The half of the AI Act that switched on

On 2 August 2026 the European Union turned on its AI labelling rules and left the expensive half of the AI Act switched off until December 2027. What was designed as a product-safety regime for consequential machines — hiring, credit, education, borders — starts life as a disclosure regime for synthetic media. That swap moves the compliance bill away from European deployers and onto the output pipelines of a handful of global model providers, who can absorb it. The cost is not measured in euros: it is the sixteen months in which Europe holds a rulebook it has chosen not to enforce.

AT

AUTHOR

Alicante Tech Vibes ***Research Labs***

~8 MIN READ · ALL FIGURES SOURCED & DATED

2 Aug 2026 ARTICLE 50 TRANSPARENCY DUTIES BECOME APPLICABLE EU-WIDE	**€15M / 3%** MAXIMUM FINE: WHICHEVER IS HIGHER, OF GLOBAL ANNUAL TURNOVER	**16 months** DELAY OF STAND-ALONE HIGH-RISK DUTIES, TO 2 DEC 2027
8 of 27 MEMBER STATES WITH A DESIGNATED CONTACT POINT, MAR 2026	**~190** SIGNATORIES TO THE TRANSPARENCY CODE OF PRACTICE, END JUL 2026	**8%** EUROPE'S SHARE OF GLOBAL AI VENTURE FUNDING, H1 2026

I. What switched on, and what quietly did not

Two things happened in Brussels in the same fortnight, and only one of them was announced as news. On **2 August 2026**, Article 50 of the AI Act became applicable: providers must design systems so a person knows they are talking to a machine, and must mark synthetic audio, image, video and text in a machine-readable way; deployers must label deepfakes and disclose AI-written text published on matters of public interest. The Commission published its implementing guidelines on 20 July 2026, thirteen days before the duties bit. Penalties reach €15 million or 3% of worldwide turnover, whichever is higher.

The second thing is what did not happen. The obligations that give the Act its teeth — risk management, data governance, human oversight and conformity assessment for high-risk systems — were due the same day. They are now due 2 December 2027 for stand-alone Annex III systems and 2 August 2028 for AI embedded in regulated products, under a Digital Omnibus agreed politically on 6 May 2026 and confirmed by member states on 13 May. Systems already on the market before 2 August 2026 get until **2 December 2026** to add machine-readable marking. So the Act that arrived this month is the labelling Act. The safety Act is a 2027 problem.

II. A product-safety law wearing a disclosure law's clothes

The AI Act was drafted on the logic of CE marking: classify by risk, impose engineering duties on the risky tier, verify before market entry. Sequencing matters to that logic. By moving transparency first and safety last, the EU has inverted it — the regime now regulates what AI output *says about itself* for at least sixteen months before it regulates what AI systems *do to people*.

The strategic consequence is a change in who is regulated. Article 50 duties fall mainly on model providers and content platforms — a concentrated set of large firms with unified global pipelines. The high-risk duties would have fallen on thousands of European deployers: HR departments running CV screening, banks scoring credit, universities scoring exams, ministries triaging asylum files. That constituency has been given a reprieve, and it is the constituency that lobbied. A Corporate Europe Observatory and LobbyControl analysis found 69% of the Commission's AI-related meetings in 2025 were with industry groups against 16% with civil society. Read charitably, the delay buys time for

harmonised standards that genuinely are not ready. Read plainly, the EU discovered that its own firms were the ones about to be regulated, and blinked.

III. The label is the weakest part of the machine

The honest technical read is that Article 50 has been switched on before the technology it depends on is dependable. Machine-readable marking rests on watermarking and provenance metadata, and both are removable. A 2025 survey of watermarking for AI-generated images by Cao and colleagues at Queen's University concludes that current systems show notable weaknesses in robustness and security — vulnerable both to removal attacks that regenerate an image through a diffusion or autoencoder model, and to forgery attacks that stamp a false watermark onto content no model produced. The WAVES benchmark it cites exposed previously unrecognised failure modes in every leading scheme tested. Separately, provenance metadata in the C2PA style is routinely stripped when files pass through social platforms, which is precisely the path that matters for deepfakes.

The regulatory scaffolding is voluntary. The Commission's Code of Practice on Transparency of AI-generated Content, published 10 June 2026, had roughly 190 signatories by end-July; signing brings legal certainty and lighter-touch supervision, not immunity. The predictable equilibrium: compliant labels on content from firms that were never the problem, and none on content from actors who strip them. This is a real obligation with a real fine attached, landing on a substrate that a competent adversary defeats in an afternoon.

