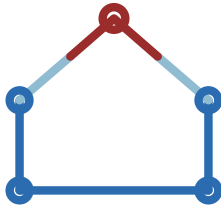


Principal Brief // two-minute read

HOUSEHOLD & STAFF

CHILD PROTECTION

AGGREGATION & IDENTITY RESOLUTION



# Your Family Never Posted Together

*It does not matter. Their separate, ordinary posts can be joined into one picture of your household, and the platform already maintains part of that picture: a model that estimates who is close to whom in real life, refreshed weekly across trillions of connections.*

## WHAT META DISCLOSED

In March 2026 Meta's engineers described the system behind Facebook's friend bubbles. Two models estimate how close you feel to each connection in real life. Meta says one runs weekly across trillions of connections.

## WHY IT REACHES YOU

You did not publish a file on your family. Your mother posted a newborn photo, a school a roster, a friend a birthday tag. Joined, and read by a system that answers in plain language, they can name your children and place your house.

## THREE THINGS TO TAKE FROM THIS

1

### The household, not the account

Locking down your profile does not bound this. The link that identifies you sits on a relative's page, a school site, or a property record you do not control.

2

### Your children are in it without posting

A child with no account is still named through adult accounts and school records. A 2018 Children's Commissioner report estimated parents would post about 1,300 images of a child before age 13.

3

### The dangerous answer is the true one

AI rules police false and banned output. A correct answer about which child is yours passes them, and is what an extortionist needs.

## DO THIS

### ☐ YOU, THIS MONTH

Ask your footprint what an adversary would ask: who are my children, where do I live now, whose voice is public.

### ☐ YOU AND YOUR HOUSEHOLD

Agree a verification phrase and a call-back rule for urgent requests about money, travel, or safety. Assume the caller can imitate a familiar voice.

### ☐ YOU, WITH YOUR FAMILY OFFICE

Have the household audited as one connected graph, and ask relatives and schools to remove identifiable images of your minors.

## NOT THE STORY YOU MAY HAVE HEARD

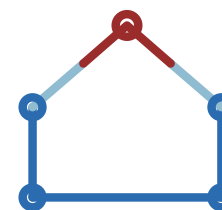
A viral video showing Meta AI naming a creator's children is circulating as proof of a lookup tool. It is one unverified account, and parts may be the model guessing. The documented capability is enough.

## HOW SURE WE ARE

High confidence that platforms model real-world family closeness and that general models infer personal details from ordinary posts. Moderate confidence on how far chat products will expose those inferences.

CB-247

# The family graph is the file: what gets assembled, not what gets posted



*"The platform already models who is close to whom in real life. AI turns that model and your scattered posts into an answer."*

*A friend-recommendation cue is the wrapper. The asset underneath is a maintained estimate of which people are close to each other offline, refreshed weekly across trillions of connections. Once a plain-language interface sits on that estimate, the question is no longer what you posted. It is what your household adds up to.*

**BLUF.** In March 2026 Meta's engineering organization published the architecture behind Facebook's friend bubbles. Two machine-learning models estimate how close a viewer feels to each of their connections, one of them trained on a survey that asks users whether they feel close to a specific person in real life and refreshed on a regular cadence. Meta states that this model runs weekly inference over trillions of person-to-person connections. The visible feature is a small row of faces on a video. The durable asset is a continuously maintained model of offline family and friendship structure.

Public argument about AI and family privacy has settled on two questions: whether a model says something false, and whether a child is old enough to be on the platform. Neither question reaches this exposure. The harm here comes from accurate output about people who never consented, produced by assembly, the joining of scattered records into one identified person and household. The governing frame for this brief: the risk sits in the assembly, not in any input.

## THE SIGNAL // WHAT IS ACTUALLY DOCUMENTED

Four public records, none of them speculative, establish that household assembly is a built and operating capability rather than a forecast.

META ENGINEERING • 18 MAR 2026

The survey-based closeness model draws on social-graph features (mutual friends, connection strength, interaction patterns) and user attributes (behavioral and demographic signals including user-provided location, number of friends, and number of posts shared). Meta describes the purpose as building a more complete picture of real-world relationships, and states the model runs weekly inference over trillions of person-to-person connections across Facebook friends. Training labels come from asking users whether they feel close to a named connection offline, with proxy questions such as how often two people communicate.

