

Gli impegni tecnici assunti da Meta sono davvero efficaci e risolutivi per la tutela dei minori?

Sono misure che, sulla carta, vanno oltre tutto ciò che Meta aveva fatto finora. Ma quando si scende nel testo depositato in giudizio, l'efficacia di quasi ognuna di esse poggia su un presupposto fragile.

Fonte primaria: [Proposed] Consent Judgment ed Exhibit A (Settlement Agreement), *In re Social Media Adolescent Addiction / Personal Injury Products Liability Litigation*, MDL n. 3047, Case n. 4:23-cv-05448-YGR, U.S. District Court, Northern District of California, depositato il 26 agosto 2026.

Dove si trovano, nell'accordo, gli impegni tecnici di Meta?



Tutti gli obblighi tecnici assunti da Meta sono concentrati nella **Sezione II (Injunctive Provisions)**, dal § II.A al § II.G. È a quel testo che rinviano i riferimenti riportati in calce a ogni slide.

CIÒ CHE RESTA FUORI

L'analisi non prende in considerazione i profili economici e risarcitori (importi, riparto tra Stati, quantificazione del danno). Sono già ampiamente dibattuti e non dicono nulla sull'efficacia delle misure. Il fuoco è sul piano tecnico: che cosa Meta si è obbligata a fare, entro quando e con quale margine di inefficacia già previsto dal testo.

L'accordo è già operativo? E quando arriva davvero ciò che è stato annunciato?

L'orologio non è ancora partito. Ogni scadenza decorre dall'**Effective Date** (il primo giorno lavorativo successivo all'ingresso del consent judgment) e il giudice non lo ha approvato seduta stante.

- **4 mesi** Pause programmate durante l'uso e opzione di feed cronologico
- **6 mesi** Limiti orari, blocco notturno, primi strumenti di rilevazione degli under 13
- **12 mesi** Framework di accertamento dell'età e relative soglie di errore
- **24–36 mesi** Modello di rilevazione under 13 applicato all'intera base account
- **5 / 10 anni** Durata minima dei limiti di tempo e durata complessiva dell'accordo

Vale la pena notare l'ordine: la misura da cui dipendono tutte le altre (sapere chi è minorenne) è l'ultima ad arrivare.

Quanti minori possono legittimamente restare fuori dalle protezioni?

COME SI LEGGE LA TABELLA

L'accordo fissa una soglia massima di **falsi positivi**, cioè la percentuale di minori che il sistema può classificare come adulti. La scadenza indica **entro quando** ciascuna soglia va raggiunta, contata dall'Effective Date.

SCADENZA E METODO	13-15 ANNI	16-17 ANNI
Entro 12 mesi tecnologia commerciale di terzi	3%	10%
Entro 12 mesi tecnologia sviluppata da Meta	7%	14%
Entro 24 mesi tecnologia sviluppata da Meta	5%	10%

Fino al 14% dei sedicenni e diciassettenni può quindi essere trattato come adulto per il primo anno, in piena conformità con l'accordo.

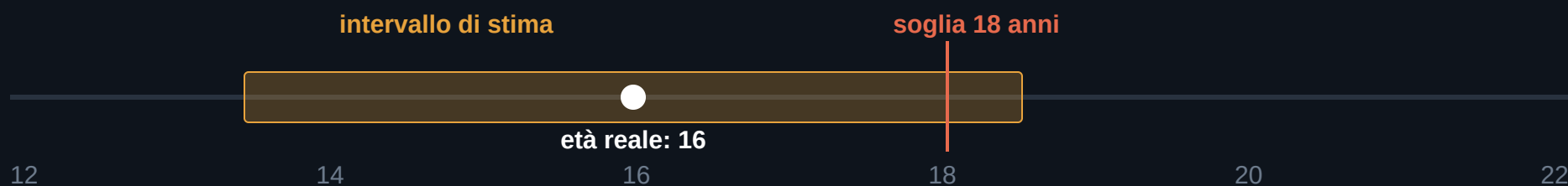
LA CRITICITÀ

Non si tratta di un margine d'errore residuo, ma di una quota di fallimento scritta nel testo e accettata dalle parti. Le soglie, inoltre, sono calcolate «excluding method circumvention»: **le tecniche di elusione non sono contemplate nel conteggio**. Il valore che ne risulta è quindi una proiezione statistica in condizioni controllate, non la misura di quanti minori riescano davvero a presentarsi come adulti.

