

AI Regulation & Risk



The tier is set by the party that pays for it

Every obligation in a graded scheme hangs off one field the regulated body fills in about itself, and no scheme in force requires a countersignature.

The party that would carry the cost of the higher tier assigns the tier, and assigns it downward. Dutch central-government bodies rated 155 of 287 live or in-development systems minimal risk and 30 high risk, and held a documented risk assessment for 132 of the 287; the national audit office records the incentive plainly, since a system that is not high risk escapes the strict requirements¹. The other published self-declaration bounds nothing: 445 of 3,611 US federal use cases are declared high-impact, 12.3 per cent².

Who applies the rule moves the number further than which rule is applied. Under one definition inside one government, the top-tier share runs from 0 of 247 at one department to 215 of 367 (58.6 per cent) at another, with 2 of 447 at the largest estate in the file². Across three national schemes it runs 2.7 per cent (41 of 1,538), 9.5 per cent (39 assessments against 412 registered systems) and 12.3 per cent, a factor of 4.5^{2,3,4,5}. A national average staffs a plan for neither end, so forecast intake body by body.

No published share has been audited. Of the 3,611 rows in the largest inventory, 426 leave the tier field blank²; a state audit service has opened an investigation into the quality and completeness of registrations in its own national register⁶; and no register field anywhere records who assigned a tier or when². Declaration quality has never been held constant, so the 4.5-fold spread across schemes cannot be attributed to scheme design.

The year's intake is four numbers multiplied together: estate size, times the share the scope test admits to a register, times the share the tier test promotes, minus the share rebutted back out. A different party sets each factor, and clearance capacity is fixed in a separate budget round². High-risk duties for stand-alone systems apply from 2 December 2027 under the Act as amended, and that is when the intake arrives^{7,8}. Set the admitted volume against a published volume-times-price derivation.

WHO DECLARES THE TIER

0 to 59%

top-tier share across bodies under one definition, against 445 of 3,611 overall²

ARGUED OUT OF THE TOP TIER

110 of 555

presumed high-impact cases declassified in one year, 19.8 per cent²

TIER MOVEMENT, TWO YEARS

32 down, 8 up

of 1,064 systems matched across two annual inventories^{2,9}

HOW WE KNOW

Two governments' machine-readable registers, read on 29 August 2026. Every figure here is self-declared and none has been audited. Units of count differ across bodies, one department enumerating individual clinical models where another reports at programme level, so part of the 0-to-59 spread is granularity; 426 blank rows also mean some apparent zeros are non-reporting.

Cost is charged per organisation, not per system

Assessment cost is charged once per organisation, so the number of bodies that must build assessment machinery prices the scheme.

The first assessed system in an organisation costs €193,000 to €330,050 one-off plus €71,400 a year, or €117,750 to €174,800 where an existing quality-management system is upgraded; each further system costs €13,000 to €37,500, a fifteenth of the first at the low end of the study's range and a ninth at the high end¹⁰. Promoting one system in each of ten bodies costs five to six times what promoting ten systems in one body costs.

Three routes carry three prices. Third-party review runs at €400 an hour against €32 for the applicant's own staff, putting one stand-alone system at €23,000, of which €18,200 is the reviewer's fee¹⁰, and one assessor reading documentation clears 88 to 220 files in a 220-day year¹¹. The in-house route is the only one measured in the field, at six contact hours for a group of six to eight in each of fifteen public bodies, so one assessment costs 36 to 48 participant-hours and a hundred promoted systems 3,600 to 4,800 hours drawn from operating units, plus one facilitator post¹². Self-assessment with no reviewer, the route Article 27 enacts, costs the deploying body its own hours and the state nothing^{10,13}.

The budgets set against that price are headcounts with no derivation attached. One central body fell from a proposal of 448.15 million zloty and 100 staff to an enacted ten-year ceiling of 247.79 million against 70 posts^{14,15}; one supervisory agency carries 31 established posts, five of them systems-analyst posts¹⁶; one state spread the duty across 13 market surveillance authorities on a €1.5 million start-up line and published a headcount for none¹⁷. Seventeen of the 19 states with a designated or proposed model chose that route¹⁸.