FTC 6(B) STAFF REPORT • SEP 2024

The Commission's study of nine social media and video streaming services found mass collection about users and, in some cases, non-users. In his accompanying statement, Commissioner Bedoya cited staff findings that companies used consumer data to infer income, relationships, whether a person is divorced, and whether they have children, and noted that most services claimed to have no child users while their own systems could estimate ages above thirteen.

Off-the-shelf large language models inferred personal attributes including location, income, and sex from ordinary public comments, reaching 84.6 percent top-1 accuracy for GPT-4 across attributes and up to 85 percent top-1 and 95 percent top-3 overall, at roughly one hundredth the cost and one two-hundred-fortieth of the time a human analyst required. Follow-on work at NeurIPS 2024 extended the finding to benign images that contain no people, and ICLR 2025 work found commercial anonymization tools misaligned with what current models can undo.

FBI IC3 PSA I-120525-PSA · 05 DEC 2025

The Bureau warned that criminals alter photographs taken from social media and other public sites to serve as fabricated proof-of-life imagery in virtual kidnapping for ransom schemes, contacting families by text with claims of abduction and immediate ransom demands. An earlier IC3 advisory (03 Dec 2024) documented generative AI used to scale tailored social engineering and to clone the voices of relatives.

**SO WHAT** The capability that turns your family's scattered posts into an answer is documented by the platform itself, by the regulator, and by peer review. Nobody needs to prove a viral video to establish it.

THE MECHANISM // FOUR FUNCTIONS BEHIND ONE INTERFACE

What changed is not the volume of data but the collapse of four previously separate skills into a single interface that requires none of them. Search engines already made documents retrievable, brokers already assembled structured records, and investigators already knew how to pivot between sources. Each of those steps cost time, money, or expertise. A conversational system performs all four at conversational speed for a user with no investigative training.

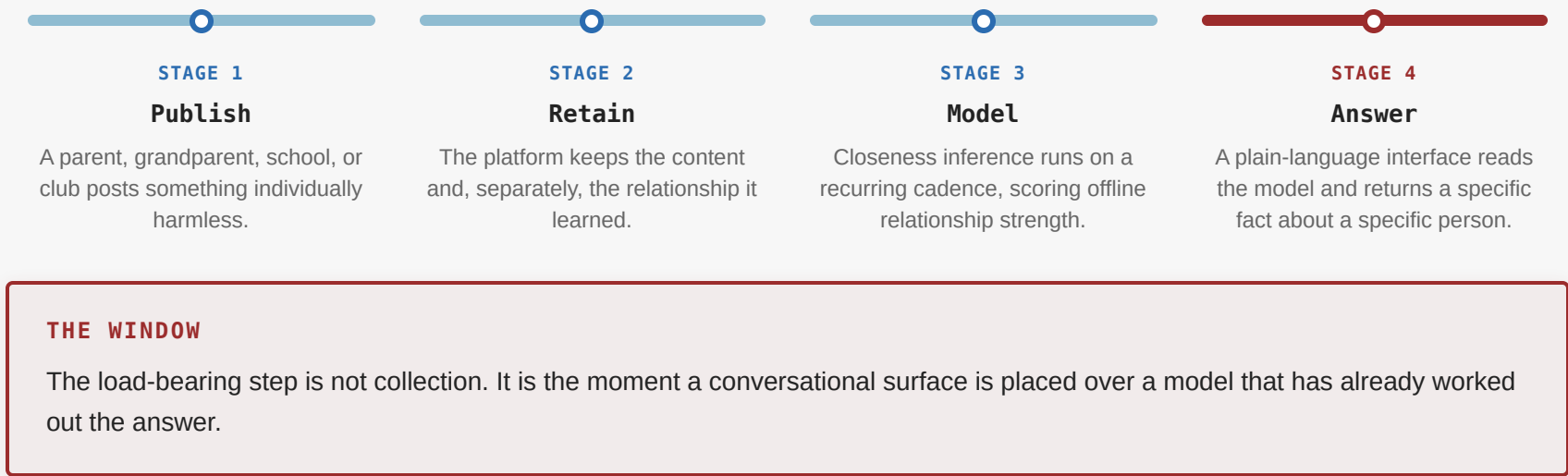
FIGURE 1 // THE FOUR FUNCTIONS, AND WHAT EACH DOES TO A HOUSEHOLD

|          |                                                                                                                         |
|----------|-------------------------------------------------------------------------------------------------------------------------|
| RETRIEVE | Finds relevant posts, images, records, and mentions, including old ones nobody remembers publishing.                    |
| JOIN     | Matches records that refer to the same person, connecting a child on one account to a parent or grandparent on another. |
| INFER    | Estimates what no single source states: current residence, routine, relationship, age, likely school, plausible wealth. |
| EXPLAIN  | Returns a coherent answer in ordinary language to a user who needs no investigative training.                           |

The graph is not the dossier. It is the join layer that lets scattered records become one, which is why a contact map matters more than its contents: it is what gives every other record household meaning.

