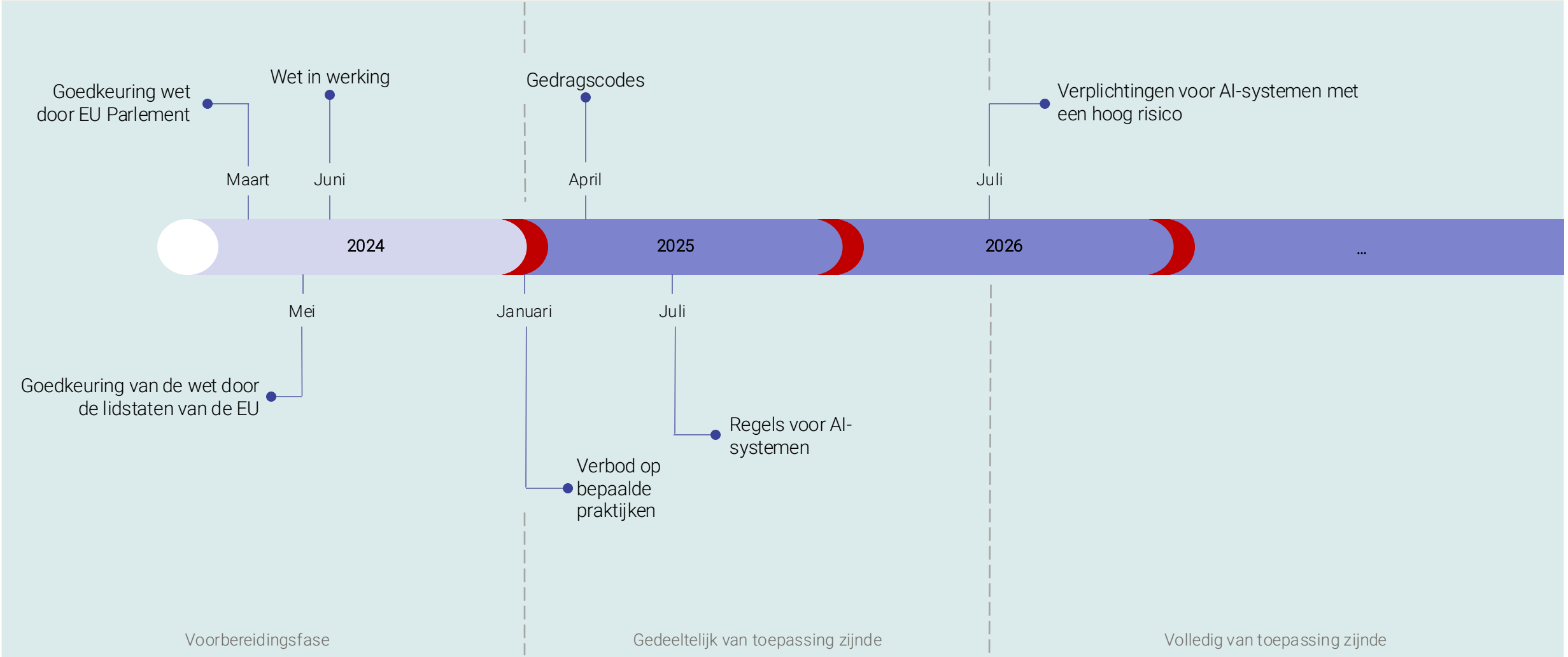
Solventa

# EU AI Act

TOELICHTING OP DE WET

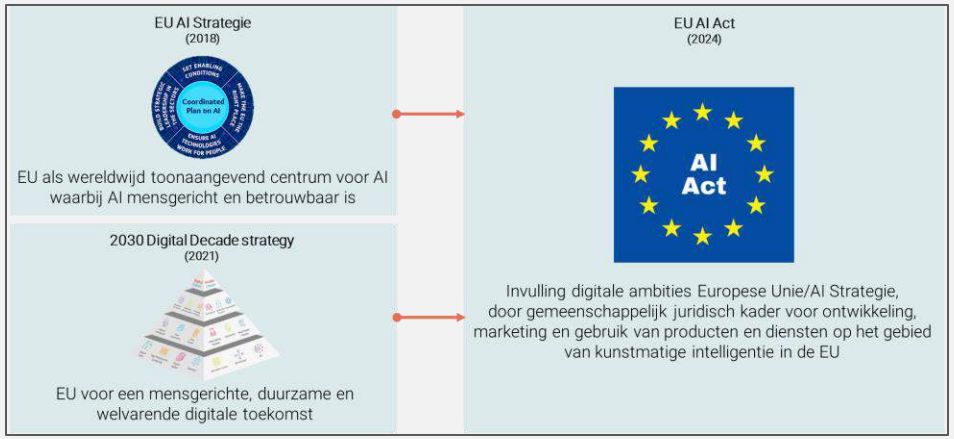
# Timeline goedkeuring en invoering EU AI Act




Wat betekent wat? *Inwerkingtreding* betekent dat de wet officieel van kracht en juridisch bindend is. *Van toepassing zijn* betekent dat de wet daadwerkelijk wordt toegepast in de praktijk.

# Context en doel

*Waar gaat het over?*

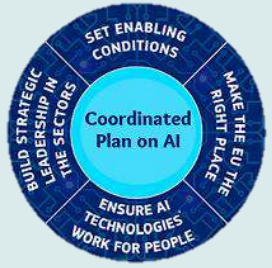



Context: waar komt het vandaan


Doel: wat willen we ermee bereiken

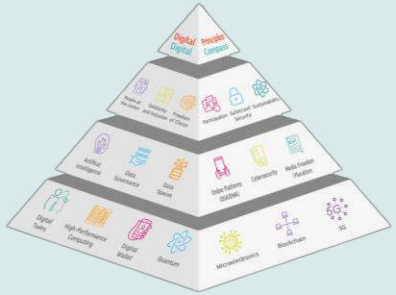
# Context

EU AI Strategie
(2018)




EU als wereldwijd toonaangevend centrum voor AI waarbij AI mensgericht en betrouwbaar is

2030 Digital Decade strategy
(2021)



EU voor een mensgerichte, duurzame en welvarende digitale toekomst

EU AI Act
(2024)


Invulling digitale ambities Europese Unie/AI Strategie, door gemeenschappelijk juridisch kader voor ontwikkeling, marketing en gebruik van producten en diensten op het gebied van kunstmatige intelligentie in de EU