The EU answered the bottleneck by removing the reviewer. Article 27 requires a public-body deployer to assess fundamental rights impacts and notify the completed template, and specifies no review of it by any authority¹³. The duty was absent from the 2021 proposal¹⁹ and appears nowhere in the support study behind every cost figure the Commission published¹⁰; the state receives a template nobody is obliged to read.

FIRST SYSTEM AGAINST TENTH

9x to 15x

€193,000 to €330,050 to enter the assessed tier, then €13,000 to €37,500 a system, low paired with low and high with high¹⁰

REVIEWER AGAINST OWN STAFF

12.5 to 1

€400 an hour against €32; the hours behind the reviewer's step trace to one 2021 interview, the rates to the study's own table¹⁰

WHAT THE RATIO ASSUMES

The fixed block is dominated by a provider-side quality-management duty under Article 17, which a public-sector deployer may not carry at all⁷; that would shrink the fixed-to-marginal ratio. All figures are 2021 ex-ante and were never tested against outturn.

A domain boundary mis-prices in both directions at once

Widening or narrowing the domain list cannot fix a boundary drawn on domain, because the two errors move in opposite directions when the list changes size.

Obligations attach to the area of use, and the size and reversibility of a wrong output change nothing about which tier a system lands in. A tool that reformats an eligibility application sits inside the same Annex III area 5 as the system that decides the eligibility²⁰. Escape from that area is permitted on four narrow grounds only, in writing and before deployment, so the low tier carries a documentation cost of its own²¹.

Over-inclusion spends the assessor's year on systems whose worst output is a badly formatted form, and the assessor-year is the binding quantity: at the Commission's own price of one to two and a half days a file, one assessor doing nothing else clears 88 to 220 files¹¹. Under-inclusion leaves the highest-volume public-facing systems with a disclosure line, no human review, no appeal route and no monitoring duty. A chatbot that decides nothing about any named person can misinform at population scale, and no domain list reaches it.

The one scheme in force that scores magnitude fails the other way, by settling in its own lower half. Of 39 published algorithmic impact assessments, 37 print a level: 15 at Level I and 22 at Level II, none above the 50 per cent cut point on a ladder whose boundaries sit at 25, 50 and 75 per cent of the maximum. A 15 per cent self-scored mitigation deduction was applied in 9 of the 34 assessments printing both a raw and a current score, taking one from 58 to 49^{4,22}. The two-reviewer and sign-off rungs that make that ladder expensive have never been executed by the government that wrote them^{4,22}.

Only the disclosure tier has ever scaled. One national register holds 1,538 entries against 39 scored assessments in another, forty to one^{3,4}, and a third register took fifteen months to move from 59 records to 143 for the whole of central government²³. Observed assessment output runs one to two orders of magnitude below the theoretical ceiling, at about eight published assessments a year across an entire federal government^{4,11}.

ONE ASSESSOR YEAR

88 to 220

files, at the Commission's price of one to two and a half days per documentation review; every over-included case is drawn from that number¹¹

A SELF-SCORED LADDER

37 of 37

published assessments printing a level sit at Level I or II, none past the 50 per cent cut point, in six years^{4,22}

WHAT IS NOT KNOWN

No one has scored four classification schemes over one fixed inventory of 200-plus systems, so the top-tier volume a state must fund is unknown to within an unknown factor. The Dutch register and the US federal use-case inventory are both large enough to run it on. The within-scheme evidence sets one condition on the design: the same team has to score every case, or the experiment measures the declaring body again.

No register records who assigned a tier, when, or whether anyone checked it

The clearance chain and the audit trail are different objects, and only the first exists.

No register schema treats assignment as an act with an author. Neither year of the largest published inventory, 54 columns then 36, carries a field for who assigned a classification, when, or whether anyone checked it^{2,9}.

Independent review is mandatory for deployed top-tier cases and is recorded for 36 of 367, with 242 leaving the required field blank^{2,24}. The field asks whether the use case was reviewed; it records nothing about the tier, so none of the 36 is a check on a classification.

Where a rebuttal is counted, it runs one way. In one year 110 of 555 presumed high-impact cases came back determined not high-impact, 19.8 per cent, at rates of 47.1, 21.9, 86.7 and 0 per cent across four bodies^{2,24}. No field records a rebuttal a chief officer refused, so the reversal rate on the reversals is recorded nowhere.

