

A/C Priv

Reviewers Approval (please add yourself or others if not listed)

- ☐ Legal -
 - ☐ Rachel Lieber
 - ☐ Einat [REDACTED]
 - ☐ Product Legal - Mel [REDACTED] ?
- ☐ Public Affairs
 - ☐ Stephanie [REDACTED]
- ☐ FB Product
 - ☐ Growth - Amee [REDACTED] and/or Robert d'Apice
 - ☐ FI - Robert [REDACTED]
- ☐ IG Product
 - ☐ IGWB - Rishi [REDACTED]
 - ☐
- ☐ XI
 - ☐ Gabriel [REDACTED]
- ☐ Policy
 - ☐ USPP - Don [REDACTED]
 - ☐ Safety - Ravi Sinha
 - ☐ IG - Swithin [REDACTED]
 - ☐ FB - Aleah [REDACTED]

CONFIDENTIAL
CONFIDENTIAL

META3047MDL-247-00000530
METANMAG-095-00000530
PLTF EX-02215_0001

PLTF EX
PLTF EX-02215

Decision: Next Steps in Implementing Adult - Minor Recommendations Policy

Oct 2, 2023

Issue for Decision

Determine our approach to adult - minor recommendations on actor-level recommendations surfaces like Facebook’s People You May Know (PYMK) and Instagram’s Accounts You May Follow (AYMF).

Context

Following the first Wall Street Journal [article](#), the Child Safety Task Force has been working to tighten our existing child safety protections. This included work to ensure that Risky Adults (i.e those with specific Malicious Child Safety Actor - MaCSA scores) and minors are not recommended to each other. Implementation of the final segment of this work (recommending teens to Risky Adults on FB’s PYMK) has been [blocked](#) by product based on Voice concerns. We have made extensive external statements that this filtering is in place, and based on these and user safety concerns,

PRIV	Legal	PRIV
PRIV		

Other segments of our existing adult-minor Recommendations Policy, which preclude recommendations of adult-minor pairs without mutual connections and with large age gaps, have also never been implemented due to technical feasibility concerns.

Following publication of the first article, we received several letters from members of Congress regarding the Wall Street Journal’s allegations, including alleged deficiencies in our approach to child safety on our platforms, and scope of the Task Force’s work. U.S Public Policy is concerned that the Risky Adult mitigation will not be compelling as we continue to engage with external stakeholders going forward on this topic (which seems likely, given that the Wall Street Journal informed us of plans to public a second article in its child safety series in short order), and Policy is advocating that we not only complete the remaining Risky Adult filters, but also launch filters targeting adults and minors with large age gaps and/or no mutual connections. We would like to resolve this issue prior to responding to a [letter](#) from the Senate Judiciary Committee on this topic.

Key Decision Factors

1. **Safety.** Reducing the likelihood that Risky Adults discover and make connections with minors.
2. **External Statements & Pressures.** External regulatory and PR pressure regarding child safety and our own external statements regarding adult <> minor recommendations.
3. **Proportionality & Voice Impact.** Impact to voice and the proportionality of the impact to our users.
4. **Execution Complexity.** Complexity of consistent implementation across apps and surfaces and potential for future compliance gaps.

Options

	Option	Safety	External Statements & Pressures	Proportionality & Voice Impact	Execution Complexity
1	Do not recommend any minors <> adults in PYMK or AYMf unless: (1) they are within a two-year age gap and (2) the adult does not have a Risky MaCSA score and (3) they have at least one mutual connection. This option allows us a clear external narrative that we do not recommend adults and minors to each other, with the exception of the same two-year age gap and mutual connection logic as our <u>private messaging restrictions</u>. XFN Support:	Safest Option: By not recommending adult - minor pairs, we are not reliant on signals to detect risky actors, offering the maximum protection to minors. MaCSA filtering for close-in-age pairs would also reduce risk.	Broader than Statements: Records of external communication on this issue indicate our messaging focused on Risky Adults, <u>not</u> All Adults, so this would go beyond previous statements. In closed door settings, we may have made statements beyond Risky Adult filtering; we have not found documentation of this. In 2021, Adam Mosseri defended adult <> minor recommendations to Congress. (details below) External Environment: This option would enable our strongest response to external stakeholders. Increased public and regulatory scrutiny and limited external trust make recommendations of adults and minors with large age gaps difficult to defend. However, this option may inadvertently imply that there is more risk inherent in adult <> minor connections that we think	User value The primary reason people in general (and teens specifically) use Facebook is to connect with Family and Friends (add key stat). Teens, in particular, often use Facebook as a way to connect with family specifically, who are often older (add key stat). Today, 31% of new teen friendships are with 18-24 and 33% are with 25+. PYMK drives 80-90% of all Facebook friendships. This suggests that removing these recommendations may result in teens losing a significant source of value they get from Facebook. Note that this change wouldn't affect public figures in PYMK but would in AYMf (IG).	

