Crypto, TradFi & Digital

Spécial IA et données personnelles : le KIDS Act européen et la première violation notifiée attribuée à un agent IA

Décrypter la réglementation financière, IA et données pour anticiper, se conformer et décider

Marchés Compagnons IA pour mineurs, agents IA attaquants : la protection des données face aux nouveaux comportements de l'IA

Les éditions précédentes du Regulatory Brief ont documenté la montée en application des cadres européens de l'IA et des données : l'AI Act au 2 août 2026 (Édition #18), la recommandation CNIL sur les pixels de suivi (Édition #19), la fragmentation réglementaire mondiale (Édition #21), la phase de mécanique fine de l'enforcement (Édition #23) et les premiers cas limites soumis aux autorités (Édition #24). Cette vingt-cinquième édition prolonge cette trajectoire avec deux actualités qui, chacune à sa manière, déplacent la frontière entre l'IA comme outil et l'IA comme acteur dont le droit doit saisir les comportements.

La première actualité est législative. Le 17 septembre 2026, la Commission européenne a présenté une proposition de règlement visant à renforcer la protection des mineurs dans l'environnement numérique, dite KIDS Act. Le texte ne se limite pas aux réseaux sociaux : il couvre également les plateformes vidéo, les jeux en ligne, les magasins d'applications ainsi que les chatbots et compagnons fondés sur l'IA. Il articule la vérification de l'âge, la protection des données dès la conception et, point remarquable, un régime particulier pour les compagnons IA accessibles aux mineurs.

La seconde actualité est opérationnelle. L'autorité espagnole de protection des données, l'AEPD, a indiqué avoir reçu la première notification en Espagne d'une violation de données personnelles, qui aurait été menée de manière largement autonome par un agent d'IA utilisant un modèle de langage connu. L'affaire n'introduit aucune obligation nouvelle, mais elle met à l'épreuve les obligations existantes des articles 32 à 34 du RGPD, à commencer par le caractère « approprié » des mesures de sécurité face à des attaques dont la vitesse et l'échelle changent de nature. Dans les deux cas, le

droit des données personnelles est confronté à des comportements de l'IA, relations simulées d'un côté, chaînes d'actions autonomes de l'autre, qu'il doit désormais qualifier et encadrer.

Le signal faible

L'IA cesse d'être seulement un système à réguler : elle devient un acteur dont les comportements, relations simulées avec un mineur ou attaques autonomes contre un système d'information entrent directement dans le champ du droit des données personnelles.

Le KIDS Act en fournit une première illustration. En prévoyant que les compagnons IA accessibles aux mineurs ne puissent pas simuler des relations susceptibles de provoquer une dépendance émotionnelle, et qu'ils ne conservent pas par défaut le contenu des conversations antérieures pour poursuivre une relation dans le temps, la proposition traite la « mémoire » conversationnelle comme un objet juridique à part entière. La limitation de cette mémoire est présentée simultanément comme une mesure de sécurité des mineurs et comme une restriction de l'accumulation de données personnelles sensibles : la conception du comportement du système devient elle-même une exigence réglementaire.

L'incident notifié à l'AEPD en donne l'illustration inverse. Lorsqu'un agent d'IA enchaîne de manière largement autonome reconnaissance, exploitation d'une vulnérabilité, accès aux données et modification des systèmes, la question n'est plus seulement de savoir si l'organisation a subi une attaque, mais si ses mesures de sécurité étaient « appropriées » au sens de l'article 32 du RGPD face à un attaquant dont la vitesse d'exécution est celle de l'automatisation. Une succession d'opérations auparavant menée par plusieurs intervenants peut être exécutée en quelques minutes. Les autorités pourraient en tirer une appréciation plus exigeante de ce standard, et les organisations doivent en tirer, dès maintenant, des conséquences sur leurs scénarios de menace, leurs dispositifs de détection et leurs procédures de notification.