Annual re-declaration moves tiers down about four times as often as up. Matching two annual inventories gives 1,064 systems present in both; 32 of the 209 in the top tier came back out, against 8 moving in, and 1,069 of 2,133 earlier entries have no counterpart at all because the earlier file carries no identifier^{2,9}. The categories changed between the two years, so part of that movement is the rulebook applied by the same people.

Where no review exists the error surfaces as disagreement between bodies. Three central-government organisations gave fingerprint comparison three different tiers, minimal, limited and high¹. More than half of 287 reported systems had no risk assessment behind the label, and nor did a third of those the organisations themselves called high risk.

Classification has no expiry. No register, at 1,538, 412 or 143 entries, carries a date of last classification or a count of tier changes^{3,5,23}, so a system admitted to the low tier while it drafted text keeps that tier when it begins ranking cases. The obligation set moves underneath the tier too: an amendment in force from 27 July 2026 lets a deployer discharge parts of the fundamental rights assessment by cross-reference to a data protection impact assessment, with no change to the system⁸.

No state has ever measured the error rate at the point every obligation hangs on, because nothing records a review of a tier. The sign-off chain that does exist publishes a role title and leaves every individual off the public record²⁵, and no administration publishes how many tier assignments were reviewed outside the delivery team or how many changed on review^{1,2}. That absence is the finding, and it is what puts the register amendment ahead of the categories.

Human oversight is the one rung ever tested, and it lowered accuracy in the laboratory

No published evaluation ties a tier obligation to a change in error, overturn or contest rate on a live government caseload.

Obligations sort into three evidence classes: measured on the obligation's own task and shown to change an outcome, measured and shown not to, and unmeasured. Human oversight has been measured and lowered accuracy; every other rung is unmeasured. The registers that record which obligations are in place record no outcome against them^{2,26}.

Human oversight is the most prescribed rung and the only one experimentally tested, and it left accuracy below the model's own in every treatment. Across 1,156 participants on a pretrial task and 732 on a lending task, mean accuracy was 0.768 against the model's 0.803 and 0.793 against 0.823; all six presentation formats beat the no-algorithm baseline and all six scored below the model alone, and participants' rating of the tool's accuracy bore no significant relationship to how accurate it was²⁷. One review counts 41 policies prescribing human oversight of government algorithms²⁸, and three further experiments across 2,854 people found no evidence that labelling advice as algorithmic changed adherence²⁹.

The one two-reviewer rule anywhere with a measured error delta is screening mammography, where a blinded second reader on 99,013 screens raised detection from 6.2 to 7.0 cancers per 1,000 and recall from 3.0 to 3.6 per cent, adding 82 cancers on 612 further recalls, about seven extra callbacks for each extra cancer³⁰. What made that number computable was a task repeated at volume with a checkable answer and a fixed arbitration step. Peer review of one system's design is a single judgement on an unrepeatable object.

Take-up of the unmeasured rungs is thin where it has been counted. Of 445 declared high-impact use cases, 227 are deployed; 16 of those record an established appeal route and 10 record all five of testing, assessment, independent review, monitoring and appeal^{2,24}. The one disclosure duty audited from outside produced the same shape: 391 employers checked, 18 audit reports and 13 transparency notices posted, and the authors state the law does not let anyone tell whether discrimination fell³¹.

HUMAN REVIEW, MEASURED

0.768

mean accuracy against the model's own 0.803 on a pretrial task, and 0.793 against 0.823 on lending²⁷

A SECOND READER'S DELTA

82 for 612

extra cancers for extra recalls on 99,013 screens, about seven callbacks per extra cancer³⁰

LIMITS OF THIS EVIDENCE

The oversight experiments used crowdworkers answering vignettes, and no official was tested on a live file. The outcome measured was prediction accuracy. Defect detection was never tested, and a reviewer checking a processing error is doing a different job from one forecasting a person's behaviour. Absence of measurement is not evidence of absence of effect, so none of this is a reason to drop the rung.

A hundred per cent deadline compliance discloses nothing about the wait

An oversight body can meet its statutory deadline on every case and still hide the queue.