La tecnologia è all'altezza del compito che l'accordo le affida?

QUANTO SBAGLIA, IN CONCRETO

Nella stima automatica dell'età di un adolescente lo scarto medio tra età stimata ed età reale **oscilla fra circa 2,5 e 3,1 anni**, a seconda del sistema e delle condizioni di ripresa. Il NIST rileva inoltre che nessun algoritmo si impone come standard e che i tassi di errore sono quasi sempre peggiori sui volti femminili.



Su un ragazzo che ha realmente 16 anni, quel margine restituisce una stima compresa fra i 13 e i 18 anni. Uno studio DSIT del 2026 lo dice in modo diretto: anche con 1–2 anni di scarto, un dodicenne può passare per quattordicenne e un sedicenne per maggiorenne.

IL PUNTO

La tecnologia è meno affidabile esattamente sulla soglia che l'accordo le chiede di presidiare.

Per proteggere i minori occorre prima identificare tutti gli altri?

OBIETTIVO

Individuare chi ha meno di 18 anni e chi ne ha meno di 13



CONSEGUENZA

Sottoporre a valutazione **l'intera base utenti**, adulti compresi



STRUMENTI

Documento d'identità, **biometria facciale** o inferenza comportamentale

L'ECCEZIONE NASCOSTA NEL TESTO

L'accordo prevede minimizzazione e cancellazione dei dati dopo la verifica, ma consente a Meta di **conservare i dati dei bambini sotto i 13 anni** per addestrare il proprio modello di rilevazione.

A fronte di questa facoltà, gli Attorney General hanno rinunciato ad agire in giudizio su quel punto specifico: una rinuncia limitata (che lascia teoricamente esistente la pretesa, ma impegna a non farla valere) riferita al **COPPA**, la legge federale statunitense del 1998 che disciplina la raccolta dei dati personali dei minori di 13 anni.

In sintesi: più dati di bambini trattenuti, meno possibilità di contestarlo. Il bilanciamento è esplicito, non implicito.

Che cosa accade se cambia il segnale di geolocalizzazione?

L'IMPEGNO ASSUNTO

Gli obblighi valgono per gli utenti **situati negli Stati firmatari**. Il consent judgment aggiunge che nulla di quanto pattuito costituisce standard di diligenza o precedente in Stati non aderenti o in qualsiasi giurisdizione estera.

LA CRITICITÀ

Una restrizione ancorata alla giurisdizione è, per costruzione, sensibile a tutto ciò che ne altera il rilevamento. Analizzando il divieto australiano di accesso ai social per gli under 16 (in vigore dal 10 dicembre 2025, ripreso nella slide successiva), il **Cato Institute**, centro studi statunitense di orientamento liberale-libertario con sede a Washington, ha collocato l'uso di VPN e di tecnologie che mascherano la posizione tra le tecniche per sottrarsi a restrizioni su base territoriale. È un rilievo riferito a quel caso, che qui non viene generalizzato.

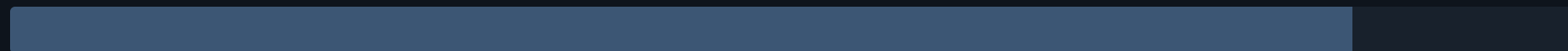
Va aggiunto che nessuna delle contromisure anti-elusione previste dall'accordo lavora sulla localizzazione: il § II.A.7 limita il numero di tentativi di verifica, non il perimetro geografico.

Esiste già un esperimento comparabile? Sì, e i numeri sono questi.

AUSTRALIA, DIVIETO DI ACCESSO AI SOCIAL PER GLI UNDER 16

In vigore dal 10 dicembre 2025, con sanzioni fino a 49,5 milioni di dollari australiani. Il regolatore ha pubblicato la prima valutazione ufficiale a tre mesi, condotta su oltre 4.100 famiglie.

Under 16 presenti su almeno una piattaforma vietata (prima) **85,9%**



Under 16 presenti su almeno una piattaforma vietata (dopo 3 mesi) **81,5%**



-4,4 punti percentuali in tre mesi il risultato della misura più severa mai adottata al mondo. Oltre metà dei ragazzi riferisce di non aver subito alcun controllo dell'età.

