

Source dossier

Supervision was a schedule — a monitor that varies when it looks beats one that looks often, and omitting oversight became a crime in Italy twelve days ago

Edition 2026-10-12 · narrative: Oversight is a schedule you must be able to evidence, not a control you can declare · compiled 2026-10-12

Claims index

Claim	Statement	Status	Sources
CLAIM-2026-10-12-01	Cadence, a runtime monitor that varies its inspection schedule, resolves 42.33% of 300 SWE-bench Lite tasks on mini-swe-agent against 17.00% for the same agent with no monitor.	verified	arxiv2026cadence
CLAIM-2026-10-12-02	On Moatless the same monitor resolves 39.67% against 24.00% unmonitored, a 15.67 percentage-point gap.	verified	arxiv2026cadence
CLAIM-2026-10-12-03	The abstract's "+25.33%" and "+15.67%" are percentage-point differences in resolve rate, not relative improvements; relative, the first is approximately +149%.	verified	arxiv2026cadence
CLAIM-2026-10-12-04	Fixed-interval monitoring on mini-swe-agent is not monotone in the interval: every 5 steps gives 34.00%, every 3 gives 33.33% and every 7 gives 26.33%.	verified	arxiv2026cadence
CLAIM-2026-10-12-05	On mini-swe-agent the monitored run cost 1.03M tokens per task against 1.42M unmonitored, while on Moatless it cost 0.48M against 0.32M.	verified	arxiv2026cadence
CLAIM-2026-10-12-06	On Moatless the strongest baseline monitor reaches 35.33% against Cadence's 39.67%, and the paper reports that gap as not statistically significant.	verified	arxiv2026cadence
CLAIM-2026-10-12-07	Italy's D.Lgs. 9 settembre 2026, n. 160, in force since 30 September 2026, makes the omission of human-oversight measures for a high-risk AI system a criminal offence under a new Art. 437-bis of the criminal code.	verified	normattiva2026reati
CLAIM-2026-10-12-08	The same decree adds Art. 25-vicies to D.Lgs. 231/2001, making AI-committed offences predicate offences for corporate liability, with disqualification sanctions that include a ban on advertising.	verified	normattiva2026reati

Claim	Statement	Status	Sources
CLAIM-2026-10-12-09	Under D.Lgs. 160/2026 a court may order a provider or user to produce records and technical documentation about the algorithm, may draw adverse inferences from an unjustified refusal, and applies a rebuttable presumption of the causal link in the victim's favour.	verified	normattiva2026reati
CLAIM-2026-10-12-10	D.Lgs. 7 ottobre 2026, n. 179 was published in Gazzetta Ufficiale Serie Generale n. 234 of 8 October 2026 and enters into force on 23 October 2026.	verified	normattiva2026ade guamen
CLAIM-2026-10-12-11	That decree designates AgID and ACN as Italy's national AI authorities, AgID as notifying authority and ACN, the national cybersecurity agency, among the market surveillance authorities.	verified	normattiva2026ade guamen, bundesam tfu2026kimig
CLAIM-2026-10-12-12	Under D.Lgs. 179/2026 Art. 41, decisions on hiring, dismissal, performance evaluation and discipline may not rest solely on automated processing, and the final decision is reserved to a natural person exercising effective and autonomous authority.	verified	normattiva2026ade guamen
CLAIM-2026-10-12-13	AI Act Article 14 requires high-risk systems to be designed so they can be effectively overseen by natural persons while in use, and binds Annex III systems from 2 December 2027 and Annex I systems from 2 August 2028.	verified	euaiactexpl2026human, eurlex2026regulatio
CLAIM-2026-10-12-14	Article 14's oversight duty runs to natural persons, so an LLM runtime monitor inspecting an agent trajectory does not on its own discharge it.	verified	euaiactexpl2026human
CLAIM-2026-10-12-15	The EUR 35,000,000 or 7% ceiling quoted in coverage of the Italian decrees is AI Act Article 99(3), for the Article 5 prohibitions; operator duties and Article 50 transparency sit at Article 99(4), EUR 15,000,000 or 3%.	verified	euaiactexpl2026penalties
CLAIM-2026-10-12-16	Regulation (EU) 2026/1744, the Digital Omnibus on AI, in force 27 July 2026, moved Annex III high-risk to 2 December 2027, Annex I to 2 August 2028 and member-state sandboxes to 2 August 2027.	verified	eurlex2026regulatio

Only claims marked **verified** may carry a post. **single-source** may appear as hedged context; **disputed** must not be published at all.