A published compliance rate is compatible with every decision landing at the ceiling of the clock. One of the few statutory-clock regimes that publishes case timing received 1,324 notifications in a year, reviewed 1,220, called in 54 and cleared 1,166, and took every decision inside the statutory 30 working days, publishing a median of 29 for the 54 it called in and no time at all for the 1,166 it cleared, 95.6 per cent of those reviewed³². Acceptance onto the clock carries no deadline of its own and took a median 11 further working days³².

The assessment machinery the AI Act reuses has no deadline on the assessor and clears in years. As of 28 February 2026, the private conformity assessors the EU designates, called notified bodies, had received 32,898 applications and issued 18,010 certificates, with 62 per cent reporting 13 to 18 months for a quality-management certificate and 10 per cent past 24 months on the fuller assessment³³. That is known only because a recurring survey of the assessors was commissioned; no statute requires them to report a clearance time³⁴.

Four published quantities make delay legible, and a compliance rate is none of them: the time on the untimed intake stage, the distribution over all cases including those cleared, the count disposed of by expiry of the clock, and a periodic survey of the assessors. The Regulation clocks none of this, setting no deadline on intake, none for the authority under Article 46 and none for sandbox admission under Article 57^{35,36}.

The expiry default decides which of two invisible shortages a state gets. One statute requires an individual opinion within 30 days, or 60 for a case declared particularly complicated, counts failure to issue as an opinion agreeing with the applicant, and refunds the 150 zloty fee in full, while disapplying both the fee and the default on its own sandbox route where demand is highest¹⁴. The Regulation carries its own opt-in version, a 30-day tacit approval for real-world testing plans³⁷: one Member State switched it on in July 2026³⁸, another left the 30 days without a consequence, its administrative law resolving a case by silence only where a special provision says so³⁹. Seven of 27 Member States have any enabling instrument to tabulate¹⁸.

DECISION ON DAY 29

29 of 30 days

median call-in decision, behind 11 further working days of untimed intake³²

ASSESSOR WITH NO DEADLINE

13 to 18 months

reported by 62 per cent of notified bodies for a quality-management certificate³³

WHAT IS NOT PUBLISHED

No AI authority anywhere publishes a clearance time. Both regimes here are analogies: investment screening is a call-in judgement on a transaction at roughly 1,300 cases a year, and the notified-body distributions run across bodies, not across cases, so a body issuing three certificates counts equally with one issuing three hundred. The natural carrier for an assessments-per-filled-post ratio already exists in the two-yearly resource report under Article 70(6)⁴⁰.

4.1

What sets a system's tier: its domain, a score it gives itself, or a countable property of its decisions?

Two of the three rules are in force, and both are answered by the body that pays for the answer.

	A · Domain list with written rebuttal	B · Self-scored magnitude score	C · Countable decision properties
BENEFITS	Already law; answerable from a use-case description, so no specialist is consumed ^{20,21}	The one instrument in force that scores magnitude; score published with the level ^{4,22}	Every input observable and disputable from outside the delivery team ⁴¹
TRADE-OFFS	Blind to magnitude: a form reformatter and a benefit decision share a tier ²⁰	Consumes the specialist at classification; about eight assessments a year ⁴	No jurisdiction runs it; no inherited categories, template or comparator
RISKS	Rebuttal rate set by the party paying, 0 to 86.7 per cent by body ²	A self-scored discount settles in the lower half; 37 of 37 at Level I or II ^{4,22}	Thresholds drafted blind; a body that counts its own decisions sets its own tier
CONDITIONS	Rebuttal registered, counted and published with the grounds relied on ²¹	Mitigation credit recorded separately and barred from moving a level ²⁶	Annual count of decisions issued, contested and changed as the price of register entry ²⁴

RECOMMENDATION

Adopt Option C. Key each level to annual decision volume, power to deny or withdraw a benefit, and whether a human signs the outcome; use the domain list only to route a case into scoring^{20,21}, and bar mitigation credit from moving a level²². Score exposure independently of denial power, so a high-volume public-facing system that decides about no named person still picks up monitoring and a correction route⁴¹.

OWNER · MINISTER SPONSORING THE ENABLING INSTRUMENT

DECIDE BY · IN FORCE BEFORE 2 DECEMBER 2027

RESIDUAL RISK · THRESHOLDS WITHOUT A COMPARATOR

No jurisdiction runs this test, so commit in the instrument to revising the thresholds after the first year of returns. Define the counting unit in the template and audit it in the same sampled pass as the rebuttals^{2,24}.