# Doel

“Maak van Europa een wereldleider in de ontwikkeling, en in het gebruik,  van veilige, betrouwbare en mensgerichte kunstmatige intelligentie..

**Bescherming van grondrechten**
Waarborging van grondrechten van EU-burgers bij de inzet van AI, zodat AI-systemen geen inbreuk doen op privacy, gelijke behandeling en menselijke waardigheid

**Transparantie en verantwoording**
Afdwingen van transparantie bij bedrijven over waar & hoe AI ingezet is, en hoe menselijk toezicht en verantwoording geborgd is

**Veiligheid en betrouwbaarheid**
Bevordering veiligheid en betrouwbaarheid AI door risico-gebaseerde eisen te stellen aan AI-systemen

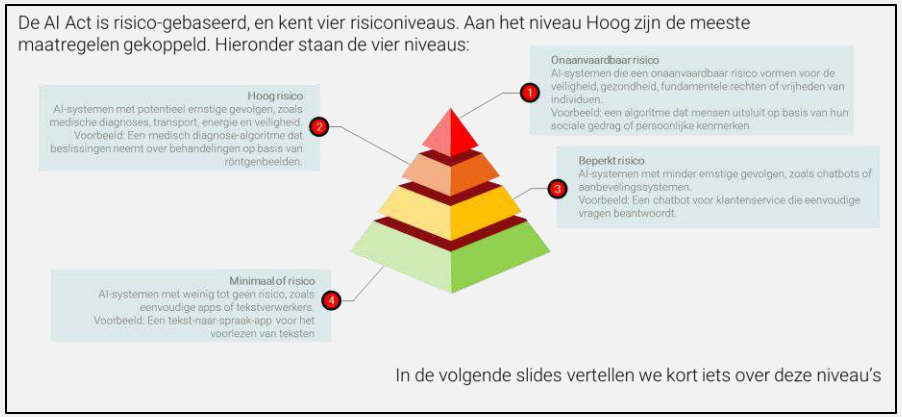
**Innovatie en concurrentievermogen**
Stimuleren van innovatie op het gebied van AI door een gelijk speelveld te creëren voor bedrijven