			to be the case.		
2	Do not recommend adults <> minors in PYMK or AYMF if (1) no mutual connections; (2) the adult is 35+ and minor is under 16; or (3) the adult has a risky MaCSA score. This option follows the <u>private messaging precedent</u> but allows for a wider age gap. It also prevents the recommendation of Risky Adults, a measure we have made extensive external statements about. <i>XFN Support: Public Affairs ("balances voice and safety")</i>	Enhanced Protection in Case of MaCSA Gaps: This approach (our current, unimplemented internal policy) enhances child safety protections while carving out some space for younger adult <> minor discovery. Any gaps in risky adult detection would be reduced by the additional filters.	Broader than Statements: Records of external communication on this issue indicate our messaging focused on Risky Adults, <u>not</u> All Adults, so this would go beyond previous statements. We understand that in closed door settings we have discussed this policy; we have not found specific documentation of this. External Environment: This option is defensible from a comms perspective and more adequately balances voice and safety considerations. This would not enable our strongest response to external stakeholders.	Similar to option 1, this option may result in teens losing a significant source of value they get from Facebook and Instagram.	
3	Do not recommend Risky Adults and minors to each other in PYMK and AYMF. Ensure FB completes the work to close the one remaining gap in Risky Adult (.3 MaCSA score) <> Minor recommendations filtering: surfacing teens to Risky Adults in PYMK. IG has completed this already on AYMF. <i>XFN Support:</i>	Dependent on MaCSA Detection: Gaps in MaCSA detection could introduce the possibility of minors connecting with risky adults who have not engaged in activity included in the MaCSA criteria.	Consistent with Statements: In our Comms channels, messaging to advertisers, and Congressional statements we have made extensive statements indicating that bi-directional Risky Adult <> minor filtering is in place on both apps; this option is consistent with our existing commitments. This option is a minimum bar to complete the Task Force's work - the gap in this filtering <u>required</u> us to tell the European Commission that the TF's work was not yet complete, but we committed to finishing it. External Environment: This option will not allow a sufficient response to	Most Narrowly Targeted Option: While teens likely get significant value from connections with adults on Facebook, the policies to remove risky individuals will not affect the vast majority of benign relationships. Note: This is yellow due to concerns about the NFX policy (which categorizes someone as MaCSA risky if they have [REDACTED] from minors within [REDACTED]; these reports include topics like a deleted friend request, and unfriend, and XX). If we remove that criterion or share examples to demonstrate that NFX is capturing true risky behavior would turn	

		external stakeholders, who Additionally, gaps in MaCSA's recall may lead to Risky Adults <> teen recommendations, generating further external pressure if regulators or reporters find risky recommendations - we have seen gaps in risky Group filtering generate further press inquiries.	this to green.	
--	--	--	----------------	--

Additional Policy and Impact Analysis

External Statements

Outside Counsel	PRIV	Legal	PRIV
PRIV			

1. Outside Counsel
- | |
|------|
| PRIV |
|------|

PRIV

PRIV

2. Comms:

a. Blogs Indicate Risky Adult <> Minor Filtering on FB and IG:

- i. Several Newsroom and blog posts include references to bi-directional Risky Adult <> minor filtering on both IG and FB. For example:
 1. Blog on CyberTips made to NCMEC: "We don't show teen accounts in Instagram Explore, Reels or 'Accounts Suggested For You' or in 'People You May Know' recommendations on Facebook to these adults, and we don't show accounts belonging to these adults to teens in these places"
 2. Blog on safety protections: "In addition to our existing measures, we're now testing ways to protect teens from messaging suspicious adults they aren't connected to, and we won't show them in teens' People You May Know recommendations. A "suspicious" account is one that belongs to an adult that may have recently been blocked or reported by a young person, for example." This announcement was picked up in press.
- b. **PYMK language was included in background information shared with Jeff Horwitz for the initial WSJ story:** "Specific restrictions we apply to potentially suspicious adults may include: We don't show teen accounts in Instagram Explore, Reels or 'Accounts Suggested For You' or in 'People You May Know' recommendations on Facebook to these adults, and we don't show accounts belonging to these adults to teens in these places."

3. Advertisers:

- a. **Commitments on Risk Adults on Instagram:** Following the WSJ article, our messaging focused on Risky Adults on Instagram, in particular our measures to prevent recommendations to or from minors, new measures to prevent Risky Adult <> Risky Adult recommendations, and Risky Adult to other Adult Recommendations.

Internal Policy

Our external Recommendations Policy does not address adult-minor recommendations; however, our internal version states that we do not recommend:

- *Accounts or entities that are suspected of engaging in inappropriate interactions with children, the sexualization of children, engaging in exploitative intimate imagery and sextortion, or solicitation of CSAM, nude imagery or sexualized imagery of real or fictional children, or offline sexual encounters with children.*
- *We will not recommend the accounts of minors under the age of **16 to adults over the age of 35**, and vice-versa.*
 - *Implementation: not possible today, should be in consultation with PYMK surface team and Child Safety.*
- *We will not recommend minors to non-minors, and vice-versa, in cases where the accounts share **no mutual friends**.*
 - *Implementation: not possible today, should be in consultation with PYMK surface team and Child Safety."*

A 2022 internal review noted that these components of the policy were not implemented due to technical feasibility.

Appendix

Proposed responses to Congressional letters for each option:

Option 1:

Another significant focus of our work to help keep teens safe is by putting in place safeguards, restrictions and limitations with regard to their interactions with adults on our platform. We already have steps in place to make it more difficult for adults to find and follow teens. For example, we restrict people over 19 years old from sending private messages to teens who do not follow them. And when teens now join Instagram, we prohibit people who they do not follow from tagging or mentioning them, or including their content in Reels Remixes or Guides.

We are also launching additional protections to make it more difficult for adults to find and follow teens. Going forward, adults' and minors' accounts will not be recommended to each other on Accounts You May Follow. Teens who are 18 or 19 years-old will still be able to find friends and classmates within a two-year age gap if they have a mutual connection and we do not have any indication that adult is exhibiting suspicious behavior.

For adults exhibiting potentially suspicious behavior, we go much further. We have developed technology to identify accounts demonstrating potentially suspicious behavior—for example, accounts belonging to adults that may have recently been blocked or reported by teens—and help prevent them from engaging with teens. We look at numerous signals to find and label these accounts and apply a number of restrictions to them, up to and including disabling the accounts.

For example, we work to avoid showing young people's accounts to these adults in Explore, Reels, or "Accounts Suggested For You" and work to avoid showing accounts belonging to these adults to teens in these places. These adults also will not be able to see comments from young people on anyone's posts, nor will young people be able to see their comments on any public posts. In addition, we use technology to automatically disable accounts exhibiting a high level of suspicious signals. We also notify teens if these adults attempt to follow or message them.

And we are not stopping there. To help keep predators from attempting to use our service to connect with one another, we are taking additional steps to help ensure accounts that belong to potentially suspicious adults are not recommended to *anyone* (including teens and now also adults) in “Accounts You May Follow.” Once this change comes into effect, the following restrictions will be applicable on Instagram:

- These accounts will not be recommended to each other in Search unless people search the exact username.
- These accounts will not be recommended to each other in “Accounts You May Follow.”
- If these accounts land on another account belonging to a potentially suspicious adult, they will not be able to follow the account.
- These accounts will not be shown comments by other potentially suspicious adults on public posts, or on their own accounts.
- These accounts will not be shown to one another in surfaces where we recommend content to people, like Explore and Reels.
- These accounts will not be visible to each other when scrolling through the list of people who have liked a post or when looking at an account’s Followers or Following list.