Il confronto è a maggior ragione: quello australiano è un **divieto di legge** assistito da sanzioni, mentre l'accordo statunitense è meno stringente e ha natura contrattuale.

A quali servizi si applicano davvero i limiti di tempo?

DENTRO IL PERIMETRO

Facebook e Instagram: feed, storie, reel, esplora

Limite di 2 ore al giorno

Blocco notturno 00:00–06:00

Notifiche silenziate in orario scolastico

FUORI DAL PERIMETRO

Messaggistica diretta: Instagram Direct e Messenger

WhatsApp, Threads, Meta AI, Horizon

Contenuti **long-form** da 22 minuti in su

Funzionalità IA e chatbot, esclusi per definizione

1,4 milioni

segnalazioni di **adescamento online** ricevute nel 2025 dalla CyberTipline, categoria che comprende grooming e sextortion

137 al giorno

le segnalazioni di **sextortion finanziaria** nel 2025, in crescita del 37% sull'anno precedente

528.000

le CyberTip inviate da Facebook, Instagram e Threads nel solo secondo trimestre 2025 per **interazioni inappropriate con minori**

Sono condotte che si consumano in conversazione privata, cioè nel canale che l'accordo lascia scoperto: la messaggistica diretta non conosce né limiti né blocco notturno né modalità scolastica.

Il feed algoritmico viene disattivato per i minori?

L'IMPEGNO ASSUNTO

Entro quattro mesi il minore potrà **scegliere** un feed cronologico non personalizzato, con una richiesta esplicita entro dieci giorni dall'identificazione dell'account e poi ogni novanta giorni.

IMPOSTAZIONE PREDEFINITA

Feed **algoritmico**, cioè la funzione al centro della contestazione



ALTERNATIVA

Feed **cronologico**, disponibile ma da attivare

LA CRITICITÀ

Resta un'opzione, e il precedente è agli atti del dibattito: il responsabile di Instagram ha ammesso di aver promosso strumenti di sicurezza per adolescenti senza rivelare i bassi tassi di adozione emersi nei test. Si torna così all'architettura opt-in che aveva già fallito una volta.

Chi tiene materialmente in piedi il sistema di protezione?

1. LA PROTEZIONE

Il minore non può allentarla da solo



2. LA CHIAVE

Solo un **genitore supervisore** può sbloccarla



3. IL CONTROLLO

Che sia adulto lo stabilisce il **modello di età di Meta**

22%

dei genitori europei utilizzava gli strumenti di controllo parentale esistenti (12% in Lituania, 14% in Repubblica Ceca)

47% / 35%

dei genitori statunitensi usa i controlli parentali su smartphone e su console; circa metà su tablet

L'accordo affida al genitore la funzione di presidio, ma non prevede alcun obiettivo di adozione: si limita a impegnare Meta a «continuare ad adottare misure volte a incoraggiare» l'iscrizione alla supervisione.

LA CRITICITÀ

Il cerchio si chiude male. La responsabilità viene spostata su un adulto che i dati mostrano largamente non attrezzato, e a custodire la via d'uscita è la stessa tecnologia fallibile della slide 05. Resta fermo che il genitore è e deve restare responsabile della supervisione: proprio per questo la sua alfabetizzazione sui rischi e sulle contromisure è un presupposto, non un dettaglio.

Meta deve ottenere un risultato o soltanto seguire una procedura?

CLAUSOLA	EFFETTO GIURIDICO
§ II.A.3.a	Presunzione di conformità se impiega un metodo di accertamento certificato di terzi
§ II.B.5.d	Sulle pause conserva la discrezionalità ed è presuntivamente conforme se documenta la valutazione
§§ II.E.2, II.F.2	Non è in violazione se un miglioramento ragionevolmente progettato non raggiunge l'efficacia attesa
§ II.A.6.b.iii	L'obbligo sul modello under 13 è condizionato alla permanenza in vigore dell'orientamento FTC

IL DATO CHE PESA DI PIÙ

In nessuna di queste clausole compare un **indicatore di efficacia misurabile**. Non esiste un solo obiettivo espresso in termini di riduzione del danno, di minori effettivamente intercettati o di esposizione evitata.

Ciò che viene verificato è la **conformità procedurale**: se il processo è stato svolto e documentato. Un'architettura che può risultare pienamente conforme anche restando priva di risultati.