...door het definiëren van een gemeenschappelijk juridisch kader voor de ontwikkeling, marketing en gebruik van producten en diensten op het gebied van kunstmatige intelligentie in de EU”

# Essentie

*Wat zegt het?*


Waar gaat het over: toepassingsgebied van de wet en definitie van AI




Risico-gebaseerde aanpak

# Waar gaat het over?

Wanneer van toepassing
- Iedereen die AI-systemen maakt, gebruikt, importeert of distribueert in de EU (ook als de systemen buiten de EU zijn gemaakt)

Wanneer niet van toepassing:
- Militaire, defensie- of nationale veiligheidsdoeleinden
- Gebruik door buitenlandse overheden of internationale organisaties voor rechtshandhaving en justitiële samenwerking, (zolang deze de rechten van individuen beschermt).
- Wetenschappelijk onderzoek en ontwikkeling
- Nog niet op de markt uitgebracht
- Gebruik voor persoonlijke, niet-professionele activiteiten
- Gratis/open source-licenties (tenzij deze hoog risico-dragend zijn)

Definitie van AI
"Software die is ontwikkeld aan de hand van een of meer van de volgende technieken en benaderingen: Machine learning, logica- en kennisgebaseerde systemen, statistische benaderingen en Bayesiaanse methoden. Waarbij geldt dat deze software voor een bepaalde reeks door mensen gedefinieerde doelstellingen output kan genereren, zoals inhoud, voorspellingen, aanbevelingen of beslissingen die van invloed zijn op de omgeving waarmee wordt geïnterageerd". Simpel gezegd: voorspellende software

# Voor wie is het belangrijk?

De EU AI Act is van toepassing op *aanbieders* en *gebruikers* van AI systemen. Deze partijen hoeven niet in de EU gevestigd te zijn: als de *output* van een AI-systeem wordt toegepast in de EU, dan is de wet van toepassing.

### Aanbieders van AI-systemen

- De partij (natuurlijk persoon of rechtspersoon) die een AI systeem binnen de EU op de markt brengt of als service aanbiedt (ongeacht of deze partij in gevestigd in of buiten de EU)
- De EU AI Act stelt aan deze partij de zwaarste eisen. Dit beschrijven we verder in deze presentatie

### Gebruikers van AI-systemen

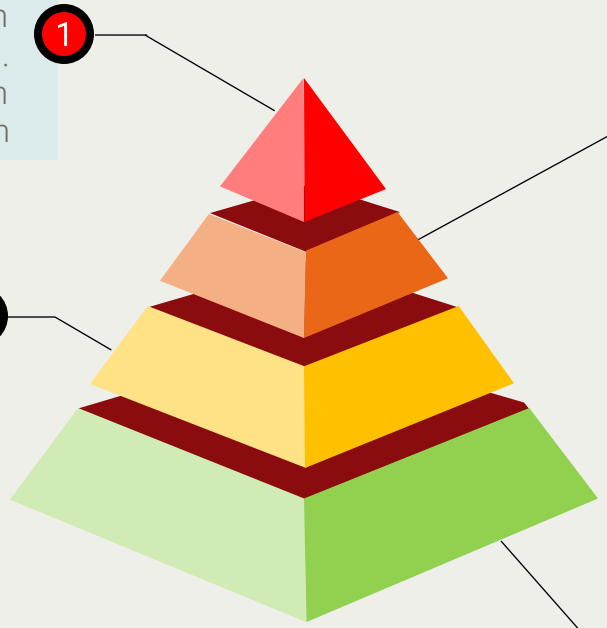
- Gebruiker ("deployer"): partij (natuurlijk persoon of rechtspersoon) die een AI systeem *toepast*. Het gaat hierbij dus niet om de eindgebruikende partij (bijvoorbeeld: jij en ik) maar om de partij die in zijn dienstverlening AI-systemen toepast (bijvoorbeeld je energieleverancier die AI gebruikt voor het optimaliseren van de aansturing van je warmtepomp).
- De EU AI Act stelt ook eisen aan gebruikers van een hoog-risico AI-systeem. Ook dit is verderop in deze presentatie beschreven