Option 2:

Another significant focus of our work to help keep teens safe is by putting in place safeguards, restrictions and limitations with regard to their interactions with adults on our platform. We already have steps in place to make it more difficult for adults to find and follow teens. For example, we restrict people over 19 years old from sending private messages to teens who do not follow them. And when teens now join Instagram, we prohibit people who they do not follow from tagging or mentioning them, or including their content in Reels Remixes or Guides.

We are also launching additional protections to make it more difficult for adults to find and follow teens. Going forward, adults’ over 35 and minors under 16 years-old will not be recommended to each other on Accounts You May Follow. We will also prohibit adults and minors without at least one mutual friend from being recommended to each other.

For adults exhibiting potentially suspicious behavior, we go much further. We have developed technology to identify accounts demonstrating potentially suspicious behavior—for example, accounts belonging to adults that may have recently been blocked or reported by teens—and help prevent them from engaging with teens. We look at numerous signals to find and label these accounts and apply a number of restrictions to them, up to and including disabling the accounts.

For example, we work to avoid showing young people’s accounts to these adults in Explore, Reels, or “Accounts Suggested For You” and work to avoid showing accounts belonging to these adults to teens in these places. These adults also will not be able to see comments from young people on anyone’s posts, nor will young people be able to see their comments on any public posts. In addition, we use technology to automatically disable accounts exhibiting a high level of suspicious signals. We also notify teens if these adults attempt to follow or message them.

And we are not stopping there. To help keep predators from attempting to use our service to connect with one another, we are taking additional steps to help ensure accounts that belong to potentially suspicious adults are not recommended to *anyone* (including teens and now also adults) in “Accounts You May Follow.” Once this change comes into effect, the following restrictions will be applicable on Instagram:

- These accounts will not be recommended to each other in Search unless people search the exact username.
- These accounts will not be recommended to each other in “Accounts You May Follow.”
- If these accounts land on another account belonging to a potentially suspicious adult, they will not be able to follow the account.
- These accounts will not be shown comments by other potentially suspicious adults on public posts, or on their own accounts.
- These accounts will not be shown to one another in surfaces where we recommend content to people, like Explore and Reels.
- These accounts will not be visible to each other when scrolling through the list of people who have liked a post or when looking at an account’s Followers or Following list.

Option 3:

Another significant focus of our work to help keep teens safe is by putting in place safeguards, restrictions and limitations with regard to their interactions with adults on our platform. We already have steps in place to make it more difficult for adults to find and follow teens. For example, we restrict people over 19 years old from sending private messages to teens who do not follow them. And when teens now join Instagram, we prohibit people who they do not follow from tagging or mentioning them, or including their content in Reels Remixes or Guides.

For adults exhibiting potentially suspicious behavior, we go much further. We have developed technology to identify accounts demonstrating potentially suspicious behavior—for example, accounts belonging to adults that may have recently been blocked or reported by teens—and help prevent them from engaging with teens. We look at numerous signals to find and label these accounts and apply a number of restrictions to them, up to and including disabling the accounts.

For example, we work to avoid showing young people’s accounts to these adults in Explore, Reels, or “Accounts Suggested For You” and work to avoid showing accounts belonging to these adults to teens in these places. These adults also will not be able to see comments from young people on anyone’s posts, nor will young people be able to see their comments on any public posts. In addition, we use technology to automatically disable accounts exhibiting a high level of suspicious signals. We also notify teens if these adults attempt to follow or message them.

And we are not stopping there. To help keep predators from attempting to use our service to connect with one another, we are taking additional steps to help ensure accounts that belong to potentially suspicious adults are not recommended to *anyone* (including teens and now also adults) in “Accounts You May Follow.” Once this change comes into effect, the following restrictions will be applicable on Instagram:

- These accounts will not be recommended to each other in Search unless people search the exact username.

- These accounts will not be recommended to each other in “Accounts You May Follow.”
- If these accounts land on another account belonging to a potentially suspicious adult, they will not be able to follow the account.
- These accounts will not be shown comments by other potentially suspicious adults on public posts, or on their own accounts.
- These accounts will not be shown to one another in surfaces where we recommend content to people, like Explore and Reels.
- These accounts will not be visible to each other when scrolling through the list of people who have liked a post or when looking at an account’s Followers or Following list.