4.2

Where does the state spend its one check on a self-assigned tier?

Coverage and depth trade directly, at 3,611 reviews against 110². Every option here needs register fields that do not exist yet⁹.

	A • Countersignature before go-live	B • Sampled audit of declassifications	C • Register cross-comparison
BENEFITS	Acts before harm; no system enters service on an unchecked tier	Targets the act with the incentive; 110 rebuttals against 3,611 use cases ²	Costs no review body; works on published register data and scales with it ³
TRADE-OFFS	Dearest per system; equal effort on the 2,630 uncontested rows and the 110 contested ones ²	Sees nothing that was never flagged	Detects disagreement; blind to uniform understatement
RISKS	Degrades into notification; independent review recorded for 36 of 367 deployed cases ^{2,24}	Reversal counts noisy while the sample is small	Read as a league table, bodies copy each other's tiers and the signal dies
CONDITIONS	Register names the individual who signed, and the classification date ²⁵	Rebuttal registered with a written justification before deployment ²¹	Register entries carrying a technique field and a purpose field ³

RECOMMENDATION

Adopt Option B and run Option C first, because the count of divergent clusters is what sizes the sample; keep countersignature for the top tier alone¹. Size the sample against the rebuttal count, 19.8 per cent of presumed cases where it has been measured, stratify by body, and publish two integers a year: the number reviewed and the number reversed².

OWNER • DESIGNATED MARKET SURVEILLANCE AUTHORITY, WITH THE REGISTER OWNER RUNNING THE CROSS-COMPARISON PASS

DECIDE BY • FIRST ANNUAL WORK PROGRAMME

RESIDUAL RISK • THE BODIES THAT FLAG NOTHING

A body reporting no presumed cases contributes no rebuttals and escapes the sample, so treat that as the audit finding in itself². Annual re-declaration against a persistent identifier is the only mechanism producing tier movement anywhere and has no reviewer in it, so it supplies data and no assurance⁹.

4.3

What happens when the review clock expires?

The clock, the fee refund and the expiry default are three separable drafting choices, and the default alone decides what a shortage looks like from outside. Settle it before the first case arrives.

	A • Deemed approval with fee refund	B • Clock with no stated consequence	C • Clock, published distributions, confined deemed consent
BENEFITS	Applicant gets a date; the state pays 150 zloty per lapsed case ¹⁴	Nothing deploys on the authority's inaction; the power to say no late survives	A date for the applicant, a number for the finance ministry, a distribution for the public
TRADE-OFFS	No queue statistic and no complainant	A waiting list only the body can see	Commits the state to publishing figures that will be bad for the first years
RISKS	Backlog becomes unreviewed approvals at statutory speed	100 per cent compliance reported at a median of 29 days out of 30 ³²	A compliance rate published alone recreates the failure it was meant to fix
CONDITIONS	Confine to opinions on facts the applicant supplied ¹⁴	Only defensible with published clearance distributions	Per-tier case record and a fixed reporting template ⁴⁰

RECOMMENDATION

Adopt Option C. Clock intake as well as decision, publish median and 90th-percentile clearance times by tier each quarter over all cases including those cleared, and annually the count disposed of by expiry^{32,40}. Confine deemed consent to tiers requiring no specialist review, keep the fee refund as the applicant’s remedy in every tier¹⁴, and where the assessors sit outside the state, fund a recurring survey of them, the only instrument that has surfaced an assessment backlog³³.

OWNER • DESIGNATED AUTHORITY, WITH THE SPONSORING MINISTER SETTLING THE DEFAULT IN THE ENABLING INSTRUMENT

DECIDE BY • REPORTING TEMPLATE FIXED BEFORE 2 DECEMBER 2027

RESIDUAL RISK • THE ARTICLE 60 BINARY

One binary is already drafted: opt in to the 30-day tacit approval for real-world testing plans, or leave the 30 days without a consequence³⁷. One Member State opted in in July 2026 and another did not, so the answer varies by state^{18,38,39}.

Five register fields come before the categories

Auditing the people who fill in the tier buys more than redrawing the ladder, so the register amendment is first and the tier definitions are fourth.