# Risicogebaseerd

De AI Act is risico-gebaseerd, en kent vier risiconiveaus. Aan het niveau Hoog zijn de meeste maatregelen gekoppeld. Hieronder staan de vier niveaus:

**Onaanvaardbaar risico**
AI-systemen die een onaanvaardbaar risico vormen voor de veiligheid, gezondheid, fundamentele rechten of vrijheden van individuen.
Voorbeeld: een algoritme dat mensen uitsluit op basis van hun sociale gedrag of persoonlijke kenmerken

**Beperkt risico**
AI-systemen met minder ernstige gevolgen, zoals chatbots of aanbevelingssystemen.
Voorbeeld: Een chatbot voor klantenservice die eenvoudige vragen beantwoordt.

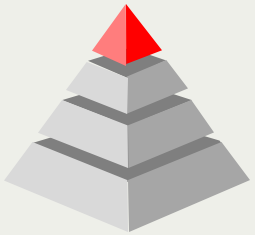



**Hoog risico**
AI-systemen met potentieel ernstige gevolgen, zoals medische diagnoses, transport, energie en veiligheid.
Voorbeeld: Een medisch diagnose-algoritme dat beslissingen neemt over behandelingen op basis van röntgenbeelden.

**Minimaal of geen risico**
AI-systemen met weinig tot geen risico, zoals eenvoudige apps of tekstverwerkers.
Voorbeeld: Een tekst-naar-spraak-app voor het voorlezen van teksten

In de volgende slides vertellen we kort iets over deze niveaus

# Onaanvaardbaar risico



AI-toepassing is verboden bij systemen die:

- beslissingen van mensen manipuleren of hun kwetsbaarheden uitbuiten.
- mensen evalueren of classificeren op basis van hun sociale gedrag, persoonlijke eigenschappen of biometrische gegevens
- het risico van een persoon om een misdrijf te plegen voorspellen op basis van profilering of persoonlijkheidskenmerken
- Datasets aanleggen voor gezichtsherkenning door gezichtsbeelden van internet of camerabeelden te verzamelen.
- Emoties op de werkplek of in onderwijsinstellingen afleiden.
- Real-time biometrische identificatie op afstand in openbare ruimtes

Systemen die AI gebruiken om een van bovenstaande te kunnen, mogen dus niet in de handel gebracht worden, in gebruik gesteld worden, of gebruikt worden binnen de EU

# Hoog risico

Wanneer valt een AI-systeem in de categorie Hoog risico?

- Het AI-systeem is de beveiligingscomponent van een bepaalde categorie producten waar volgens de EU richtlijnen een zogeheten **conformiteitsverklaring** nodig is. En dat zijn zo'n beetje alle producten die iets doen of waar je iets mee kan doen, en wat schade oplevert als dat misgaat. Staafmixers, grasmaaiers en hartimplantaten, om er maar een paar te noemen
- Het AI-systeem is zelf onderdeel van zo'n product

Of als gebruikt in een van de volgende use cases:

- Toegestane remote biometrische identificatie (zie ook risicoclassificatie Onaanvaardbaar risico)
- Beheer en exploitatie kritieke infrastructuur (wegverkeer, levering van water, gas, verwarming, en elektriciteit)
- Onderwijs en beroepsopleiding
- Werkgelegenheid, personeelsbeheer en toegang tot zelfstandige arbeid
- Toegang tot en gebruik van essentiële particuliere diensten en openbare diensten en uitkeringen
- Rechtshandhaving
- Migratie, asiel en grenscontroles
- Rechtspraak en democratische processen

# Hoog risico

Waar moeten AI-systemen in deze categorieën aan voldoen?