Status of Minor <> Adult Filtering on High Priority Recommendations Surfaces

Status of Filtering Across Rec Surfaces (For WSJ Purposes)

App	Surface	Are we filtering?			
		1. MaSCA → Anyone	2. MaSCA → MaSCA	3. MaSCA → Minor	4. Minor → MaSCA
Facebook	a. People You May Know (PYMK)	Yes (>[REDACTED])	Yes (>[REDACTED])	Yes (>[REDACTED])	[REDACTED] [Blocked] Testing >[REDACTED] threshold currently (not launched)
	b. In-Feed Recommendations (IFR)	Yes (>[REDACTED])	Yes (>[REDACTED])	Yes (>[REDACTED])	Yes (>[REDACTED])
	c. Reels	Yes (>[REDACTED])	Yes (>[REDACTED])	Yes (>[REDACTED])	Yes (>[REDACTED])
	d. Watch	Yes (>[REDACTED])	Yes (>[REDACTED])	Yes (>[REDACTED])	Yes (>[REDACTED])
	e. Search (unconnected)	Yes (>[REDACTED])	Yes (>[REDACTED])	Yes (>[REDACTED])	Yes (>[REDACTED])
	e. Groups You Should Join (GYSJ)	Yes (1. [REDACTED] with MACSA [REDACTED] AND [REDACTED] with MACSA [REDACTED] 2. [REDACTED] of members > [REDACTED] 3. [REDACTED] of members > [REDACTED] 4. [REDACTED] [REDACTED])	N/A	N/A	N/A
	g. Pages/Events (PYML/EYSJ)	Yes [REDACTED] admin/hosts is MaSCA >[REDACTED])	N/A	N/A	N/A

App	Surface	Are we filtering?			
		1. MaSCA → Anyone	2. MaSCA → MaSCA	3. MaSCA → Minor	4. Minor → MaSCA
Instagram	a. AYMf	Yes (> [REDACTED])	Yes (> [REDACTED])	Yes (> [REDACTED])	Yes (> [REDACTED])
	b. Reels	No. Pending user-facing transparency experience for account level soft action. WIP over Q4	Yes (> [REDACTED])	Yes (> [REDACTED])	Yes (> [REDACTED])
	c. Explore		Yes (> [REDACTED])	Yes (> [REDACTED])	Yes (> [REDACTED])
	d. Search (via Hashtags)		Yes (> [REDACTED])	Yes (> [REDACTED])	Yes (> [REDACTED])

Precedent / Parallels on Other Apps

Recommendations:

- All Adults on Horizon Worlds:** On Horizon Worlds, minors and all adults are blocked from surfacing eachothers' PYMK tabs and adults <> minors are prevented from inviting to parties unless they are already mutual followers.
- Risky Adults in VR:** In VR, we only prevent adults <> minors pairs from surfacing in PYMK based on suspicious adult score (similar to MaCSA).
- No Age on WhatsApp:** On WhatsApp, we do not have accurate age collection, so this issue is not relevant.
- No Minors in FB Creators To Follow:** Minors are excluded from being recommended in or receiving recommendations from the Creators To Follow Notifs.
- Upcoming PYMK Unit on Messenger:**

Private Messaging (note: these are not Recommendations, so are not a direct parallel):

1. **Two-Year Age Gap on IGD:** *On IGD, we permit people aged 18 and 19 to send private messages to their peers within a two-year age gap to allow for connections between classmates and friends, for example, a 19 year old may message teens aged 17 and older. Otherwise, we restrict people over 19 years old from sending private messages to teens who don't follow them.*
2. **Mutual Friends on Messenger:** *On Messenger, we limit adults from sending private messages to teens if they are not Friends or share a mutual Friend.*
 - a. *Messenger is also exploring permitting the two-year age gap, as IG does, by the end of H2.*

References

- Letter from SJC
- Communication with advertisers post-WSJ report on IG
- Outside Counsel PRIV
- XFN escalation document on MaCSA filter on PYMK
- Internal Recommendations Policy