasets
	<ul style="list-style-type: none"> <li>Hou bij wat de milieu-impact (zoals energie- en resourceverbruik) is van het onderzoek, zodat ecologische implicaties inzichtelijk(er) worden.</li> </ul>	bevorderen, maar ook maatschappelijke problemen aanpakken.			duurzaamheid en maatschappelijke waarden.
<b>8. Verantwoording</b>	<ul style="list-style-type: none"> <li>Documenteer stappen van het onderzoeksproces, inclusief keuzes in modelselectie, trainingsparameters, datasets en interpretatie van resultaten, zodat deze achteraf inzichtelijk en te verantwoorden zijn;</li> <li>Publiceer onderzoeksresultaten en methodologieën in begrijpelijke taal en maak deze toegankelijk voor geïnteresseerden, zodat het onderzoek transparant en controleerbaar is;</li> <li>Denk na over mechanismen voor verantwoording zodat anderen het onderzoek kunnen evalueren op aspecten zoals technische validiteit en impact.</li> </ul>	<ul style="list-style-type: none"> <li>Stel vast wie binnen het onderzoeksteam verantwoordelijk is voor verschillende aspecten van het generatieve AI-proces;</li> <li>Documenteer stappen van het onderzoek, inclusief keuzes en aannames;</li> <li>Laat het onderzoek en de toepassing van AI controleren door een onafhankelijke derde om de kwaliteit en maatschappelijke verantwoordelijkheid te waarborgen.</li> </ul>	<ul style="list-style-type: none"> <li>Houd bij hoe en waarom de generatieve AI wordt ingezet, inclusief de specifieke rol bij taalondersteuning, om verantwoording af te kunnen leggen over de toepassing;</li> <li>Zorg dat er een duidelijk proces is waarin menselijke controle en beoordeling worden toegepast op de uitkomsten;</li> <li>Erken en corrigeer eventuele fouten of tekortkomingen (menselijke controle).</li> </ul>	<ul style="list-style-type: none"> <li>Documenteer stappen van het onderzoek, inclusief keuzes in modelselectie, datasetgebruik, trainingsparameters en interpretatie van resultaten, zodat deze achteraf inzichtelijk en te verantwoorden zijn;</li> <li>Maak de resultaten toegankelijk zodat het onderzoek transparant en controleerbaar is;</li> <li>Zorg dat er een duidelijk proces is waarin menselijke controle en beoordeling worden toegepast op de uitkomsten;</li> </ul>	<ul style="list-style-type: none"> <li>Stel vast wie binnen het onderzoeksteam verantwoordelijk is voor verschillende aspecten van het generatieve AI-proces;</li> <li>Zorg voor voldoende gedetailleerde informatie waarin staat hoe de synthetische datasets zijn gegenereerd, de gebruikte modelinstellingen en wat de originele datasets zijn;</li> <li>Zorg dat er een duidelijk proces is waarin menselijke controle en beoordeling worden toegepast op de uitkomsten.</li> </ul>