- Risicomanagement: er moet een systeem voor risicobeheer vastgesteld, uitgevoerd, gedocumenteerd en in stand gehouden zijn
- Data governance: trainings-, validatie- en testdata zijn relevant, representatief en zo veel als mogelijkl foutenvrij en compleet in relatie tot het doel waarvoor het systeem de data gebruikt
- Technische documentatie: er is beschreven op welke manier het systeem in overeenstemming is met de wet, en geeft alle informatie die nodig is voor de autoriteit om dit te kunnen beoordelen
- Registratie: AI-systemen loggen hun activiteiten zodat de werking van het systeem traceerbaar is
- Transparantie en informatieverstrekking aan gebruikers: er zijn gebruiksinstructies voor de partijen die het AI-systeem implementeren/toepassen over hoe zij kunnen voldoen aan de aan hun gestelde wettelijke eisen
- Menselijk toezicht: het AI-systeem biedt middelen en mogelijkheden voor menselijk toezicht bij gebruik van het systeem
- Nauwkeurigheid, robuustheid en cyberbeveiliging: Het AI-systeem is ontworpen en gebouwd om te voldoen aan “een passend niveau” van deze vermelde kwaliteitskenmerken
- Aanvullend eisen met betrekking tot centrale registratie en post-market monitoring. Deze staan op de volgende slides

# Hoog risico

Waar moeten aanbieders van AI-systeem in deze categorieën aan voldoen?

a. AI-systemen moeten aan specifieke normen voldoen (zie voorgaande slide)
b. Contactgegevens moeten beschikbaar zijn.
c. Moet een kwaliteitsmanagementsysteem hebben.
d. Moet bepaalde documenten en logboeken bijhouden.
f. AI-systeem moet worden gecontroleerd op naleving van de regelgeving (conformiteit) voordat het wordt verkocht of gebruikt. *Dit is een hele belangrijke eis, want dit triggert het hele governance-proces waarbij allerlei instanties gaan controleren of het AI-systeem voldoet aan de wettelijke eisen, zorgt voor registratie in de EU-database etc. Dit leggen we in een volgende slide verder uit*
g. Product moet worden voorzien van een CE-markering om aan te tonen dat het voldoet aan de EU-normen.
i. Product moet worden geregistreerd (zie volgende slide)
j. Eventuele problemen die bij gebruik gemeld zijn moeten worden opgelost en bij informatieverzoeken moet de nodige informatie verstrekt worden
k. Moet bewijzen dat het product aan de normen voldoet als de autoriteiten daarom vragen.
l. Het AI-systeem moet toegankelijk zijn volgens EU-richtlijnen

Het begrip “Aanbieder” moet je een beetje ruim zien: de regels gelden voor de producten van het systeem zelf, de importeur, en de distributeur ervan. Elk van deze eis is in de wet verder uitgelegd en toegelicht. Het voert een beetje ver om dit hier allemaal op te nemen. De details zijn te vinden in titel III – AI-systemen met een hoog risico, hoofdstuk 3 - Verplichtingen van aanbieders en gebruikers van ai-systemen met een hoog risico

# Hoog risico

Waar moeten gebruikers van AI-systeem in deze categorieën aan voldoen?

a. De juiste maatregelen treffen om het systeem te gebruiken zoals bedoeld door de aanbieder
b. Menselijk toezicht houden op gebruik van het systeem
c. Borgen van het gebruik van de juiste data
d. De operatie van het systeem monitoren en bij het vermoeden van risicovol gedrag van het systeem de aanbieder hierover informeren
e. Logfiles bewaren voor een nog vast te stellen periode
f. Uitvoeren van een DPIA (data protection impact assessment) voor het systeem gebruikt wordt

