

Europese AI-Act: veilige, eerlijke en transparante AI-systemen

Regels voor AI-systemen	Risico	Soort systemen	Toezicht	Handhaving
Niet van toepassing	Onacceptabel Risico = VERBODEN	(Real-time) biometrische identificatiesystemen op afstand in openbare ruimtes. Biometrische categoriseringssystemen en/of voorspellend politiewerk met kenmerken zoals geslacht, etnische afkomst, religie, eerder crimineel gedrag Emotieherkenning bij wetshandhaving, grenscontrole, werkplek en onderwijs. Biometrische gegevens van sociale media of videos tbv gezichtsherkenning. Sociale scores die leiden tot discriminatie in sociale contexten. AI die kwetsbaarheden van een specifieke groep mensen uitbuit.	Onaangekondigde inspecties door Autoriteit Persoonsgegevens (AP)	30 miljoen euro of 6 procent van de jaarlijkse wereldwijde omzet
VERPLICHT contact met toezichthouder over kwaliteits- en risicomanagement, Data governance Kwaliteitseisen voor trainings-, test- en validatiegegevens, Technische documentatie en logging, Transparantie t.a.v. systeemarchitectuur, algoritme en werking, Menselijk toezicht, Conformiteitsverklaring, CE-markering en registratie in EU-databank	Art 6: Hoog Risico Voldoen aan EU-regels tijdens hele levensduur	Biometrische en biometrisch ondersteunde systemen, Systemen tbv beheer en exploitatie van kritieke infrastructuur, Algemeen- en beroepsonderwijssystemen, Systemen van openbare (overheids) diensten, Systemen tbv werkgelegenheid, personeelsbeheer, politie, migratie, asiel en grenscontrole, justitie en democratische processen, fraude-detectie	Onaangekondigde inspecties ter plaatse en op afstand door AP, gecoördineerd door EU-AI-Raad → ook sector specifieke toezichthouders	30, 20 of 10 miljoen euro of 6, 4 of 2 procent van de jaarlijkse wereldwijde omzet, afh van soort overtreding
Consumenten informeren over het AI-systeem, Gedragscode (aanbevolen)	Gelimiteerd risico <i>Transparantie eisen</i> tijdens de hele levensduur	AI-systemen die rechtstreeks met mensen communiceren: o.a. emotieherkenningssystemen, biometrische categoriseringssystemen, door AI gegenereerde of gewijzigde inhoud die lijkt op echte personen, objecten, plaatsen of gebeurtenissen (deep fake), Foundation models en LLMs, Generatieve AI (zoals ChatGPT).	Onaangekondigde inspecties ter plaatse en op afstand door AP, gecoördineerd door EU-AI-Raad → ook sector specifieke toezichthouders	10 miljoen euro of 2 procent van de jaarlijkse wereldwijde omzet, i.h.g.v. onvolledige info
Gedragscode (aanbevolen)	Minimaal Risico: <i>Geen verplichtingen</i>	Alle systemen die niet in bovenstaande categorieën vallen: systemen t.b.v. voorspellend onderhoud, Spam filters, video games, magazijn systemen, research & development (in sandbox), systemen die geen impact hebben op individuele personen	N.v.t.	N.v.t

Ontwikkelaars, importeurs, leveranciers van AI-systemen passen zelf de regels toe en bepalen zelf de risico-classes (self-assessment)

Copyright © 2023 , drs Kees van den Tempel, AI-Labs BV