Focus 1 : Union européenne, le KIDS Act relie vérification d'âge, protection des données et compagnons IA

Le 17 septembre 2026, la Commission européenne a présenté une proposition de règlement destinée à renforcer la protection des mineurs dans l'environnement numérique. Le texte couvre les réseaux sociaux, les plateformes vidéo, les jeux en ligne, les magasins d'applications ainsi que les chatbots et compagnons fondés sur l'IA.

Un régime d'âge gradué et des obligations de conception

La proposition prévoit notamment l'interdiction de créer un compte sur un réseau social avant 13 ans ; entre 13 et 14 ans, un accès limité sous supervision parentale ; et l'accès autonome à partir de 15 ans. Les plateformes devraient démontrer qu'elles ont vérifié l'âge des utilisateurs et conçu leurs services de manière adaptée aux mineurs. Le texte prévoit également une limitation du profilage et de la personnalisation fondée sur le suivi des mineurs, ainsi que la mise à disposition d'au moins un fil de contenus ne reposant pas sur leur profilage. Une précision importante s'impose : la restriction principale porte sur la création de comptes sur les réseaux sociaux, et non sur un accès général à Internet avant 13 ans.

Une vérification d'âge conçue pour préserver la vie privée

La Commission souhaite recourir à des dispositifs certifiés de vérification de l'âge, dont l'application européenne de vérification de l'âge puis, à terme, le portefeuille européen d'identité numérique. Le service devrait uniquement recevoir une preuve du franchissement d'un seuil d'âge, sans accéder à l'identité complète, à la localisation ou à l'historique de l'utilisateur. Le projet mentionne notamment l'utilisation de preuves cryptographiques à divulgation nulle de connaissance. Cette architecture tente de concilier une vérification obligatoire avec les principes de minimisation et de protection des données dès la conception.

Elle ne supprime cependant pas toutes les questions RGPD. Restent notamment ouvertes la qualification des différents acteurs (responsable de traitement, sous-traitant ou éventuelle responsabilité conjointe), la conservation des preuves de vérification, la sécurité des prestataires d'assurance d'âge, le traitement des utilisateurs ne disposant pas de document d'identité compatible, et les modalités de recours en cas d'erreur sur l'âge.

Un régime particulier pour les compagnons IA

La proposition prévoit que les chatbots et compagnons IA accessibles aux mineurs ne puissent pas simuler des relations susceptibles de provoquer une dépendance émotionnelle. Par défaut, ils ne devraient pas non plus conserver et réutiliser le contenu des conversations antérieures pour poursuivre une relation dans le temps. Les fournisseurs devraient procéder à des évaluations des risques avant le lancement du service et maintenir une surveillance après sa mise sur le marché.

Pour les moins de 13 ans, l'accès devrait passer par des outils parentaux. La limitation de la mémoire conversationnelle est ainsi présentée à la fois comme une mesure de sécurité des mineurs et comme une restriction de l'accumulation de données personnelles sensibles.

L'application du dispositif reposerait sur les structures du DSA, ainsi que sur l'AI Act pour les systèmes concernés. Des sanctions pouvant atteindre 6 % du chiffre d'affaires annuel mondial sont annoncées. Une précaution de lecture s'impose : il s'agit d'une proposition de la Commission, pas encore d'un règlement adopté. Le Parlement européen et le Conseil peuvent encore modifier substantiellement le texte.

Ce qu'il faut surveiller

- Le **parcours législatif de la proposition** au Parlement européen et au Conseil, et les modifications susceptibles d'être apportées aux seuils d'âge, au périmètre des services couverts et au régime des compagnons IA.
- Le **déploiement des dispositifs certifiés de vérification d'âge** (application européenne, portefeuille d'identité numérique) et les positions des autorités de protection des données sur leur conformité au RGPD.
- L'**articulation avec le DSA et l'AI Act**, notamment la répartition des compétences de contrôle et l'application aux fournisseurs de compagnons IA déjà présents sur le marché.