Gebruiker in deze context is dus de organisatie die AI inzet (jouw energieleverancier bijvoorbeeld), en niet de eindgebruiker (jij & ik)

# Hoog risico

Aanvullende aspecten

Voor AI-systemen met een hoog risico gelden er nog een aantal aanvullende aspecten. De meest belangrijke behandelen we hier

## Centrale registratie

- Er komt een centrale database voor hoog-risico AI systemen (en systemen die wel voldoen aan de criteria maar een lage impact hebben en daardoor niet als hoog-risico zijn gelabeld)
- De aanbieder van zo'n systeem moet zichzelf en het systeem hierin registreren voordat hij het op de markt brengt. Dit is onderdeel van het conformiteitsproces
- Overheidsinstanties die een hoog-risico AI systeem gaan gebruiken, moeten dit ook in de centrale database registreren. Deze data is publiek inzichtelijk, met enkele uitzonderingen (bijvoorbeeld rechtshandhaving en migratie)

## Post-market monitoring

- Aanbieders van AI-systemen met een hoog risico moeten een monitoringsysteem opzetten als het systeem op de markt komt. Het doel hiervan is om te borgen dat de aanbieder aan de regelgeving blijft voldoen
- Het monitoringssysteem moet gegevens verzamelen en analyseren over de prestaties van het systeem zolang het in bedrijf is
- Het monitoringsysteem moet gebaseerd zijn op een plan dat deel uitmaakt van de technische documentatie
- De Europese Commissie zal een model voor dit plan ter beschikking stellen. Dit is er dus nog niet!

Als er al een monitoringsysteem bestaat op grond van andere wetgeving dat voldoet aan de hier gestelde eisen, dan kan de aanbieders dat systeem inzetten hiervoor

## Markttoezicht

- AI-systemen zijn onderworpen aan dezelfde regelgeving als andere economische producten
- Markttoezichtautoriteiten moeten jaarlijks verslag uitbrengen aan de Commissie en de nationale mededingingsautoriteiten over eventuele problemen of verboden praktijken
- AI-systemen met een hoog risico zijn onderworpen aan extra toezicht, vooral die welke worden gebruikt door financiële instellingen of voor wetshandhavingsdoeleinden
- De autoriteiten kunnen gezamenlijke activiteiten voorstellen om de naleving te bevorderen en niet-naleving te signaleren
- Onder bepaalde voorwaarden kunnen de autoriteiten toegang krijgen tot de broncode van risicovolle AI-systemen

# General purpose AI

En hoe zit het met “general purpose AI” (GPAI) zoals ChatGPT?

- De wet herkent het concept “general purpose AI”. De definitie hiervan is een beetje vaag, maar het er wordt hierbij bedoeld: AI-modellen die algemeen van aard zijn en een brede toepassingscontext hebben. ChatGPT dus, bijvoorbeeld
- Daarbovenop heeft de wet het over “general purpose AI met systeemrisico”. De bepaling of een model systeemrisico heeft en hoe het proces voor deze bepaling loopt, is in algemene bewoordingen beschreven. Het is duidelijk dat de EU hier nog mee worstelt. De EU Commissie houdt zich dan ook het recht voor om allerlei uitwerkingen (annexen genaamd) nog aan te passen
- Hoe dan ook, als eenmaal, via een vrij ingewikkeld proces met allerlei ja-maar-wat-als-constructies, is vastgesteld dat een “general purpose AI” model systeemrisico heeft, dan gelden er de volgende aanvullende eisen voor het bedrijf dat dit model levert:
    - Bijhouden gedetailleerde gegevens bijhouden van de ontwikkeling en het testen van hun model
    - Informatie verstrekken aan andere bedrijven die hun AI willen gebruiken (waarbij hun intellectuele eigendom niet vrijgegeven hoeft te worden) (Bovenstaande regels gelden overigens weer niet voor AI-modellen die open source zijn en gratis beschikbaar zijn voor het publiek, tenzij ze systeemrisico’s met zich meebrengen)
    - Waar nodig: samenwerken met de Europese Commissie en de nationale autoriteiten.
    - Naast eisen m.b.t. technische documentatie en informatieverstrekking naar toepassende partijen zijn er eisen aangaande de Copyright Directive en de voor de training gebruikte data