Cadence: Strategic Guidance for Coding Agents

PRIMARY

Publisher	arXiv
Published	2026-10-08
Retrieved	2026-10-12
Cite key	arxiv2026cadence
URL	https://arxiv.org/abs/2610.12269

Assessment. Read from the abs record and the v1 HTML rendering at arxiv.org/html/2610.12269v1. Affiliations taken from the paper's own front matter, not from any aggregator. Identifier verified against the story before any figure was used: title and YYYY both match an 8 October 2026 submission.

Claims drawn from this source. CLAIM-2026-10-12-01, CLAIM-2026-10-12-02, CLAIM-2026-10-12-03, CLAIM-2026-10-12-04, CLAIM-2026-10-12-05, CLAIM-2026-10-12-06

Front matter

Authors and affiliations. Minxing Wang (Singapore Management University, Singapore); He Ye (University College London, United Kingdom); Earl T. Barr (University College London, United Kingdom); Yintong Huo (Singapore Management University, Singapore).

Version. v1, Thu, 8 Oct 2026 16:34:02 UTC. Single version. cs.SE.

Argument

Runtime monitors improve LLM-based coding agents by inspecting their execution trajectories and issuing corrective guidance when they detect misbehaviour. The paper's claim is that existing monitors are badly scheduled rather than badly informed: they use either fixed inspection intervals, which miss severe failures while wasting effort on healthy runs, or rigid heuristic rules, which miss complex reasoning errors.

Cadence has two parts. A **two-tier intervention module** issues *advisory* guidance for normal runs and minor lapses, and *replacement-level* guidance for severe misbehaviour. An **inspection scheduler** tightens supervision after replacement-level guidance and relaxes it after advisory-level guidance. The inspection schedule is therefore a function of the run's observed health rather than of step count.

Measured results

Benchmark: 300 SWE-bench Lite tasks. Agents: mini-swe-agent and Moatless.

Resolve rate, mini-swe-agent:

- Cadence 42.33%
- Fixed-Interval k=5 34.00%

- Fixed-Interval k=3 33.33%
- LivePlan 26.67%
- Fixed-Interval k=7 26.33%
- Static Rule-Based 24.00%
- Vanilla (no monitor) 17.00%

Resolve rate, Moatless:

- Cadence 39.67%
- LivePlan 35.33%
- Fixed-Interval k=3 33.00%
- Fixed-Interval k=7 31.00%
- Fixed-Interval k=5 30.67%
- Static Rule-Based 25.67%
- Vanilla (no monitor) 24.00%

The abstract reports the gains over Vanilla as "+25.33%" and "+15.67%". The tables show these as percentage-point differences in resolve rate: $42.33 - 17.00 = 25.33$ and $39.67 - 24.00 = 15.67$. The abstract also states that "Cadence achieves the highest resolve rate among all evaluated monitors."

On Moatless, LivePlan is the strongest baseline at 35.33% and the paper reports its gap to Cadence as **not statistically significant**.

Token cost

Tokens per task:

- mini-swe-agent: Cadence 1.03M, Vanilla 1.42M, LivePlan 1.47M
- Moatless: Cadence 0.48M, Vanilla 0.32M, LivePlan 0.51M

Token leverage, resolved tasks per million tokens: mini-swe-agent Cadence 0.41 against Vanilla 0.12; Moatless Cadence 0.83 against Vanilla 0.75.

The direction of the cost result differs by agent: supervision was cheaper than no supervision on mini-swe-agent and more expensive on Moatless.

Note on the fixed-interval ordering

On mini-swe-agent the fixed-interval results are not monotone in k: k=5 (34.00%) is above k=3 (33.33%), and k=7 (26.33%) is below both. Shortening the interval is therefore not the lever the paper identifies; adapting it is.

Reati, responsabilita e risarcimenti connessi all'utilizzo dell'intelligenza artificiale (D.Lgs. 160/2026)

PRIMARY

Publisher Normattiva

Published 2026-09-15

Retrieved 2026-10-12

Cite key normattiva2026reati

URL <https://www.normattiva.it/eli/id/2026/09/15/26G00179/ORIGINAL>

Assessment. Italian official register entry. The decree's substantive provisions were read from Italian legal-press summaries (Lavoro si, diritto.it, aodv231); the date, number and Gazzetta Ufficiale reference are from the register's own ELI identifier and record. Prison-term ranges differ between an English-language secondary report and the Italian legal press, so no term is published in the edition.