The join step is therefore the one that matters, and it can survive deletion. Removing a photograph removes a copy. It need not remove the association a system learned from that photograph, the copy on a relative's account, the cached or indexed version, or the household link a broker already sold. A source can be deleted while a conclusion drawn from it persists in the system that derived it.

FIGURE 2 // FROM AN ORDINARY POST TO AN ACTIONABLE ANSWER



WHY THIS REACHES A PRINCIPAL // THE HOUSEHOLD IS THE PERIMETER

A footprint program scoped to the principal will miss the paths most likely to reconstruct the principal. The identifying link routinely sits on a node the principal does not control: a mother's page, a school's event listing, a club roster, a property record, a friend who tags reliably. Suppressing the principal's own records leaves those bridges intact, and a bridge is worth more to an adversary than a dozen isolated mentions, because it is what converts two separate clusters into one household.

Children carry this exposure without contributing to it. A child may hold no account, publish nothing, and consent to nothing, and still be named, aged, located, and associated through adult accounts and institutional records. Research for the Children's Commissioner for England found that roughly 1,300 photographs and videos of the average child are posted online before the child turns 13, overwhelmingly by adults. The privacy subject and the data contributor are different people, which is precisely the case that consent frameworks handle worst.

The privacy subject and the data contributor are different people.

For a principal, the operative threat is not embarrassment. It is that assembly produces exactly the inputs a targeted approach needs: which child belongs to whom, who is trusted enough to impersonate, which residence is current, when the household is away, which relative will answer an urgent call, and what school, vehicle, pet, or family phrase makes a pretext credible. The FBI's proof-of-life advisory closes the loop, because the family photographs that establish who is who are the same photographs used to fabricate evidence that one of them has been taken.

► THE PRACTICAL TEST

The test is not whether a system holds a record on you. It is whether a stranger with no skill can ask a plain-language question and receive a *true* answer about a child who has no account.

SO WHAT If your protective work searches for you rather than for your household, it is auditing the wrong perimeter, and the person who has to change that scope is whoever writes the engagement.

Platform, broker, and adversary have incompatible motives and converge on the same technical primitive. That convergence is why a feature built in good faith and an attack built in bad faith draw on the same underlying work.

FIGURE 3 // DIFFERENT INCENTIVES, IDENTICAL PRIMITIVE

| ACTOR     | INCENTIVE                   | WHAT ASSEMBLY GIVES THEM                                                                                                                                             |
|-----------|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| PLATFORM  | Relevance and engagement    | A maintained, refreshed model of who matters to whom offline, feeding ranking, recommendation, search, and increasingly conversational answers.                      |
| BROKER    | A sellable household record | Tolerance for messy inputs. A perfect matching row is no longer required, because a model reconciles variants, ranks candidates, and infers the missing link.        |
| ADVERSARY | A sufficient answer         | Targeting context and pretext material: the child's name, the trusted relative, the current address, the voice sample, the family detail that makes a call credible. |

THE GOVERNANCE GAP // WHAT THE LARGEST CHILD-SAFETY SETTLEMENT DOES NOT TOUCH

The regulatory instruments now in force against this platform address attention and age, not assembly. On 26 August 2026 a bipartisan coalition of 51 state and territorial attorneys general announced a proposed settlement resolving claims that Meta designed Instagram and Facebook to drive compulsive use by children and teens and misled the public about the resulting harms. Meta will pay at least 12.1 billion dollars over ten years, rising to as much as 17.1 billion only if Snap, TikTok and YouTube reach comparable settlements of their own. The parties filed a proposed consent judgment the same day in the Northern District of California; it remains subject to the court's approval. The negotiated remedies are default daily time limits for users under 18, a block on overnight use that only a parent can lift, enhanced age assurance, additional parental supervision tools, and an independent compliance auditor. Not one of them constrains cross-account assembly of a household, retention of a derived relationship after the source post is deleted, or the return of an accurate answer about a minor. The largest state consumer protection settlement since tobacco was negotiated entirely around how long a child may scroll and how old the platform believes they are.