Focus 2 : Espagne, première notification à l'AEPD d'une violation attribuée à un agent IA autonome

L'autorité espagnole de protection des données, l'AEPD, a indiqué avoir reçu la première notification en Espagne d'une violation de données personnelles qui aurait été exécutée de manière largement autonome par un agent d'IA utilisant un modèle de langage connu. Selon les informations rendues publiques, l'attaque aurait commencé avec des identifiants valides. L'agent aurait ensuite identifié des vulnérabilités du système, accédé à des données personnelles, modifié certaines de ces données et consulté des documents de facturation. Les infrastructures du fournisseur du modèle ne seraient pas elles-mêmes compromises et rien n'indique que le modèle aurait été conçu à des fins malveillantes. La qualification exacte des faits reste en cours d'examen.

Pourquoi l'incident est réglementairement important

L'affaire n'introduit pas de nouvelle obligation, mais elle met à l'épreuve les obligations existantes des articles 32 à 34 du RGPD : mise en œuvre de mesures de sécurité adaptées au risque ; documentation et notification de la violation à l'autorité compétente, lorsqu'elle est requise ; notification, en principe dans les 72 heures suivant la prise de connaissance ; et information des personnes concernées lorsqu'un risque élevé subsiste.

L'utilisation d'un agent IA change surtout la vitesse et l'échelle de l'attaque. Une succession d'opérations auparavant menée par plusieurs intervenants peut être exécutée en quelques minutes : reconnaissance, exploitation d'une vulnérabilité, accès aux données et modification des systèmes. Cela peut conduire les autorités à apprécier plus sévèrement le caractère « approprié » des mesures de sécurité au titre de l'article 32.

Les enseignements opérationnels pour les organisations

Pour les organismes concernés, les principaux enseignements sont les suivants : intégrer les agents autonomes dans les scénarios de menace et les analyses de risques ; renforcer l'authentification multifacteur et le principe du moindre privilège ; détecter les enchaînements d'actions inhabituelles à une vitesse compatible avec l'automatisation ; prévoir des mécanismes de confinement automatique ; adapter les procédures internes afin que le délai de notification ne soit pas retardé par la difficulté à comprendre le comportement de l'agent ; et conserver les journaux nécessaires pour reconstituer la chaîne d'actions.

Une précaution de lecture s'impose : il ne s'agit pas encore d'une décision de sanction ni d'une qualification définitive de l'AEPD. L'entreprise, le modèle et son fournisseur n'ont pas été publiquement identifiés. Il faut parler de la première notification de ce type reçue ou rendue publique par l'AEPD, et non du premier incident mondial. Les informations disponibles proviennent principalement de médias rapportant les déclarations de l'autorité.

Ce qu'il faut surveiller

- Les **suites données par l'AEPD** : qualification définitive des faits, éventuelles mesures correctrices ou procédure de sanction, et communication publique sur les enseignements de l'affaire.
- La **doctrine émergente des autorités européennes sur l'article 32 du RGPD** face aux attaques automatisées : niveau de sécurité attendu, standards de détection et de confinement, articulation avec DORA pour les entités financières.
- La **convergence avec les travaux sur la sécurité des modèles et des agents** (ESRB, ESAs, BCE, cf. Édition #15), les usages malveillants de l'IA passant du registre du risque anticipé à celui de l'incident documenté.

À retenir

Cinq points structurants à intégrer :

