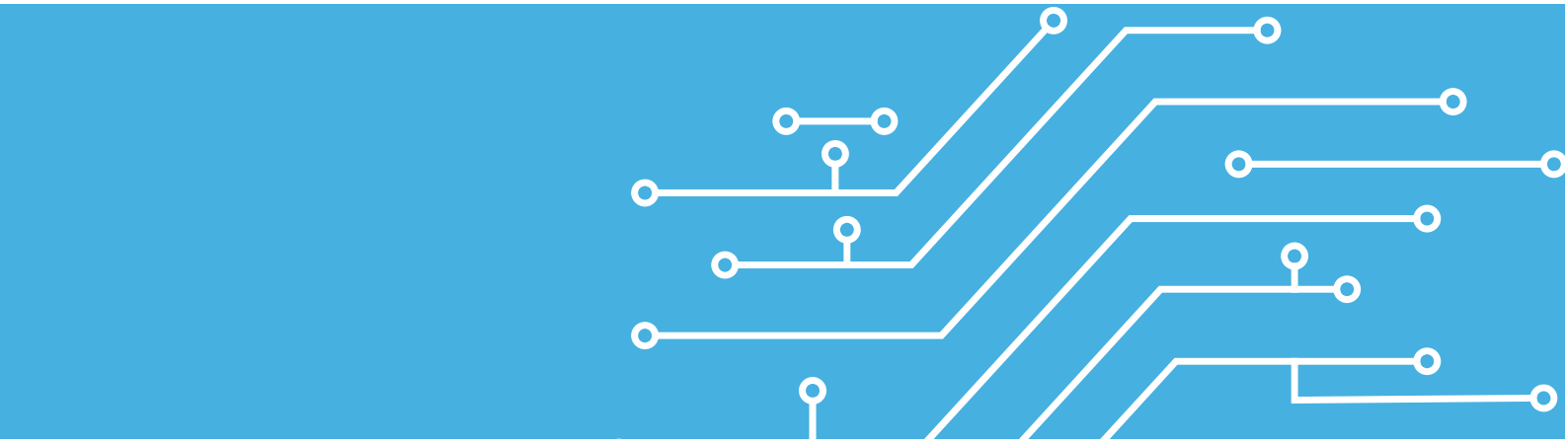


QUNICS

Deep Dive – EU AI ACT

Versie: 0.1

1. Waarom deze wet er is (en wat het wél en niet is)

De EU AI Act is opgezet als **product- en marktwetgeving**: het doel is dat AI die in de EU wordt aangeboden of gebruikt **veilig**, **betrouwbaar** en **in lijn met fundamentele rechten** is. De wet is dus geen "ethische richtlijn", maar een juridisch afdwingbaar kader met:

- **verboden toepassingen**
- **zware verplichtingen voor hoog-risico AI**
- **transparantie-eisen** voor bepaalde interacties en synthetische content
- **apart regime** voor general-purpose AI (GPAI) modellen

Belangrijk: de AI Act is **risicogebaseerd en use-case-gedreven**. Het gaat minder om "welke techniek" en meer om "waarvoor gebruik je het".

---

2. Reikwijdte: wie valt eronder?

De AI Act heeft een **extraterritoriale werking**: ook partijen buiten de EU kunnen eronder vallen als hun AI-output binnen de EU wordt gebruikt.

Kernrollen in de keten

De wet definieert rollen die je verplichtingen bepalen:

- **Provider (aanbieder)**: ontwikkelt een AI-systeem of GPAI-model en zet het op de markt of in gebruik onder eigen naam/merk.
- **Deployer (gebruiker/exploitant)**: gebruikt het AI-systeem onder eigen verantwoordelijkheid.
- **Importer**: brengt een systeem van buiten de EU de EU-markt op.
- **Distributor**: levert door binnen de keten.

"Provider by behavior" (belangrijke valkuil)

Je kunt *provider* worden door gedrag, bijvoorbeeld als je:

- je **eigen merk** op een hoog-risico systeem zet,
- **substantiële wijzigingen** doorvoert,
- de **beoogde bestemming** wijzigt waardoor het systeem in een hogere risicoklasse valt.

In praktijk betekent dit: integrators, implementatiepartners en grote afnemers kunnen ineens "aanbieder" worden met zwaardere plichten.

---

3. Definitie van "AI-systeem": breed, maar niet alles

De AI Act hanteert een brede definitie: een AI-systeem is een machinegebaseerd systeem dat met enige autonomie outputs genereert zoals voorspellingen, aanbevelingen, content of beslissingen die invloed kunnen hebben op fysieke of digitale omgevingen.

"Gewone" software met puur vaste regels valt er vaak niet onder, maar in twijfelgevallen is het verstandig om je classificatie te documenteren (zodat je later kunt aantonen waarom je iets wel/niet als AI beschouwde).