Product safety rules have the same blind spot from the other direction. Model governance concentrates on output that is false, harmful, or disallowed, and hallucination is what attracts scrutiny because it is what embarrasses the vendor. A correct statement of which child belongs to a principal, or which of two addresses is the current one, is none of those things. It is accurate, it is responsive, and by the vendor's own quality metric it is a good answer.

SO WHAT

Do not wait for regulation to close this one. The largest child-safety settlement in American history was resolved on screen time and age checks, and left household assembly entirely untouched.

THE STEELMAN // THE HONEST CASE AGAINST THIS READING

THE BEST CASE FOR THE OTHER SIDE

- + The closeness model is documented as a ranking and retrieval input for a video feature. Meta does not describe it as a lookup service, and nothing in the engineering post says a closeness score is queryable through a chat interface.
- + Meta has published privacy-infrastructure work on classifying assets in what it calls the AI-native era, operates under an existing consent decree, and has now accepted court-supervised obligations. These are real constraints, not nothing.
- + The viral case that prompted this analysis is a single first-person account with no independent reproduction. Some of what it appears to show is better explained by content that was never fully deleted, by a relative's surviving copy, or by the model confabulating a plausible location.
- + Inference accuracy measured on Reddit corpora does not transfer cleanly to a specific product with its own retrieval scope, safety layer, and refusal behavior.

Concede all four. The claim in this brief is not that Meta ships a household lookup, and any argument that depends on the viral video is not worth making. The claim is narrower and harder to dismiss: the components exist, are first-party documented, are trained specifically on offline relationship strength, and are refreshed weekly at population scale, while the governance now binding the same company addresses time limits and age assurance. What increasingly separates a ranking feature from an answerable household query is product integration and permissioning, not the absence of the underlying analytic capabilities, and integration decisions are not the thing a principal gets advance notice of.

ASSESSMENT

The social graph was already a map of relationships. What is new is that relationship modeling, personal-data inference, retrieval, and conversational explanation can now be combined within one product environment. That lowers the minimum skill required to conduct household intelligence to the level of typing a question.

FIGURE 4 // ANALYTIC CONFIDENCE BY CLAIM

|                                                                                                                                                                         |                                            |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|
| Platforms maintain and refresh models of offline family closeness, not merely on-platform connections.                                                                  | STRONG / FIRST-PARTY DOCUMENTED            |
| General-purpose models infer personal attributes from ordinary text and images at accuracy and cost that put household assembly within reach of an unskilled adversary. | STRONG / PEER-REVIEWED                     |
| A child with no account and no disclosure can still be named, aged, and located through adult accounts and institutional records.                                       | STRONG / STRUCTURAL                        |
| Conversational interfaces will increasingly surface household-level inferences that previously stayed internal to ranking, ads, and fraud systems.                      | MODERATE / PRODUCT-DEPENDENT               |
| Deleting a source post does not remove the association a system derived from it.                                                                                        | MODERATE / NOT EXTERNALLY VERIFIABLE       |
| The circulating Meta AI video demonstrates a working household lookup capability.                                                                                       | REJECTED / SINGLE UNREPRODUCED ACCOUNT     |
| Locking down your own accounts bounds your family's exposure.                                                                                                           | REJECTED / THE GRAPH IS NOT ACCOUNT-SCOPED |

POSTURE

**FOR PRINCIPALS**

- Ask answer-shaped questions of your own footprint rather than record-shaped ones. Not "what records exist" but "who are my children, which relatives post them, where do I live now, whose voice is publicly available."
- Establish a family verification phrase and an out-of-band call-back rule for any urgent request involving money, travel, medical decisions, or safety. Assume the caller knows correct names, relationships, and recent events, and can reproduce a familiar voice.
- Treat relatives, household staff, and your children's institutions as inside your perimeter, and ask them directly to remove identifiable images of minors. Your own deletions do not reach their pages.