The minister owning the national algorithm register adds five fields to the return template: a persistent system identifier, the date the tier was assigned, the individual who assigned it, the individual who countersigned it for top-tier entries alone under the market surveillance authority's own duty, and the named trigger that forces a re-run^{2,3,5,9,23,25}. Set the triggers as a change in decision role, affected population, data source or volume of use. Publish annually the tier changes up and down, the assignments reviewed outside the delivery team, the number changed on review, and each new field's completion rate: 36 of 367 deployed top-tier cases carry a recorded review, 242 leaving the field blank². **Before the next annual inventory cycle.**

The register owner runs one pass a year over its own published data, clusters entries sharing a technique and a purpose, and refers every cluster whose tiers disagree to the market surveillance authority. Three central-government organisations gave fingerprint comparison three different tiers¹. Publish the clusters examined and the number referred, never a league table by body^{3,6}. **In the first annual work programme, ahead of the sampled audit.**

The designated market surveillance authority re-reviews a fixed sample of cases argued out of the top tier, publishing the number reviewed and the number reversed each year. Size the sample against the rebuttal count, 19.8 per cent of presumed cases where measured, and stratify by body^{2,24}. Require a written justification for each rebuttal, registered before deployment under Article 6(3), and publish it²¹. **In its first annual work programme.**

The minister sponsoring the enabling instrument keys each level to annual decision volume, power to deny or withdraw a benefit, and whether a human signs the outcome, uses the domain list only to route a case into scoring^{20,21}, bars mitigation credit from moving a level²², and scores exposure independently of denial power, so a high-volume public-facing system deciding about no named person still carries monitoring and a correction route⁴¹. Publish each tier's thresholds and the obligations attaching at it. **In force before 2 December 2027⁸.**

5.1

The template amendment precedes the categories

Every check above needs fields no register carries. The cross-comparison pass costs no review body and sizes the sampled audit, so it runs second. The categories are redrawn last, because within one definition the top-tier share runs from 0 to 59 per cent across bodies, while the choice of scheme moves it by a factor of 4.5^{2,3,4}.

Fund the cadre against a derivation, and book no benefit that has not been measured

A headcount loses the spending round; expected annual volume times unit price, with the utilisation assumption stated, is the object that survives it.

The finance ministry funds assessment work as one standing facilitator cadre serving all deploying bodies, and publishes the derivation alongside the establishment. Leave decision powers with the sectoral bodies and concentrate the assessment work, so few organisations pay the fixed cost of €193,000 to €330,050 plus €71,400 a year while many draw on the cadre at €13,000 to €37,500 a system¹⁰. Assume attendance sets throughput: the six contact hours run across five meetings, the substantive sessions two to three weeks apart¹². Publish annually the cases admitted, cases cleared and filled technical posts. State on the costing that the unit price behind every European figure traces to one 2021 interview, and measure time and motion on the first ten promoted systems. Set each year's admitted top-tier volume against that published derivation before the establishment is voted. **In force before 2 December 2027, when the intake arrives⁸.**

The designated authority publishes, quarterly and by tier, median and 90th-percentile clearance times over all cases including those cleared, and annually the count disposed of by expiry³². Clock intake as well as decision. The enabling instrument confines deemed consent to tiers requiring no specialist review, keeps the fee refund as the applicant's remedy in every tier¹⁴, and settles the tacit-approval binary explicitly^{37,38}. Where the assessors sit outside the state, fund a recurring survey of them as a standing budget item³³. **Reporting template fixed before 2 December 2027.**

The department sponsoring a decision system states in the business case that the human-review and peer-review rungs carry no assumed error reduction, and commissions one measurement: route a random share of one decision type through review, and publish the changed-outcome rate on both sides within twelve months of go-live^{27,28,30}. Add the counter that makes that possible, decisions issued, contested and changed per system per year, to the same template amendment. Defend the rung meanwhile as a process guarantee owed to the person affected. Check the top rung is executable at all: two qualified peer reviewers in an administration with five systems-analyst posts means reviewing a colleague's own work¹⁶. **At business-case approval.**

Four quantities are missing, each answerable from data a state already holds once the five register fields exist: the top-tier share four schemes would put on one fixed inventory, the price of moving one system from the disclosure tier to the assessed tier, the number of self-assigned tiers reviewed and the number changed, and the change in error, overturn or contest rate from any tier obligation.