---

4. Het risicomodel: vier lagen

De AI Act hanteert vier hoofdcategorieën (niet altijd strikt exclusief, omdat transparantieverplichtingen óók bovenop andere eisen kunnen gelden):

4.1 Verboden AI (unacceptable risk)

Dit zijn toepassingen die de EU onverenigbaar acht met fundamentele rechten of maatschappelijke waarden. Voorbeelden in de wet (hoog-over):

- manipulatie die mensen wezenlijk schaadt of autonomie ondermijnt
- sociale scoring
- bepaalde vormen van biometrische categorisatie (bijv. afleiden van gevoelige kenmerken)
- grootschalige of ongerichte biometrische dataverzameling (zoals scraping voor gezichtsdatabases)
- emotieherkenning op werk/onderwijs (met beperkte uitzonderingen)

Praktisch gevolg: je hebt een "stoplicht" nodig in governance: als use-case in de verboden zone zit, houdt het op—ook als de techniek werkt.

4.2 Hoog-risico AI (high-risk)

Dit is het zwaartepunt van de AI Act: hier zitten de meeste verplichtingen.

Er zijn grofweg twee routes naar "hoog-risico":

A) Annex III: hoog-risico use-cases
Denk aan domeinen zoals:

- biometrie (in specifieke gevallen)
- kritieke infrastructuur
- onderwijs en examinering
- werk & HR (selectie, beoordeling, ontslag, taaktoewijzing)
- essentiële diensten (kredietwaardigheid, verzekeringsrisico's, noodoproepen-triage)
- law enforcement
- migratie/asiel/border control
- rechtspraak en democratische processen

B) Annex I: AI als safety component in gereguleerde producten
Bijv. medische hulpmiddelen, voertuigen, machines/robots, luchtvaart/rail, liften, PPE, radio equipment, etc. (koppeling met bestaande productwetgeving).

Opt-out uitzondering (belangrijk):
Sommige systemen in een hoog-risico domein kunnen onder voorwaarden worden uitgesloten als ze géén significante risico's creëren (maar profiling blijft doorgaans hoog-risico).

4.3 Transparantie-risico (limited risk / transparency)

Voorbeelden:

- chatbots: gebruiker moet weten dat hij met AI praat
- deepfakes/synthetische media: labeling of herkenbaarheid
- bepaalde AI-gegenereerde outputs: duidelijke informatieplicht

Deze plichten kunnen **bovenop** hoog-risico eisen komen.

4.4 Laag/overig risico

Voor veel "normale" bedrijfs-AI (zoals spamfilters, inventory forecasting, predictive maintenance) geldt geen specifieke AI Act-verplichting—maar wél vaak andere wetgeving (AVG, cybersecurity, consumentenrecht, productveiligheid).

---

5. Hoog-risico verplichtingen: wat moet je echt regelen?

Voor hoog-risico AI gelden verplichtingen over de hele lifecycle: ontwerp → training → validatie → uitrol → monitoring.

5.1 Risk Management System (RMS)

Je moet een formeel systeem hebben dat risico's identificeert, evalueert, mitigeert en monitort, inclusief:

- **safety-by-design** (risico's voorkomen in ontwerp)
- **mitigatie & control measures**
- **informatie voor veilig gebruik** (instructies, beperkingen, training)

5.2 Data governance & datakwaliteit

Eisen draaien om:

- representativiteit, relevantie en kwaliteit van datasets
- bias-risico's en onbedoelde effecten
- traceerbaarheid van data (herkomst, bewerkingen, splits)
- passende maatregelen voor privacy en security

In praktijk betekent dit: je moet *dataset discipline* hebben zoals in gereguleerde software/medtech: datalijnen, datadocumentatie, testdekkingen, drift monitoring.

5.3 Technische documentatie

Je moet kunnen aantonen:

- doel en beoogde use-case
- systeemarchitectuur en modelkeuzes
- trainings- en testmethoden
- performance metrics, limieten, robuustheid
- risicobeoordeling en mitigaties
- human oversight ontwerp

5.4 Logging & traceerbaarheid

Voor bepaalde hoog-risico systemen moeten logs automatisch of controleerbaar worden vastgelegd zodat je achteraf kunt reconstrueren:

- input/outputs
- relevante systeemstatus
- incidenten en afwijkingen

5.5 Transparantie en informatie aan deployers

Providers moeten deployers voldoende informatie geven om het systeem veilig te gebruiken, inclusief beperkingen en juiste context.

5.6 Human oversight (menselijke controle)

Je moet aantonen dat:

- mensen het systeem kunnen begrijpen in gebruikscontext
- mensen kunnen ingrijpen of override-en
- escalatie en fallback zijn geregeld

Belangrijk: "human in the loop" is niet automatisch goed. Het gaat om **effectieve** controle, inclusief training, UI/UX, workflow en bevoegdheden.

5.7 Accuracy, robustness & cybersecurity

De AI Act legt nadruk op:

- prestaties binnen opgegeven grenzen
- weerstand tegen manipulatie/attacks (o.a. adversarial, data poisoning)
- security-by-design en passende mitigaties

---

6. Conformiteitsbeoordeling: self-assessment vs notified body

Voor hoog-risico AI moet je aantonen dat je voldoet voordat je de markt op gaat of inzet.