Chi può concretamente far valere questi impegni?

LO STRUMENTO PREVISTO	IL LIMITE CHE LO ACCOMPAGNA
Auditor indipendente accesso a dati grezzi e documenti interni	È ingaggiato e retribuito da Meta
Cinque relazioni annuali con piani correttivi obbligatori	Sono confidenziali e non divulgabili senza preavviso a Meta
Legittimazione ad agire in caso di inadempimento	Solo Stati firmatari e Corte: nessuna azione privata per genitori e minori
Ingiunzione sulle dichiarazioni ingannevoli in tema di sicurezza	Serve la maggioranza dello State Committee , 30 giorni di preavviso, ed è esclusa dall’audit

Sulla carta è uno strumento più penetrante di qualunque report di trasparenza volontario. Nel testo, ogni leva è accompagnata dal contrappeso che ne riduce la portata.

E chi deve segnalare o chiedere la rimozione di un contenuto, che cosa trova?

QUALCOSA C'È, ED È CIRCOSCRITTO

L'accordo interviene sulle segnalazioni: strumenti in-app, risposta sulla decisione entro sei ore nel 90% dei casi, appello quando la revisione non ravvisa violazioni e un flusso semplificato per segnalare account di under 13.

COPERTO DALL'ACCORDO

Segnalazioni dei **minori titolari di account**

Risposta in 6 ore nel 90% dei casi, ma **solo in inglese o spagnolo**

Appello se la revisione non ravvisa violazioni

Flusso semplificato per segnalare under 13, aperto anche a chi non è utente

NON COPERTO

Chi agisce **per conto di altri**: genitore, tutore, difensore

Allegazione di prove, limiti di caratteri, formato dell'istanza

Categorie diverse dall'**elenco chiuso**: 6 su Facebook, 5 su Instagram

Punti di contatto diretti, escalation, revisione umana garantita

L'ASIMMETRIA CHE COLPISCE DI PIÙ

Uno standard di accessibilità nell'accordo c'è, ed è pure preciso: «Reasonably Accessible» significa raggiungibile in tre gesti, senza scorrere, in una posizione intuitiva. Ma viene imposto alle impostazioni del feed, non al percorso di segnalazione e reclamo.

Il problema degli strumenti di reclamo era già noto?

24.10.2025

I meccanismi di segnalazione di Facebook e Instagram **imporrebbero passaggi non necessari** e impiegherebbero *dark patterns*, con l'effetto di confondere e dissuadere chi segnala.

24.10.2025

I meccanismi di reclamo **non consentirebbero di fornire spiegazioni o prove a sostegno**, limitando l'efficacia dell'impugnazione.

29.04.2026

Lo strumento per segnalare account di under 13 sarebbe **poco intuitivo**, richiederebbe numerosi passaggi e non garantirebbe un *follow-up* adeguato.

PERCHÉ QUESTA SLIDE CONTA

Sono risultanze **preliminari** della Commissione europea, non un accertamento definitivo di violazione. Ma provengono da un'autorità pubblica, il procedimento è tuttora aperto e riguarda esattamente i meccanismi che l'accordo statunitense ha lasciato fuori.

Era il terreno più semplice e meno costoso su cui dimostrare un sacrificio tecnico reale: rendere davvero accessibile il canale attraverso cui un genitore, un tutore o un difensore chiede la rimozione di un contenuto che riguarda un minore.

Perché le misure più incisive non partono subito?

FASE I · ATTIVA

Blocco notturno **00:00–06:00**
Limite **2 ore** complessive



CONDIZIONE · FUORI DAL CONTROLLO DI META

Snap, TikTok e YouTube devono assumere obblighi equivalenti, sottoporsi ad audit indipendente per cinque anni e adottare requisiti di età non meno restrittivi



FASE II · SOSPESA

Blocco notturno **22:00–07:00**
Limite **60 minuti** per app

LA LETTURA CRITICA

Se il rischio per i minori è tale da giustificare un'azione promossa da 52 Attorney General, la protezione più forte non si subordina alla condotta dei concorrenti: si adotta e basta.

Condizionarla sposta il piano dall'*accountability* alla **parità competitiva**, e lascia la tutela più efficace nelle mani di soggetti che non hanno firmato nulla.