Sources cited in this brief

IN ORDER OF CITATION

- 1** Algemene Rekenkamer, *Focus op AI bij de rijksoverheid*, 2024.

<https://www.rekenkamer.nl/site/binaries/site-content/collections/documents/2024/10/16/focus-op-ai-bij-de-rijksoverheid/PAC+Focus+op+AI+bij+de+rijksoverheid.pdf>

- 3** Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, *Algoritmeregister dashboard*, 2026.

<https://algoritmes.overheid.nl/nl/dashboard>

- 5** Treasury Board of Canada Secretariat, *Government of Canada AI Register*, 2026.

<https://open.canada.ca/data/en/dataset/fcbc0200-79ba-4fa4-94a6-00e32facea6b>

- 7** European Union, *Regulation (EU) 2024/1689, Artificial Intelligence Act*, 2024.

https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401689

- 9** US Office of Management and Budget, *2024 Federal Agency AI Use Case Inventory*, 2025.

<https://github.com/ombegov/2024-Federal-AI-Use-Case-Inventory>

- 11** European Commission, *Impact Assessment accompanying the AI Act proposal, SWD(2021) 84 final*, 2021.

<https://artificialintelligenceact.eu/wp-content/uploads/2022/06/IA-COM-Impact-Assessment-1-21-April.pdf>

- 13** European Union, *AI Act Article 27, fundamental rights impact assessment*, 2024.

https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02024R1689-20260727#art_27

- 15** WNP.pl, *Spór o nowy urząd do nadzoru AI*, 2025.

<https://www.wnp.pl/tech/ministerstwo-finansow-daje-czerwone-swiatlo-dla-nowego-urzedu-do-nadzoru-ai-wytyka-koszty,963322.html>

- 2** US Office of Management and Budget, *2025 Federal Agency AI Use Case Inventory*, 2026.

<https://github.com/ombegov/2025-Federal-Agency-AI-Use-Case-Inventory>

- 4** Treasury Board of Canada Secretariat, *Algorithmic Impact Assessment collection*, 2026.

<https://search.open.canada.ca/opendata/?collection=aia>

- 6** Binnenlands Bestuur, *Overheden worstelen met mogelijke hoog risico AI-systemen*, 2025.

<https://www.binnenlandsbestuur.nl/digitaal/overheden-worstelen-met-mogelijke-hoog-risico-ai-systemen>

- 8** Official Journal of the European Union, *Regulation (EU) 2026/1744, Digital Omnibus on AI*, 2026.

<https://eur-lex.europa.eu/eli/reg/2026/1744/oj/eng>

- 10** European Commission DG CNECT (CEPS, ICF and Wavestone), *Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe*, 2021.

<https://artificialintelligenceact.eu/wp-content/uploads/2022/06/IA-COM-Impact-Assessment-3-21-April.pdf>

- 12** Muis, Kamphorst and Straatman, *Responsible AI innovation in the public sector*, Journal of Responsible Technology, 2025.

<https://research-portal.uu.nl/ws/files/261655372/1-s2.0-S2666659625000149-main.pdf>

- 14** Dziennik Ustaw, *Ustawa o systemach sztucznej inteligencji*, 2026.

<https://eli.gov.pl/api/acts/DU/2026/1003/text/O/D20261003.pdf>

- 16** Portal de Transparencia, *Relación de puestos de trabajo, AESIA*, 2026.

https://transparencia.gob.es/content/dam/transparencia_home/publicidadactiva/organizacionempleo/rpt/actual/mtdf/aesia/260101-RPT-XEX02-52269-PF.pdf

17 Houses of the Oireachtas, *Consideration of the DETE Estimate for 2026*, 2026.

https://data.oireachtas.ie/ie/oireachtas/committee/dail/34/select_committee_on_enterprise_tourism_and_employment/submissions/2026/2026-03-04_briefing-note-peter-burke-t-d-minister-for-enterprise-tourism-and-employment-department-of-enterprise-to-tourism-and-employment_en.pdf

19 European Commission, *Proposal for a Regulation on artificial intelligence, COM(2021) 206 final*, 2021.

<https://artificialintelligenceact.eu/wp-content/uploads/2021/08/The-AI-Act.pdf>

21 European Union, *AI Act Article 6, classification rules*, 2024.

https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02024R1689-20260727#art_6

23 GOV.UK, *Algorithmic transparency records*, 2026.

<https://www.gov.uk/algorithmic-transparency-records>

25 Department for Science, Innovation and Technology, *Algorithmic Transparency Recording Standard: guidance for public sector bodies*, 2025.