> ***A large part of high-risk AI systems will never have to comply with the rules.***
>
> BRAM VRANKEN, CORPORATE EUROPE OBSERVATORY, ON THE NON-RETROACTIVITY OF THE DELAYED HIGH-RISK REGIME · TECH POLICY PRESS, 2 APRIL 2026

IV. Who actually pays, and how much

Strip out the headline numbers and the AI Act's cost was always narrower than its reputation. The CEPS analysis supporting the Commission's impact assessment put compliance at roughly 17% of overhead on AI spending — but only for high-risk systems, which the Commission expected to be 5–15% of all AI

systems, and only for firms with no existing quality management system. CEPS costed initial setup at €193,000–€330,000 plus about €71,400 a year in maintenance for such a firm, and dismantled the widely circulated claim of €30 billion-plus in EU-wide costs as an extrapolation that applied the high-risk overhead figure to all AI spending. Those estimates date from 2021 and predate both the generative-AI wave and the Act's final text — treat them as the right order of magnitude, not a current quote.

That reframes what the delay is worth. Deferring high-risk duties by sixteen months does not save European industry a headline sum; it defers a moderate, concentrated cost on a minority of systems. What it does change is *timing*: the firms that already invested in conformity infrastructure to hit August 2026 now hold a stranded asset for a year and a half, while their slower competitors bank the float. Regulation that slips punishes the compliant early mover. That is the quiet tax here, and it falls on exactly the European vendors the Act was supposed to advantage.

V. The sixteen-month window

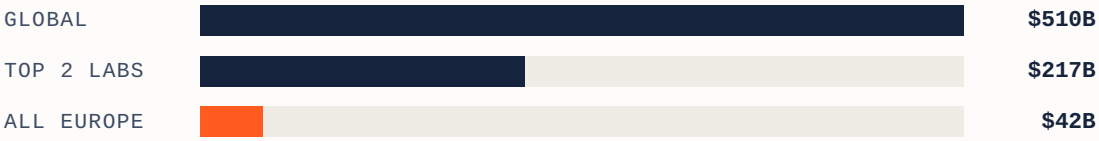
The sharper second-order effect is behavioural. Because the delayed regime is not retroactive, a high-risk system placed on the market before December 2027 can, on the current reading, keep operating outside the new duties unless it is substantially modified afterwards. Laura Caroli, a former AI Act negotiator, has flagged exactly this for hiring systems; German Green MEP Sergey Lagodinsky called the non-retroactivity a loophole and a weak spot. The rational response for any vendor of consequential AI is to ship into the EU now.

Enforcement capacity compounds it. The deadline for member states to designate competent authorities and single points of contact was 2 August 2025; as of March 2026 the European Parliament's research service counted 8 of 27 with a designated contact point. An obligation with a €15 million ceiling and no staffed regulator behind it is a compliance memo, not a constraint — and the Act's decentralised design means the gap will be uneven across the single market, which is itself a fragmentation risk for a law whose selling point was one rulebook.

Meanwhile the capital picture explains why Brussels feels the pressure to soften. Global AI venture funding reached about $510 billion in the first half of 2026, of which OpenAI and Anthropic alone took $217 billion — 43% of every venture dollar deployed worldwide. Europe raised roughly $42 billion, up 50% year on

year and a genuine improvement, but still around 8% of the global total. Europe is not regulating from a position of market strength; it is regulating a supply chain it does not own.

FIG. 1: AI VENTURE FUNDING, FIRST HALF OF 2026 ($BN)



BARS SCALED TO THE GLOBAL TOTAL (=100%). CRUNCHBASE DATA REPORTED BY GOHUB VENTURES, 13 JULY 2026.

VI. The precedent Brussels should fear

Europe has run this experiment before. The ePrivacy and GDPR consent rules produced the cookie banner: a universally deployed, universally ignored disclosure that changed the surface of the web without changing the underlying data economy, and which the Commission itself has since spent years trying to undo. Article 50 has the same structural risk profile — a duty discharged by displaying a notice, on content whose provenance signal degrades in transit, enforced by regulators that mostly do not yet exist.

The counter-precedent is GDPR's substantive core, which did travel: it seeded comparable laws across dozens of jurisdictions because it imposed real engineering obligations that global firms found cheaper to implement once, worldwide. That is the Brussels effect worth having, and it lives in the high-risk chapter, not in the labelling chapter. By shipping the labels first and the engineering later, the EU has exported the part of its regime that is easiest to copy cosmetically and hardest to make bite.

VII. How it plays out: three scenarios

Horizon to end-2028. Probabilities are analytical judgment, not measurement, and they assume no further omnibus is tabled before the indicators below become checkable.

1 · Plumbing, not policing ~55%

2026–2028 · BASELINE

Marking becomes default infrastructure in the major model APIs because shipping it once globally is cheaper than geofencing it, and the code of practice grows past 300 signatories. Enforcement stays nominal into 2027 with no significant Article 50 fine. The high-risk chapter lands on 2 December 2027 roughly on time but with narrowed scope. Watch: signatory count on the Commission register; whether any national authority opens a formal Article 50 proceeding before mid-2027; whether CEN-CENELEC harmonised standards publish before Q2 2027.

