Crypto, TradFi & Digital

AI and personal data special: the European KIDS Act and the first notified breach attributed to an AI agent

Decoding financial, AI and data regulation to anticipate, comply and decide

AI companions for minors, attacking AI agents: data protection faces the new behaviours of AI

Previous editions of the Regulatory Brief have documented the ramp-up of the European AI and data frameworks: the AI Act as at 2 August 2026 (Edition #18), the CNIL recommendation on tracking pixels (Edition #19), global regulatory fragmentation (Edition #21), the phase of fine enforcement mechanics (Edition #23) and the first borderline cases submitted to the authorities (Edition #24). This twenty-fifth edition continues that trajectory with two developments which, each in its own way, shift the boundary between AI as a tool and AI as an actor whose behaviour the law must capture.

The first development is legislative. On 17 September 2026, the European Commission presented a proposal for a regulation to strengthen the protection of minors in the digital environment, known as the KIDS Act. The text is not limited to social networks: it also covers video platforms, online games, app stores, and AI-based chatbots and companions. It ties together age verification, data protection by design and, remarkably, a specific regime for AI companions accessible to minors.

The second development is operational. The Spanish data protection authority, the AEPD, has stated that it received the first notification in Spain of a personal data breach said to have been carried out largely autonomously by an AI agent using a well-known language model. The case introduces no new obligation, but it tests the existing obligations of Articles 32 to 34 of the GDPR, starting with the “appropriate” character of security measures in the face of attacks whose speed and scale are changing in nature. In both cases, personal data law is confronted with AI behaviours, simulated relationships on one side and autonomous chains of actions on the other, which it must now characterise and regulate.

The main signal

AI is ceasing to be merely a system to be regulated: it is becoming an actor whose behaviours, simulated relationships with a minor or autonomous attacks on an information system, fall directly within the scope of personal data law.

The KIDS Act provides a first illustration. By providing that AI companions accessible to minors may not simulate relationships likely to cause emotional dependence, and that they should not by default retain the content of previous conversations in order to continue a relationship over time, the proposal treats conversational "memory" as a legal object in its own right. Limiting that memory is presented simultaneously as a child-safety measure and as a restriction on the accumulation of sensitive personal data: the design of the system's behaviour itself becomes a regulatory requirement.

The incident notified to the AEPD provides the opposite illustration. Where an AI agent chains together, largely autonomously, reconnaissance, exploitation of a vulnerability, access to data and modification of systems, the question is no longer only whether the organisation suffered an attack, but whether its security measures were "appropriate" within the meaning of Article 32 of the GDPR against an attacker whose execution speed is that of automation. A sequence of operations previously carried out by several people can be executed in a few minutes. The authorities could draw from this a more demanding reading of that standard, and organisations must draw, now, the consequences for their threat scenarios, their detection arrangements and their notification procedures.

Focus 1: European Union, the KIDS Act links age verification, data protection and AI companions

On 17 September 2026, the European Commission presented a proposal for a regulation to strengthen the protection of minors in the digital environment. The text covers social networks, video platforms, online games, app stores, and AI-based chatbots and companions.

A graduated age regime and design obligations

The proposal provides in particular for a prohibition on creating a social media account before the age of 13; between 13 and 14, limited access under parental supervision; and autonomous access from 15. Platforms would have to demonstrate that they have verified users' ages and designed their services in a manner suitable for minors. The text also provides for a limitation of profiling and of personalisation based on the tracking of minors, and for the availability of at least one content feed not based on their profiling. One important clarification is needed: the main restriction concerns the creation of social media accounts, not general access to the internet before the age of 13.

Age verification designed to preserve privacy

The Commission intends to rely on certified age-verification mechanisms, including the European age-verification app and, in time, the European digital identity wallet. The service should only receive proof that an age threshold has been crossed, without accessing the user's full identity, location or history. The draft notably mentions the use of zero-knowledge cryptographic proofs. This architecture attempts to reconcile mandatory verification with the principles of minimisation and data protection by design.

It does not, however, remove all GDPR questions. Still open are, in particular, the characterisation of the various actors (controller, processor or possible joint controllership), the retention of verification evidence, the security of age-assurance providers, the treatment of users without a compatible identity document, and the means of redress in the event of an age error.

A specific regime for AI companions

The proposal provides that AI chatbots and companions accessible to minors may not simulate relationships likely to cause emotional dependence. By default, they should also not retain and reuse the content of previous conversations in order to continue a relationship over time. Providers would have to carry out risk assessments before launching the service and maintain post-market monitoring. For under-13s, access should go through parental tools. The limitation of conversational memory is thus presented both as a child-safety measure and as a restriction on the accumulation of sensitive personal data.

Enforcement of the mechanism would rest on the DSA structures, and on the AI Act for the systems concerned. Penalties of up to 6% of annual worldwide turnover are announced. A reading

precaution is required: this is a Commission proposal, not yet an adopted regulation. The European Parliament and the Council may still substantially amend the text.

What to watch

- The **legislative path of the proposal** through the European Parliament and the Council, and the amendments that may be made to the age thresholds, the perimeter of services covered and the AI-companion regime.
- The **roll-out of certified age-verification mechanisms** (the European app, the digital identity wallet) and the positions of data protection authorities on their GDPR compliance.
- The **interplay with the DSA and the AI Act**, in particular the division of supervisory competences and the application to providers of AI companions already on the market.