<https://www.gov.uk/government/publications/guidance-for-organisations-using-the-algorithmic-transparency-recording-standard/algorithmic-transparency-recording-standard-guidance-for-public-sector-bodies>

27 Green and Chen, *The Principles and Limits of Algorithm-in-the-Loop Decision Making*, ACM CSCW, 2019.

<https://www.benzevgreen.com/wp-content/uploads/2019/09/19-cscw.pdf>

29 Alon-Barkat and Busuioc, *Human-AI interactions in public sector decision making*, JPART, 2023.

<https://arxiv.org/abs/2103.02381>

31 Wright and others, *Null Compliance: NYC Local Law 144*, ACM FAccT, 2024.

<https://arxiv.org/html/2406.01399v1>

18 Future of Life Institute, *Overview of all AI Act National Implementation Plans*, 2026.

<https://artificialintelligenceact.eu/national-implementation-plans/>

20 European Union, *AI Act Annex III, high-risk AI systems*, 2024.

https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02024R1689-20260727#anx_III

22 Government of Canada, *aia-eia-js scoring implementation*, 2025.

<https://github.com/canada-ca/aia-eia-js/blob/master/src/store.ts>

24 US Office of Management and Budget, *M-25-21, Accelerating Federal Use of AI*, 2025.

<https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf>

26 Treasury Board of Canada Secretariat, *Directive on Automated Decision-Making*, 2025.

<https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>

28 Green, *The Flaws of Policies Requiring Human Oversight of Government Algorithms*, Computer Law and Security Review, 2022.

<https://arxiv.org/abs/2109.05067>

30 Coolen and others, *Impact of the second reader at blinded double reading of digital screening mammograms*, British Journal of Cancer, 2018.

<https://pmc.ncbi.nlm.nih.gov/articles/PMC6134129/>

32 UK Cabinet Office, *National Security and Investment Act 2021 Annual Report 2025-26*, 2026.

<https://www.gov.uk/government/publications/national-security-and-investment-act-2021-annual-report-2025-26/national-security-and-investment-act-2021-annual-report-2025-26-html>

33 European Commission DG SANTE, *20th notified body survey, MDR and IVDR*, 2026.

https://health.ec.europa.eu/document/download/59b9d90e-be42-4895-9f6f-bec35138bb0a_en

35 European Commission AI Act Service Desk, *AI Act Article 46, derogation from conformity assessment*, 2024.

<https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-46>

37 European Commission AI Act Service Desk, *AI Act Article 60, testing in real world conditions*, 2024.

<https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-60>

39 ArsLege, *Art. 122a Kodeksu postępowania administracyjnego*, 2017.

<https://arslege.pl/milczace-zalatwienie-sprawy/k3/a104779/>

41 Government Digital Service, *ATRS Mandatory Scope and Exemptions Policy*, 2024.

<https://www.gov.uk/government/publications/algorithmic-transparency-recording-standard-mandatory-scope-and-exemptions-policy/algorithmic-transparency-recording-standard-atrs-mandatory-scope-and-exemptions-policy>

34 European Union, *AI Act Article 43, conformity assessment*, 2024.

https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02024R1689-20260727#art_43

36 European Union, *AI Act Article 57, AI regulatory sandboxes*, 2024.

https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02024R1689-20260727#art_57

38 Bundesministerium der Justiz, *Paragraph 14 KI-MIG*, 2026.

https://www.gesetze-im-internet.de/ki-mig/__14.html

40 European Union, *AI Act Article 70, national competent authorities*, 2024.

https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02024R1689-20260727#art_70

Policy Briefs



Short guidebooks for ministers and senior officials, published by Numan Policy. Each issue follows one question from evidence to decision, and is reviewed by an independent panel before publication.

HOW TO CITE

Numan Policy, *AI Regulation & Risk*, August 2026.

CORRECTIONS & UPDATES

Evidence moves. Corrections and revised editions are published at policy.numan.ai.