Claims drawn from this source. CLAIM-2026-10-12-07, CLAIM-2026-10-12-08, CLAIM-2026-10-12-09

Identification

DECRETO LEGISLATIVO 9 settembre 2026, n. 160 - "Reati, responsabilita e risarcimenti connessi all'utilizzo dell'intelligenza artificiale". Published in Gazzetta Ufficiale n. 214 of 15 September 2026. **In force 30 September 2026.**

This is the instrument that carries Italy's criminal and civil liability provisions for AI, alongside rules on police use of AI systems. It is a different instrument from D.Lgs. 179/2026, which is the authorities-and-training decree.

Criminal provisions

A new **Art. 437-bis of the criminal code** punishes the omission of technical safety measures for high-risk AI systems - in design, training, production, placing on the market, or professional use - **or the omission of human oversight measures**. The offence is a delitto. Aggravated treatment applies where national security is engaged, and a separate paragraph covers a person outside the system's development who alters its functioning. Omission with gross negligence (colpa grave) attracts a reduced penalty.

Reported prison-term ranges diverge between sources and are therefore recorded here as disputed rather than stated: an English-language secondary report gives 1-5 years for danger to life or safety and 2-8 years where State security is engaged; the Italian legal press gives 2-8 years for concrete demonstrable danger, 2-6 years for alteration by an outsider, and 3-10 years in national-security cases.

Corporate liability

Art. 25-vicies is added to D.Lgs. 231/2001 as a predicate offence covering crimes committed using AI systems. For Art. 437-bis the entity faces a fine in quotas where it gained an undue advantage, and on conviction the disqualification sanctions of Art. 9(2)(b), (c), (d) and (e) of D.Lgs. 231/2001: exclusion from benefits and funding, a ban on contracting with the public administration, suspension or revocation of authorisations, and a **ban on advertising**. Disqualification from exercising the activity itself does not apply.

Civil liability

Three provisions matter to anyone who keeps machine-generated records:

1. **Production of records.** On the injured party's request the court can order the provider or user to produce evidence, registers (registri) and technical documentation about the algorithm. Refusal without justified reason allows the court to draw adverse inferences or treat the alleged facts as proven.
2. **Presumed causation.** A rebuttable presumption of the causal link operates in the victim's favour: where the victim shows non-conformity or negligent use plus a reasonable connection to the damage, the other party must prove the contrary.
3. **Conformity is not a shield.** A conformity certificate under the AI Act is not by itself sufficient to exclude the civil liability of the producer or user.

The injured party may also sue the liable party's insurer directly, and may bring the action where they reside or are domiciled.

Adeguamento della normativa nazionale al regolamento (UE) 2024-1689 - poteri delle Autorita nazionali (D.Lgs. 179/2026)

PRIMARY

Publisher	Normattiva
Published	2026-10-08
Retrieved	2026-10-12
Cite key	normattiva2026adeguamen
URL	https://www.normattiva.it/eli/id/2026/10/08/26G00200/ORIGINAL

Assessment. Italian official register entry, read directly. Title, number, Gazzetta Ufficiale issue and the entry-into-force date are quoted from the register record itself. This is the source that settles the date error in the English-language coverage, which attributed D.Lgs. 160/2026's September dates and liability provisions to this decree.

Claims drawn from this source. CLAIM-2026-10-12-10, CLAIM-2026-10-12-11, CLAIM-2026-10-12-12

Identification

DECRETO LEGISLATIVO 7 ottobre 2026, n. 179. Official subtitle, as printed on the register record:

"Adeguamento della normativa nazionale alle disposizioni del regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale, in materia di poteri delle Autorita nazionali e di utilizzo dell'intelligenza artificiale nella formazione."

Document code (26G00200). Published in **Gazzetta Ufficiale Serie Generale n. 234 of 8 October 2026**. The register states: "Entrata in vigore del provvedimento: 23/10/2026" and "Testo in vigore dal: 23-10-2026".

It is described in the Italian legal press as the **second** adequation decree to the AI Act.

Authorities

- **Art. 1(2)(a)** defines the "autorita nazionali per l'intelligenza artificiale" as the Agenzia per l'Italia digitale (**AgID**) and the Agenzia per la cybersicurezza nazionale (**ACN**).
- **Art. 1(2)(e)** defines the "autorita nazionale di notifica" as AgID.
- **Art. 1(2)(b)** includes ACN among the "autorita nazionali di vigilanza del mercato".
- The notes to Art. 1 quote **Law 132/2025, Art. 20**: AgID and ACN "sono designate quali Autorita nazionali per l'intelligenza artificiale"; AgID "e designata quale autorita

di notifica" and ACN "e designata quale autorità di vigilanza del mercato e punto di contatto unico".