Er is sprake van een zogeheten “code of conduct”. Als een bedrijf zich aan zo’n (overigens nu nog niet bestaande) gedragsregel houdt, dan ziet de EU Commissie dit als compliant-zijn-aan-de-wettelijke-eisen

# Overige zaken

## Aanvullende eisen

De EU AI Act stelt nog een paar aanvullende eisen aan AI-systemen met beperkt risico:

- Als gebruikers met een AI-toepassing interacteren dan moet dat direct duidelijk zijn voor hen
- Als een AI-system synthetische inhoud weergeeft (dus: inhoud dit door het AI-system zelf is gegenereerd zoals muziek, video, of teksten) dan moet het system dit duidelijk aangeven aan de gebruikers
- Bij gebruik van emotieherkenning of biometrische categorisering moet het systeem de gebruikers informeren dat dit zo is, hoe het werkt, en hoe hun gegevens veilig gehouden worden (en blijven)
- Als AI-systeem deepfake content maakt en toont, dan moet het systeem dit duidelijk aangegeven bij de gebruikers

Deze eisen zijn van toepassing als mensen voor het eerst de AI-toepassing gebruiken, en de vereiste informatie moet makkelijk vindbaar zijn. Voor het opsporen van misdaden, en voor uitingen die duidelijk als kunstzinnige expressie worden getoond, gelden deze eisen niet.

Voor AI-systemen met een minimaal of geen risico stelt de EU AI Act geen verdere eisen

## AI Office

Een term die in de nieuwste versie van de EU AI Act opduikt is de AI Office. Dit bureau is in begin 2024 opgericht door de EU Commissie en heeft de volgende doelen:

- Kennis & kunde van de Unie op het gebied van AI ontwikkelen en bijdragen aan de uitvoering van de EU AI Act
- Coördinatie voeren over de invoer, monitoring en supervisie van AI binnen de EU

## Code of Conduct

Een manier om de doelstellingen van de EU AI Act te bereiken is door gedragscodes (codes of conduct):

- De AI Office stimuleert en faciliteert het opstellen van deze codes
- Deze codes bevorderen het vrijwillig toeppassen van standaarden en best practices
- Daarnaast stimuleren deze codes om AI te gebruiken op een manier die goed is voor het milieu, AI-geletterdheid bevorderd, bijdraagt aan diversiteit, en negatieve effecten op kwetsbare groepen voorkomt
- De codes moeten bij voorkeur ontstaan vanuit het werkveld (AI-bedrijven, of daaraan gerelateerd), en houden rekening met de belangen van kleine bedrijven en startups

# Governance

## Governance

Toezicht op de wet per land loopt via de zogeheten aangemelde instanties. Governance is ingericht over een aantal assen en wordt een landelijke verplichting waarbij samenhang en samenwerking op EU-niveau geborgd is

### Governance instanties

- Er komt een EU-breed AI-office met een AI-board om op EU-niveau kennis & kunde van AI op te bouwen en toe te passen. Elk EU-land heeft zitting in de board
- Op landelijk niveau worden er de zogeheten Aangemelde instanties aangewezen die toezoen op naleving van de wet (o.a. door de confomiteitsbeoordeling, zie volgende punt)

### Conformiteitsbeoordeling

- Systemen die geclassificeerd zijn als hoog-risico moeten door de Aangemelde instantie beoordeeld worden: de conformiteitsbeoordeing
- Na positieve beoordeling volgt administrateie in de centrale database voor hoog-risico AI systemen