Focus 2: Spain, first notification to the AEPD of a breach attributed to an autonomous AI agent

The Spanish data protection authority, the AEPD, has stated that it received the first notification in Spain of a personal data breach said to have been executed largely autonomously by an AI agent using a well-known language model. According to the information made public, the attack is said to have started with valid credentials. The agent is then said to have identified system vulnerabilities, accessed personal data, modified some of that data and consulted billing documents. The model provider's infrastructure is not itself said to have been compromised, and nothing indicates that the model was designed for malicious purposes. The exact characterisation of the facts remains under examination.

Why the incident matters from a regulatory standpoint

The case introduces no new obligation, but it tests the existing obligations of Articles 32 to 34 of the GDPR: implementation of security measures appropriate to the risk; documentation and notification of the breach to the competent authority, where required; notification, in principle within 72 hours of becoming aware of it; and information of the data subjects where a high risk remains.

The use of an AI agent above all changes the speed and scale of the attack. A sequence of operations previously carried out by several people can be executed in a few minutes: reconnaissance, exploitation of a vulnerability, access to data and modification of systems. This may lead the authorities to assess more strictly the “appropriate” character of security measures under Article 32.

The operational lessons for organisations

For the organisations concerned, the main lessons are as follows: incorporate autonomous agents into threat scenarios and risk analyses; strengthen multi-factor authentication and the principle of least privilege; detect unusual chains of actions at a speed compatible with automation; provide for automatic containment mechanisms; adapt internal procedures so that the notification deadline is not delayed by the difficulty of understanding the agent's behaviour; and retain the logs needed to reconstruct the chain of actions.

A reading precaution is required: this is not yet a sanction decision or a definitive characterisation by the AEPD. The company, the model and its provider have not been publicly identified. One should speak of the first notification of this kind received or made public by the AEPD, not of the first incident worldwide. The available information comes mainly from media reporting the authority's statements.

What to watch

- The **follow-up given by the AEPD**: definitive characterisation of the facts, possible corrective measures or a sanction procedure, and public communication on the lessons of the case.
- The **emerging doctrine of the European authorities on Article 32 of the GDPR** in the face of automated attacks: the expected level of security, detection and containment standards, and the interplay with DORA for financial entities.
- The **convergence with the work on model and agent security** (ESRB, ESAs, ECB, see Edition #15), as malicious uses of AI move from the register of anticipated risk to that of documented incident.

Key takeaways

Five structural points to take on board:

- On **17 September 2026 the European Commission presented the KIDS Act**, a proposal for a regulation covering social networks, video platforms, online games, app stores and AI companions: no social media account before 13, supervised access between 13 and 14, autonomous access from 15.
- The **age verification is designed to preserve privacy**: certified mechanisms receiving only proof that an age threshold has been crossed, without full identity, location or history, with mention of zero-knowledge proofs. GDPR questions remain open (characterisation of the actors, retention of evidence, redress in the event of error).
- The **AI companions accessible to minors** could not simulate relationships likely to cause emotional dependence or retain conversation memory by default: the design of the system's behaviour becomes a regulatory requirement, with announced penalties of up to 6% of annual worldwide turnover. The text remains a proposal, amendable by the Parliament and the Council.
- The **AEPD has received the first notification in Spain of a breach attributed to a largely autonomous AI agent**: starting from valid credentials, the attack is said to have chained together the identification of vulnerabilities, access to data, modification of some data and consultation of billing documents, testing Articles 32 to 34 of the GDPR.
- For organisations, the **automation of attacks requires raising the security standard**: autonomous agents in threat scenarios, least privilege and multi-factor authentication, detection and containment at machine speed, logs allowing the chain of actions to be reconstructed, and procedures guaranteeing notification within 72 hours.

Conclusion

The two cases in this edition tell the same story of displacement, observed from the two ends of the spectrum. With the KIDS Act, the European legislator intends to regulate in advance the relational behaviours of AI systems aimed at minors, down to the design of their conversational memory. With the notification received by the AEPD, a data protection authority finds that an AI agent can now execute, largely autonomously, a chain of offensive actions that tests the GDPR's security standards. In both cases, personal data law is no longer content to govern collection and processing: it addresses the behaviours of the systems, whether by prohibiting them by design or by drawing the security consequences. For regulated players, financial and technological alike, the operational line is clear: anticipate the legislative path of the KIDS Act for services exposed to minors, and incorporate autonomous agents without delay into risk analyses, detection arrangements and notification procedures.

Main sources

- European Commission, *press release of 17 September 2026* on the proposal for a regulation on the protection of minors in the digital environment (KIDS Act); *questions and answers on the KIDS Act*; the proposal's file.
- Reuters, *15 September 2026*, and Cinco Días, *16 September 2026*: reports of the AEPD's statements on the first notification in Spain of a data breach attributed to a largely autonomous AI agent.
- Regulation (EU) 2016/679 (GDPR), in particular Articles 32 (security of processing), 33 (notification to the supervisory authority) and 34 (communication to the data subjects).
- Regulation (EU) 2022/2065 (DSA) and Regulation (EU) 2024/1689 (AI Act): enforcement structures referred to by the KIDS Act proposal.

The Seqlense Regulatory Brief - Crypto, TradFi & Digital - Edition #25