- **Banca d'Italia, CONSOB and IVASS** are the market surveillance authorities for high-risk systems in their supervised sectors, under Art. 74(6) of the AI Act for Annex III point 5(b) and (c) systems used by financial intermediaries; Art. 17 of the decree gives them "poteri di vigilanza e ispettivi".
- The Garante applies the AI Regulation within its existing powers. ACN can run checks on real-world testing and conduct unannounced inspections. AgID and ACN jointly oversee the Italian regulatory sandbox.

Contrast worth recording for a German reader: the market surveillance seat that sits with ACN, a cybersecurity agency, in Italy sits with the Bundesnetzagentur in Germany under the KI-MIG.

Sanctions

Section V gives the authorities "il potere di irrogare le sanzioni e le altre misure amministrative di cui all'art. 99 dell'AI Act", within the EU-set ceilings, together with non-monetary measures such as orders to stop or remove a violation and public declarations.

Other provisions relevant to practitioners

- **Art. 41:** hiring, dismissal, performance-evaluation and disciplinary decisions may not rest solely on automated processing; the final decision "e in ogni caso riservata a una persona fisica, che eserciti un potere effettivo e autonomo". Decisions and dismissals taken solely on automated processing are void. Candidate search and screening are not treated as final hiring decisions.
- **Art. 42:** AI systems affecting work organisation, pace, methods or safety-relevant decisions must be included in the D.Lgs. 81/2008 risk assessment, with information and training on AI-specific risks.
- **Art. 48:** algorithmic use by employment centres is conditional on human supervision and transparency.
- **Art. 52:** amends Art. 98 of the Industrial Property Code to cover training data, model architectures and mathematical methods as trade secrets.
- **Arts. 43 and 47:** professional orders must add AI literacy to initial and continuing education, with six months from entry into force to adapt their rules.

Human oversight - Article 14

SECONDARY

Publisher	EU AI Act Explorer
Published	2026-08-02
Retrieved	2026-10-12
Cite key	euaiactexpl2026human
URL	https://artificialintelligenceact.eu/article/14/

Assessment. Secondary rendering of the Article's text, used for the provision and its application dates. Cite key reused from the vault's existing note for this URL.

Claims drawn from this source. CLAIM-2026-10-12-13, CLAIM-2026-10-12-14

Application dates, as printed at the foot of the page

"Comes into force 2 December 2027 (high-risk under Annex III)" / "2 August 2028 (high-risk under Annex I), according to Article 113(c)"

The provision

1. **Design duty.** High-risk AI systems shall be designed and developed so that they can be effectively overseen by natural persons during the period in which they are in use, including through appropriate human-machine interface tools.
2. **Purpose.** Oversight aims at preventing or minimising risks to health, safety or fundamental rights, including risks arising from foreseeable misuse or from interaction with other systems.
3. **Proportionality and delivery.** Measures shall be commensurate with the system's risks, level of autonomy and context of use, and may be built into the system by the provider where technically feasible, identified by the provider for the deployer to implement, or both.
4. **What the overseer must be enabled to do.** Providers must put the deployer in a position where the natural persons assigned to oversight can:
 - understand the system's relevant capacities and limitations and monitor its operation for anomalies, dysfunctions and unexpected performance;
 - remain aware of the possible tendency to automatically rely on or over-rely on the output (automation bias), in particular where the system informs or recommends decisions;
 - correctly interpret the output;
 - decide, in any particular situation, not to use the system or otherwise to disregard, override or reverse its output;
 - intervene in the operation of the system or interrupt it through a "'stop' button or a similar procedure" that allows it to come to "a halt in a safe state".

5. **Dual verification.** For Annex III point 1(a) systems, no action or decision may be taken on the basis of the identification unless it has been separately verified and confirmed by at least two natural persons with the necessary competence, training and authority - subject to the stated exception for certain law enforcement, migration, border control and asylum uses.

Reading relevant to this edition. Paragraph 1 requires oversight by *natural persons*. An LLM runtime monitor inspecting an agent trajectory is not a natural person, so it does not discharge Article 14 on its own; it may be part of the human-machine interface tooling that makes oversight effective under paragraphs 1 and 4.

Penalties - Article 99

SECONDARY

Publisher	EU AI Act Explorer
Published	2026
Retrieved	2026-10-12
Cite key	euaiactexpl2026penalties
URL	https://artificialintelligenceact.eu/article/99/

Assessment. Secondary rendering, used only for the fine ceilings and the allocation of penalty-setting to Member States. The page carries corrupted inserted glossary text in parts of paragraphs 3-5, 6 and 10; the figures were taken from the legible wording and are consistent with the Regulation as published. Cite key reused from the vault's existing note for this URL.