- La Commission européenne a présenté le **17 septembre 2026 le KIDS Act**, proposition de règlement couvrant les réseaux sociaux, les plateformes vidéo, les jeux en ligne, les magasins d'applications et les compagnons IA : pas de compte sur un réseau social avant 13 ans, accès supervisé entre 13 et 14 ans, accès autonome à partir de 15 ans.
- La **vérification d'âge est conçue pour préserver la vie privée** : dispositifs certifiés ne recevant qu'une preuve de franchissement d'un seuil d'âge, sans identité complète, localisation ni historique, avec mention de preuves à divulgation nulle de connaissance. Des questions RGPD restent ouvertes (qualification des acteurs, conservation des preuves, recours en cas d'erreur).
- Les **compagnons IA accessibles aux mineurs** ne pourraient pas simuler des relations susceptibles de provoquer une dépendance émotionnelle ni conserver par défaut la mémoire des conversations : la conception du comportement du système devient une exigence réglementaire, avec des sanctions annoncées pouvant atteindre 6 % du chiffre d'affaires annuel mondial. Le texte reste une proposition, modifiable par le Parlement et le Conseil.
- L'**AEPD a reçu la première notification en Espagne d'une violation attribuée à un agent IA largement autonome** : partie d'identifiants valides, l'attaque aurait enchaîné identification de vulnérabilités, accès aux données, modification de certaines données et consultation de documents de facturation, en mettant à l'épreuve les articles 32 à 34 du RGPD.
- Pour les organisations, l'**automatisation des attaques impose d'élever le standard de sécurité** : agents autonomes dans les scénarios de menace, moindre privilège et

authentification multifacteur, détection et confinement à vitesse machine, journaux permettant de reconstituer la chaîne d'actions et procédures garantissant la notification dans les 72 heures.

Conclusion

Les deux dossiers de cette édition racontent le même déplacement, observé depuis deux extrémités du spectre. Avec le KIDS Act, le législateur européen entend encadrer par avance les comportements relationnels des systèmes d'IA à destination des mineurs, jusqu'à la conception de leur mémoire conversationnelle. Avec la notification reçue par l'AEPD, une autorité de protection des données constate qu'un agent d'IA peut désormais exécuter, de manière largement autonome, une chaîne d'actions offensives qui met à l'épreuve les standards de sécurité du RGPD. Dans les deux cas, le droit des données personnelles ne se contente plus de régir la collecte et le traitement : il s'attache aux comportements des systèmes, qu'il s'agisse de les proscrire par conception ou d'en tirer les conséquences en matière de sécurité. Pour les acteurs régulés, financiers comme technologiques, la ligne opérationnelle est claire : anticiper le parcours législatif du KIDS Act pour les services exposés aux mineurs, et intégrer sans attendre les agents autonomes dans les analyses de risques, les dispositifs de détection et les procédures de notification.

Sources principales

- Commission européenne, *communiqué du 17 septembre 2026* relatif à la proposition de règlement sur la protection des mineurs dans l'environnement numérique (KIDS Act) ; *questions-réponses sur le KIDS Act* ; dossier de la proposition.
- Reuters, *15 septembre 2026*, et Cinco Días, *16 septembre 2026* : informations rapportant les déclarations de l'AEPD sur la première notification en Espagne d'une violation de données attribuée à un agent IA largement autonome.
- Règlement (UE) 2016/679 (RGPD), notamment articles 32 (sécurité du traitement), 33 (notification à l'autorité de contrôle) et 34 (communication aux personnes concernées).
- Règlement (UE) 2022/2065 (DSA) et règlement (UE) 2024/1689 (AI Act) : structures d'application mentionnées par la proposition KIDS Act.

ou opérationnelle propre à chaque organisation, susceptible de tenir compte de sa taille, de son secteur, de ses pratiques et de son écosystème contractuel.

Les informations présentées reflètent une analyse générale des dynamiques réglementaires à la date de publication. Les positions des autorités européennes, américaines, asiatiques et nationales peuvent évoluer rapidement ; les textes en construction ne créent pas d'obligation immédiate et ne doivent pas être traités comme du droit applicable.

Malgré les soins apportés à la sélection et à la vérification des sources, aucune garantie n'est donnée quant à l'exactitude, l'exhaustivité ou l'actualité des informations. Les pratiques opérationnelles doivent être ajustées à la lumière des positions supervisorielles à venir, des décisions de sanction et de l'évolution des cadres juridiques.

Toute décision de mise en conformité relève de la seule responsabilité de l'organisation concernée et doit, le cas échéant, être prise avec l'appui de conseils juridiques et techniques qualifiés dans chaque juridiction concernée.