- Voor veel **Annex III** use-cases is een **interne beoordeling** mogelijk (self-assessment), afhankelijk van de situatie en standaarden.
- Voor **Annex I** (safety components in gereguleerde producten) is vaak een **derde partij** (notified body) betrokken via bestaande product compliance routes.

Praktisch: organisaties moeten leren werken zoals bij productcompliance:

- requirements mapping
- evidence packs
- auditeerbare documentatie
- change control

---

7. Fundamental Rights Impact Assessment (FRIA)

Voor bepaalde hoog-risico inzet (met name publieke dienstverlening of vergelijkbare impact) geldt een **fundamental rights impact assessment** vóór inzet.

Een FRIA lijkt op een DPIA (AVG), maar focust breder op grondrechten:

- wie wordt geraakt en hoe?
- risico's (discriminatie, autonomie, toegang, fairness, etc.)
- menselijke controlemaatregelen
- mitigaties, governance en escalatie

Slimme aanpak: integreer FRIA met DPIA en je bestaande risk frameworks, zodat je geen parallelle bureaucratie bouwt.

---

8. General-Purpose AI (GPAI): aparte laag in de wet

De EU maakt onderscheid tussen:

- **GPAI modellen** (foundation models)
- **GPAI systemen/applicaties** die daarop gebouwd zijn

GPAI-modellen worden niet automatisch "hoog-risico" door het model zelf; hoog-risico hangt vooral af van de downstream use-case. Maar providers van GPAI-modellen krijgen wel verplichtingen.

8.1 Basisplichten voor GPAI-modelproviders

Onder meer:

- technische documentatie (voor toezichthouders en downstream providers)
- informatie over capabilities/limitations
- aandacht voor copyright en trainingsdata-samenvattingen
- maatregelen rond vertrouwelijkheid/IP

8.2 Systemisch risico (high-impact / systemic risk)

Voor modellen met systemisch risico gelden zwaardere eisen zoals:

- model evaluations en risk assessments
- incident reporting
- adversarial testing
- cybersecuritymaatregelen
- rapportage over energie/efficiency (in bepaalde vorm)

Voor organisaties die alleen een GPAI-model **gebruiken**: je krijgt vooral te maken met transparantie, contractuele eisen, supplier assurance en (als je het integreert in hoog-risico use-cases) met de hoog-risico verplichtingen.

---

9. Sandboxes en real-world testing

De AI Act stimuleert innovatie via **AI regulatory sandboxes**: gecontroleerde omgevingen onder toezicht van autoriteiten. Dit is vooral interessant voor:

- startups en SMEs
- partijen die hoog-risico AI willen testen met "regulatory guidance"
- organisaties die bewijs willen opbouwen richting conformiteit

Daarnaast bestaat het concept van gecontroleerde real-world tests buiten de sandbox met strikte voorwaarden (o.a. plan, toezicht, consent, incident reporting).

---

10. Handhaving en governance: wie controleert?

Lidstaatniveau

Lidstaten moeten markttoezicht en notificatie organiseren (incl. notified bodies). Autoriteiten kunnen documentatie, datasets en in sommige gevallen zelfs broncode opvragen voor onderzoek.

EU-niveau

Een EU-structuur coördineert vooral rond GPAI en consistentie tussen lidstaten.

---

11. Boetes en consequenties

De AI Act kent een **stevige sanctiestructuur** die oploopt naargelang ernst (verboden AI het zwaarst). Naast boetes kunnen autoriteiten:

- systemen laten terugroepen
- systemen van de markt laten halen
- aanvullende maatregelen opleggen

Voor bestuurders en risk committees betekent dit: AI Act compliance is niet “nice to have”, maar een core operational risk.

---

12. Implementatie: hoe pak je dit aan als organisatie?

Een werkbare aanpak in 8 bouwstenen:

1. **AI-inventaris**
    - welke systemen, waar, door wie, met welk doel
    - inclusief shadow AI en citizen development
2. **Rolbepaling (provider/deployer/etc.)**
    - vooral bij integratie, fine-tuning, eigen branding, substantiële wijzigingen
3. **Use-case classificatie**
    - verboden / hoog-risico / transparantie / laag risico
    - leg beslissingen vast (audit trail)
4. **Control framework ontwerpen**
    - RMS, data governance, logging, oversight, cyber controls
    - koppel aan bestaande ISO/ITGC/SDLC/ERM waar mogelijk
5. **Supplier & model assurance**
    - contractuele eisen, documentatie, testresultaten, incidentprocessen
6. **FRIA/DPIA integreren**
    - één geïntegreerd impact assessment proces
7. **Operationaliseren in SDLC/MLOps**
    - change management, drift monitoring, retraining governance, release gates
8. **Training & AI literacy**

- niet alleen awareness, maar rolgerichte training (dev, product, legal, risk, procurement)

---

13. Wat verandert er strategisch?

De AI Act maakt “trustworthy AI” meetbaar en toetsbaar. Dat betekent:

- meer vraag naar evidence, controls en auditability
- sterkere positie voor organisaties die hun AI lifecycle professioneel beheersen
- product- en procurementteams krijgen een grotere compliance-rol
- governance rond data, security en modelbeheer wordt een kerncompetentie