Claims drawn from this source. CLAIM-2026-10-12-15

Application date, as printed at the foot of the page

"Comes into force 2 August 2025, according to Article 113(b)"

Ceilings

- **Para. 1.** Member States lay down the rules on penalties and other enforcement measures for infringements by operators, which may include administrative fines, warnings and non-monetary measures. Penalties shall be effective, proportionate and dissuasive, taking into account the interests of SMEs including start-ups. Member States notify the rules to the Commission (para. 2).
- **Para. 3 - prohibited practices (Article 5).** Up to **EUR 35 000 000** or, for an undertaking, **7%** of total worldwide annual turnover for the preceding financial year, whichever is higher.
- **Para. 4 - operator and notified body obligations.** Up to **EUR 15 000 000** or, for an undertaking, **3%** of worldwide annual turnover, whichever is higher. This paragraph covers providers, authorised representatives, importers, distributors, deployers and notified bodies, **and the transparency obligations of Article 50.**
- **Para. 5 - incorrect, incomplete or misleading information to authorities.** Up to **EUR 7 500 000** or **1%** of turnover, whichever is higher.
- **Para. 6.** For SMEs including start-ups, the lower of the fixed amount and the percentage applies.
- **Para. 7.** Authorities weigh gravity and duration, prior fines, size and turnover, cooperation, intent or negligence, and mitigation.
- **Paras. 8-9.** Each Member State decides whether and how fines apply to its own public authorities and bodies; depending on the legal system, fines may be imposed by courts or other bodies.

Reading relevant to this edition. The EUR 35 000 000 / 7% figure that appears in coverage of the Italian decrees is the paragraph 3 ceiling for the Article 5 prohibitions.

Operator duties and Article 50 transparency sit one tier down, at paragraph 4: EUR 15 000 000 or 3%.

Regulation (EU) 2026-1744 (Digital Omnibus on AI)

PRIMARY

Publisher	EUR-Lex
Published	2026-07-27
Retrieved	2026-10-12
Cite key	eurlex2026regulatio
URL	https://eur-lex.europa.eu/eli/reg/2026/1744/oj

Assessment. Cite key reused from the vault's existing note for this URL. EUR-Lex renders as a metadata shell to the fetch proxy (recorded as D104), so the dates below are the settled ones carried by the pipeline's own run book rather than a fresh page read.

Claims drawn from this source. CLAIM-2026-10-12-13, CLAIM-2026-10-12-16

Identification

Regulation (EU) 2026/1744, the Digital Omnibus on AI. **In force 27 July 2026.**

Effect on the AI Act's application dates

- Annex III high-risk obligations moved to **2 December 2027**.
- Annex I high-risk obligations moved to **2 August 2028**.
- Member-state regulatory sandboxes moved to **2 August 2027**.

Unmoved, and settled: Article 50 transparency has applied since **2 August 2026**; the marking-of-generated-content obligation lands **2 December 2026**; Article 53 GPAI documentation has been in force since **August 2025**.

Note. EUR-Lex returns a metadata shell rather than the operative text to this pipeline's fetch path. The dates above are the run book's settled set, which is also what the Commission's own transparency guidance and the national implementing acts assume.

KI-MIG

KI-Marktueberwachungs- und Innovationsfoerderungsgesetz

PRIMARY

Publisher	Bundesamt fuer Justiz
Published	2026-07-29
Retrieved	2026-10-12
Cite key	bundesamtfu2026kimig
URL	https://www.gesetze-im-internet.de/ki-mig/

Assessment. Cite key reused from the vault's existing note for this URL. Carried for the German counterpart to Italy's ACN designation; no new reading was required for this edition.

Claims drawn from this source. CLAIM-2026-10-12-11

Identification

KI-Marktueberwachungs- und Innovationsfoerderungsgesetz (KI-MIG), the German act implementing the AI Act. **In force since 29 July 2026.**

Relevance to this edition

The KI-MIG designates the **Bundesnetzagentur (BNetzA)** as the market surveillance authority for the AI Act in Germany. Italy's D.Lgs. 179/2026, in force 23 October 2026, places the equivalent seat with the **ACN**, the national cybersecurity agency, with AgID as notifying authority.

Two member states, two institutional characters in the same statutory role. That is the practitioner-visible consequence: the authority that would eventually ask about a high-risk deployment is a telecoms and network regulator in one jurisdiction and a cybersecurity agency in the other.