**FOR FAMILY OFFICES**

- Scope exposure work to the household as a connected graph, not a roster of separately protected individuals. Prioritize bridge artifacts: a tagged photograph, an obituary, a wedding page, a team roster, a property link.
- Retest after moves, births, deaths, weddings, school changes, liquidity events, litigation, and new public roles. Those events cause relatives and institutions to publish, which is when new bridges appear.
- Distinguish copy removal from association removal in your reporting. They age differently, and a report that counts only records deleted will overstate what has actually been reduced.

**FOR SECURITY & LEGAL FUNCTIONS**

- Test by asking, not by enumerating. An audit that lists records held will not surface what a plain-language query returns about the principal's children, residence, or routine.
- Build the pretext inventory the adversary would build, then verify that your household verification protocol defeats it. Voice samples and named relationships are the two inputs worth tracking.
- This is intelligence analysis, not legal advice. Erasure rights, minor-specific state provisions, and platform retention of derived attributes vary considerably; route deletion strategy and any demand letters through counsel.

**CONFIDENCE** High confidence that major platforms build, maintain, and regularly refresh models of offline relationship strength; that commercial systems infer household composition including relationship status and whether a person has children; and that general-purpose models infer personal attributes from ordinary public text and images at low cost and high accuracy. Moderate confidence that conversational products will increasingly expose household-level inferences that previously remained internal to ranking, advertising, and fraud systems, since the components and incentives are present but product access and safeguards will vary by surface and jurisdiction. Rejected: that the circulating Meta AI video demonstrates a household lookup capability. It is a single unreproduced first-person account, ObscureIQ has not replicated the behavior or established which sources were accessed, and confabulation is a live alternative explanation. The argument in this brief is constructed so that it does not depend on that account.

**METHOD** Built from first-party engineering documentation, a federal regulatory study and an accompanying commissioner statement, peer-reviewed inference research, a federal law enforcement advisory, and the state attorneys general announcements of a proposed multistate settlement. Cross-read against the ObscureIQ doctrine that adversaries assemble answers while defenders buy systems, and against the coverage-versus-capability distinction in CB-243. The viral creator case was deliberately demoted from evidence to illustration and is not load-bearing; the creator and her children are not named here, since republishing the query that assembled a household would reproduce the harm the brief describes.

**SOURCES** Meta Engineering, *Friend Bubbles: Enhancing Social Discovery on Facebook Reels*, 18 Mar 2026. FTC, *A Look Behind the Screens: Examining the Data Practices of Social Media and Video Streaming Services*, Sep 2024, with the Statement of Commissioner Alvaro M. Bedoya, 19 Sep 2024. Staab, Vero, Balunovic and Vechev, *Beyond Memorization: Violating Privacy via Inference with Large Language Models*, ICLR 2024; related ETH Zurich work on image-based attribute inference (NeurIPS 2024) and on anonymization (ICLR 2025). FBI IC3 PSA I-120525-PSA, 05 Dec 2025, and IC3 PSA of 03 Dec 2024. Announcements of the proposed Meta multistate settlement, 26 Aug 2026 (Offices of the Attorneys General of California, Connecticut, New Jersey and the District of Columbia; *People of the State of California v. Meta Platforms, Inc.*, No. 4:23-cv-05448, N.D. Cal.). Children's Commissioner for England, *Who Knows What About Me?*, 2018. Origin illustration credited to the Instagram creator @kontheinside, 04 Sep 2026.

*Scope note: the Meta architecture details, the FTC findings, and the ICLR and NeurIPS accuracy figures are confirmed against primary sources. The settlement figures and remedy terms are taken from the state attorneys general announcements of 26 Aug 2026 and should be confirmed against the filed consent judgment before external reproduction. As of publication the proposed consent judgment awaits entry; one participating office described it as approved on the day of announcement, which conflicts with the others and with subsequent reporting, so approval is treated here as pending. Meta counts 52 attorneys general where the coalition announcements count 51; Texas settled separately and Florida did not join. The 1,300-image figure originates in a*

2018 Children's Commissioner report and is cited as an order-of-magnitude indicator, not a current measurement. The viral creator account is attributed, not asserted. This is intelligence analysis, not legal advice.

OBSCUREIQ INTERNAL ANALYSIS // ARCHIVE COMM-247-FAMILY-GRAPH-HOUSEHOLD-ASSEMBLY-2026Q3 // XREF [CB-180](#) [CB-240](#) [CB-243](#) // ACCESS PUBLIC · OPEN TO ALL READERS AT RELEASE