E la formazione di minori e genitori? Che cosa si è obbligata a fare Meta?

Nessun obbligo di educazione digitale a carico di Meta. Non un programma, non un budget, non un obiettivo di copertura, non un indicatore di efficacia. Ciò che esiste sono «risorse» ai genitori dopo che un contatto sospetto è già stato rilevato (§ II.F.4.b) e richieste periodiche di revisione delle impostazioni (§§ II.G.3–4): elementi di interfaccia, non di formazione.

MEDIAZIONE RESTRITTIVA

Imporre limiti, blocchi e divieti dall'esterno

È **tutto ciò che l'accordo prevede**: orari, tetti, default, filtri

Riduce il tempo d'uso, ma la letteratura la associa a un **impatto negativo sulle competenze digitali** dei ragazzi

MEDIAZIONE ATTIVA

Spiegare, discutere, accompagnare la scelta

È la strategia che le meta-analisi associano alla **riduzione degli esiti avversi**

L'89% dei ragazzi dichiara di sentirsi a proprio agio nel parlarne con un genitore: il canale è aperto

L'INCOERENZA INTERNA

L'accordo poggia interamente sulla colonna di sinistra e non finanzia in alcun modo quella di destra, che è però l'unica che la ricerca collega a un miglioramento degli esiti.

Che cosa **non** è questo accordo.

QUATTRO LIMITI SCRITTI NEL TESTO

- Non è un accertamento di responsabilità: Meta nega ogni addebito e l'accordo non costituisce ammissione.
- Non costituisce *standard of care* né precedente in Stati non firmatari o in giurisdizioni estere.
- Non attribuisce alcun diritto a terzi: nessuna azione privata.
- Non è permanente: i limiti di tempo durano al minimo cinque anni, l'accordo dieci.

SINTESI CONCLUSIONALE

Che Meta abbia finalmente assunto impegni più stringenti di quelli precedenti è un fatto, e va riconosciuto. Ma il perimetro è ritagliato, la durata è limitata, la verifica è procedurale e non di risultato, la parte più incisiva è appesa alla condotta dei concorrenti e il pilastro educativo semplicemente non c'è.

Rispetto a ciò che l'accordo ha permesso di evitare, il sacrificio tecnico non è proporzionato. Non è un cambio d'epoca: è un punto di partenza da verificare, misura per misura, nei prossimi mesi.

E per gli utenti europei? Quando, e se, tutto questo arriverà.

IL DATO DI PARTENZA

Nulla nell'accordo obbliga Meta a estendere una sola di queste misure fuori dagli Stati firmatari, e **nessuna data esiste per l'Europa**. Gli obblighi sono territoriali e il consent judgment esclude espressamente qualunque effetto in giurisdizioni estere.

- **autunno 2024** «Teen Accounts» introdotti su Instagram
- **aprile 2025** Estensione a Facebook e Messenger in USA, Regno Unito, Canada e Australia
- **settembre 2025** Estensione al resto del mondo, a circa dodici mesi dal primo mercato

CHE COSA SE NE PUÒ RICAVERE

Il precedente indica un rilascio a ondate: prima i mercati anglofoni, poi il resto. Se lo schema si ripetesse, le **funzioni** potrebbero arrivare in Europa nell'arco di dodici-ventiquattro mesi, anche perché l'infrastruttura di accertamento dell'età è costosa e, una volta costruita, tende a essere riutilizzata ovunque.

Ciò che **non** viaggia con esse sono le garanzie: soglie di errore, auditor indipendente e legittimazione ad agire restano contrattuali e limitati agli Stati firmatari. Un segnale in questo senso è già nel testo, dove la risposta entro sei ore è prevista solo per l'inglese e lo spagnolo.

L'esito plausibile è quindi il più scomodo: un utente europeo potrebbe trovarsi davanti la stessa interfaccia, **senza nessuno dei presidi che la rendono verificabile**.

Settlement Agreement § II.A.1; Consent Judgment § IX.C. Cronologia dei rilasci: annunci Meta di aprile e settembre 2025. Il regolatore dei media sudcoreano ha chiesto che alcune delle misure siano applicate ai giovani utenti di tutto il mondo e non solo in mercati specifici.

Avv. Nicolò Bastaroli · Digital Reputation Law · Reputation Channel