2 · An incident forces the issue ~20%

2026–2027 · ACCELERATION

A high-salience synthetic-media event — an election deepfake, a market-moving fake earnings clip — converts a dormant obligation into an enforcement priority overnight. A first fine in the eight figures makes provenance a board-level line item, and Article 50 becomes the de facto global labelling standard by default implementation. Watch: any Article 50 penalty above €10 million; a US federal or state labelling rule that references machine-readable provenance; major platforms preserving rather than stripping C2PA metadata.

3 · The rulebook keeps sliding ~25%

2027–2028 · STALL

A second omnibus narrows or further defers Annex III before it ever applies; standards slip again; labelling collapses into a boilerplate footer nobody reads. National authorities stay unstaffed and the Act loses its normative pull abroad, leaving Europe with the reputational cost of strict regulation and none of the leverage. Watch: any new Commission proposal touching Annex III before August 2027; fewer than 20 of 27 member states fully designated by August 2027; code-of-practice signatories declining rather than growing.

WHAT WOULD FALSIFY THIS THESIS

The argument here is that the Act has become a disclosure regime with weak instruments and absent enforcers. Three findings would break it. First, substantive Article 50 investigations against major providers before mid-2027 that impose real cost — none are visible today, but the Digital Omnibus did grant the AI Office new investigative, inspection and fining powers, so the capacity is being built. Second, independent measurement showing machine-readable marks surviving ordinary redistribution at high rates; the published research points the other way, but watermarking is an active field and the survey evidence cited is a 2025 snapshot. Third, the high-risk deadlines holding on 2 December 2027 with scope intact, which would show the delay was a standards-readiness pause rather than a retreat. The strongest counter-case is the Commission's own: harmonised standards genuinely were not ready, and enforcing conformity assessment against non-existent standards would have produced legal chaos rather than safety. That case is reasonable — it is also indistinguishable in the short run from the case made by the firms that wanted the delay.

WHAT TO WATCH, IN ORDER

1. **2 December 2026** — the grace period closes for generative systems placed on the market before 2 August 2026. Check whether the major providers have actually shipped machine-readable marking, or merely a visible disclaimer.
2. **2 December 2026** — the transition ends for the new Article 5 prohibition on generating non-consensual intimate imagery and CSAM. First real test of whether the AI Office enforces a hard ban rather than a disclosure duty.
3. **Q1–Q2 2027** — CEN-CENELEC harmonised standards for high-risk AI. If they slip again, the stated justification for the December 2027 date collapses in public.
4. **2 August 2027** — member state deadline for regulatory sandboxes, and a natural checkpoint on how many of the 27 have fully designated and staffed market surveillance authorities against the 8 counted in March 2026.
5. **2 December 2027** — Annex III high-risk obligations apply. Watch the preceding six months for a further omnibus; a proposal tabled before mid-2027 is the single clearest signal of the stall scenario.

SOURCES

1. EU Artificial Intelligence Act, “The EU AI Act’s Transparency Rules: A Practical Guide to Article 50,” accessed 12 August 2026 · artificialintelligenceact.eu
2. Cooley LLP, “EU AI Act: Transparency Obligations Take Effect 2 August 2026,” 3 August 2026 · cooley.com
3. Al Jazeera, “What came into force with the EU’s AI Act this week – and what didn’t,” 6 August 2026 · aljazeera.com
4. Gibson Dunn, “EU AI Act Omnibus Agreement — Postponed High-Risk Deadlines and Other Key Changes,” 2026 · gibsondunn.com
5. European Commission, “Code of Practice on Transparency of AI-generated Content,” published 10 June 2026 · digital-strategy.ec.europa.eu
6. Tech Policy Press, “EU’s AI Act Delays Let High-Risk Systems Dodge Oversight,” 2 April 2026 · techpolicy.press
7. European Parliamentary Research Service, “Enforcement of the AI Act,” 18 March 2026 · epthinktank.eu
8. CEPS, “Clarifying the costs for the EU’s AI Act,” 24 September 2021 · ceps.eu
9. Cao, Li, Zhang, Ni & Lu, “Secure and Robust Watermarking for AI-generated Images: A Comprehensive Survey,” arXiv 2510.02384, 2025 · arxiv.org
10. AI IP Protection, “Why C2PA Watermarks Fail on Social Media,” accessed 12 August 2026 · aiipprotection.org
11. GoHub Ventures, “AI Venture Funding H1 2026: US vs Europe” (Crunchbase data), 13 July 2026 · gohub.vc

METHOD NOTE: FIGURES ARE ATTRIBUTED AND DATED INLINE; WHERE INDEPENDENT SOURCES DISAGREE, THE DISAGREEMENT IS REPORTED RATHER THAN AVERAGED. COST FIGURES FROM THE 2021 CEPS IMPACT-ASSESSMENT WORK PREDATE THE ACT’S FINAL TEXT AND ARE USED AS ORDERS OF MAGNITUDE, NOT CURRENT QUOTES.

Built by the Alicante Tech Vibes community, a tope. 2026-08-12