### Toezicht en boetes

- De EU lidstaten zijn zelf verplicht toezicht te houden op de naleving van EU AI Act
- Naar alle waarschijnlijkheid zal de Autoriteit Persoonsgegevens worden aangewezen als de toezichthouder hiervoor in Nederland
- De toezichthouder kan een boete van maximaal 35 miljoen euro per overtreding opleggen of een boete van 7% van de wereldwijde omzet van het concern waartoe de overtredende onderneming behoort. Voor kleine administratieve overtredingen geldt een maximum 7,5 miljoen euro of 1,5% van de wereldwijde omzet
- Voor General Purpose AI zijn de boetes 15 miljoen of 3% van de omzet
- Ook EU-instellingen zélf kunnen beboet worden bij geconstateerde overtredingen. De Europese Toezichthouder voor Gegevensbescherming heeft hier mandaat voor

*Titel III AI-systemen met een hoog risico, Hoofdstuk 4 en 5, Titel VI Governance, Titel XII Vertrouwelijkheid en sancties*

# En nu?

## Voorbereiding op de EU AI Act

Nu duidelijk is wat de AU AI Act inhoudt en voor wie deze van toepassing is, kijken we naar hoe je je organisatie voor kan bereiden op deze nieuwe wet. De meeste organisaties vallen in de categorie Gebruikers. We beschrijven de voorbereiding dan ook vanuit dit perspectief

### Eerste fase: initiatie

Deze fase is bedoeld om de huidige toepassing van AI binnen de organisatie compliant te krijgen met de EU AI Act. Deze fase is eenmalig, en kent de volgende stappen:

- Inventariseer de gebruikte AI-systemen en classificeer deze naar Niet-toegestaan, Hoog-risico, Beperkt of Minimaal risico.
- Voor zover nog niet gedaan: ken aan elk AI-systeem een eigenaar toe. Dit moet een eigenaar in de gebruikersorganisatie zelf zijn.
- Elimineer de systemen die gecategoriseerd zijn als niet-toegestaan
- Richt de maatregelen in voor systemen die gecategoriseerd zijn als hoog-risico (zie slides “Voor wie is het belangrijk?” en “Hoog risico”)
- Richt de maatregelen in voor de systemen die vallen in de categorie Beperkt of Minimaal risico (zie slide “Overige zaken”)
    - Van vitaal belang is het hebben en houden van de juiste data als input voor AI-verwerking. Pas de data governance processen en –verantwoordelijkheden hierop aan, evenals de datakwaliteits-processen!
- Stel een AI-beleid op om toepassing van AI binnen de organisatie af te kaderen. Alhoewel dit geen wettelijke eis is, vereenvoudigt het hebben van een AI-beleid het maken van keuzes op dit gebied. Zorg voor eigenaarschap van dit beleid op C-level, en zorg dat de organisatie snapt en begrijpt wat er in het beleid staat en waarom het belangrijk is (“AI hygiëne”)

### Tweede fase: consolidatie

Deze fase is bedoeld om het gebruik van AI in de organisatie compliant te houden met de AU AI Act. Deze fase is een continuüm, en kent de volgende stappen:

- Toets bij elke risico-impact bepaling op (nieuwe of wijzigende) inzet van AI-systemen en bepaal in de impact de categorie van deze AI-toepassing. Neem de te treffen maatregelen op als uitkomst van de impactbepaling (zie slides “Voor wie is het belangrijk?” voor de eisen die aan gebruikers gesteld worden)
- Laat de verantwoordelijke eigenaar van AI-systemen geregeld de categorisering van AI-systemen actualiseren. Een andere toepassing kan namelijk leiden tot een andere risico-classificatie!
- Toets geregeld de toepassing van het AI-beleid en stel dit, desgewenst, bij. Blijf ook werken aan communicatie over en bewustwording van de AI hygiëne
    - Speciale aandacht blijft vereist voor het controleren van de kwaliteit van de data die als input wordt gebruikt. Blijf datakwaliteit controleren en zorg voor een goed werkende PDCA-cyclus voor de betreffende data