

FILED UNDER SEAL

EXHIBIT 2

Graphic Content Q1 2019 CIPCO Roadmap

This roadmap covers investments in Child Exploitation Imagery (CEI), Non-Consensual Intimate Imagery (NCII), Nudity (N&P), Non-Consensual Sexual Touching (NCST), Sexual Solicitation (SS) and Prostitution (P).

tl;dr - In Q1, the Graphic Content team will focus on:

- 1. CEI: Improve banking, re-define Tier 1 CEI, prep for Messenger CEI scanning, build ops dashboard**
 - a. Bank more content and drive more deletes per content banked
 - b. Re-define Tier 1 CEI targeting real world harm and build a classifier prototype potentially with NCMEC dataset
 - c. Prepare infrastructure to begin proactively scanning Messenger for Tier 1 CEI
 - d. Build CIPCO OPS KPI dashboard to facilitate weekly syncs, deep dives and capacity utilization
 - e. [P2] Improve our FB/IG video classifier
- 2. NCII: Fix backlog, bank NCII context, expand proactive recall, test NFX/FRX UX, build ops dashboard**
 - a. Reduce FB and IG reactive backlogs and incoming volumes
 - b. Improve classifier recall on FB and IG
 - c. Bank context with simple objects to unlock quality audits and retro-action (Q2)
 - d. Build CIPCO OPS KPI dashboard to facilitate weekly syncs, deep dives and capacity utilization
 - e. Launch and measure NCII FRX/NFX test to inform Q2 UX improvements
- 3. N&P: Reduce CO workforce spend, stabilize top-level metrics, goal on prevalence, improve GT recall**
 - a. Complete 50% reduction in CO workforce spend for N&P (\$16.8M/year savings)
 - b. Resume goaling on prevalence at <11.5 bps; Take GT recall from 58% → 80%
 - c. Maintain FP Reach < 80k/day, GT accuracy > 93.7%
 - d. Take end-to-end infra/actioning ownership and work with CO to stabilize prevalence

Mission

Problem Mission: Build a safe, inclusive, and supportive community by:

- Minimizing distribution of policy violating Nudity, NCII and CEI.
- Prioritizing efforts to quickly detect, report and action NCII and CEI that is most likely to prevent real world harm.

- Developing experiences that lessen the impact of NCII and CEI on targets and the community.
- Working with external partners to reduce NCII and CEI experiences on and off our platform and get buy-in on our approach.

Maturity Goals (H1 and Q1 Summary)

- **Reference:** [[HYPERLINK "https://fb.quip.com/MNJ2ApgAlbjY"](https://fb.quip.com/MNJ2ApgAlbjY) \h], [[HYPERLINK "https://fb.quip.com/jSqwAFmL4Sgf"](https://fb.quip.com/jSqwAFmL4Sgf) \l "SNIACAqWcNA" \h]
- **Legend:** 0 = RISK, 1 = BASIC, 2 = MEASURED, 3= OPERATIONAL

	A	B	C	D	E	F	G	H	I	J
1	PROBLEM AREA	H1 PRI	FACE BOOK			INST AGR AM			MES SEN GER	
2			H1 STAR T	Q1 GOA L	H1 GOA L	H1 STAR T	Q1 GOA L	H1 GOA L	H1 STAR T	Q1 GOA L
3	CEI	H	2	2	2	1	1	1	0	0
4	NCII	H	2	2	2 → 3	1	1	1	0	0
5	NCST	H	0	0 → 1	0 → 1	0	0	0	0	0
6	N&P	M	3	3	3	3	3	3	-	-
7	Sexual Solicitation (S)	M	1	1	1	0	0 → 1	0 → 1	-	-
8	Prostitution (P)	M	1	1	1	0	0 → 1	0 → 1	-	-

Q4 Lookback/Summary

Q4 Roadmap: [[HYPERLINK "https://fb.quip.com/tgJsAZcQHj91"](https://fb.quip.com/tgJsAZcQHj91) \h]

Q4 THEMES

1. **CEI: Step function improvement to automation, CO throughput and banking with goals to** enable A1-B1 realtime CEI detection on FB and IG, automate detection and actioning of viral memes, reduce the # of steps required to bank content/enable retroaction, and invest in our NCMEC partnership.
2. **NCII: Step function improvement to automation, CO throughput and target support with goals to** increase review/CO throughput without additional CO staff, increase action rate on proactive jobs and N&P reactive escalation jobs, ship first ever NCII video classifier, deliver recommendation (with design mocks) for improving NCII target/reporting experience and define NCII tiers.
3. **N&P: Hold the line while unifying the workforce and focusing on CO reporting efficiency gains with goals to** hit top-level metric goals for vNFX, FP Reach, and GT Accuracy, merge 100% of the proactive/reactive workforce and implement reporting efficiency gains. *[P2] Publish a CO measured prevalence metric and define a process to RCA prevalence spikes/changes.*

Q4 Highlights

1. Maturity

- a. **CEI.** We went from 1 → 2 on FB and 0 → 1 on IG.
- b. **NCII.** We hit our goals of 1 → 2 on FB and 0 → 1 on IG.
- c. **Sexual Solicitation & Prostitution.** We hit our goals of 0 → 1 on FB for SS & P.

2. CEI

- a. **Proactive Coverage.** We now have real-time proactive detection across FB photos, IG photos, Messenger Kids photos, and FB Video. We achieved our goal of proactively detecting >5k NCMEC reportable contents/week with >20% A1/B1 rate.
- b. **MMS Matching.** We improved MMS on IG by shipping PDQ client-side support (3-5x increase on selected IG surfaces) and enabling/fixing a gap in direct messages (increased MMS coverage by ~100M IG DMs/day).
- c. **Virality.** In addressing decision inconsistency, we launched virality-based checkpointing and reduced appeals by 50%.

3. NCII

- a. **Proactive Detection.** Boosted NCII photo detection significantly. Increased FB accuracy from 28 → 40% and IG accuracy from 21% → 46% with higher volume. Launched NCII video detection on both FB & IG with almost no data at inception. We are now detecting ~100 NCII videos a day!
- b. **Queue Prioritization.** Change in Q-Pri on IG queue with huge backlog have increased action rate from ~3% to ~22% in Q4.
- c. **Escalation Guidelines.** CO has launched a new escalation guideline which reduced incoming IG reactive volume by >60%.

4. N&P

- a. **Workforce Merge.** Merged CO proactive, reactive and prevalence measurement workforce. We now have a CO measured prevalence that can be compared to PDO! [[HYPERLINK "<https://fburl.com/datainsights/gvtu7nsj>" \h]]
- b. **Capacity Reduction.** Developed a capacity plan to reduce workforce spend by 50% while keeping top line metric goals. Shipped Pages disaggregation rollback saving 850 hours/week.
- c. **Top-line Metric Goals.** Met vNFX, FP Reach goals and hit our aggressive GT Precision goal of ~93%

Q4 LOWlights

1. Maturity

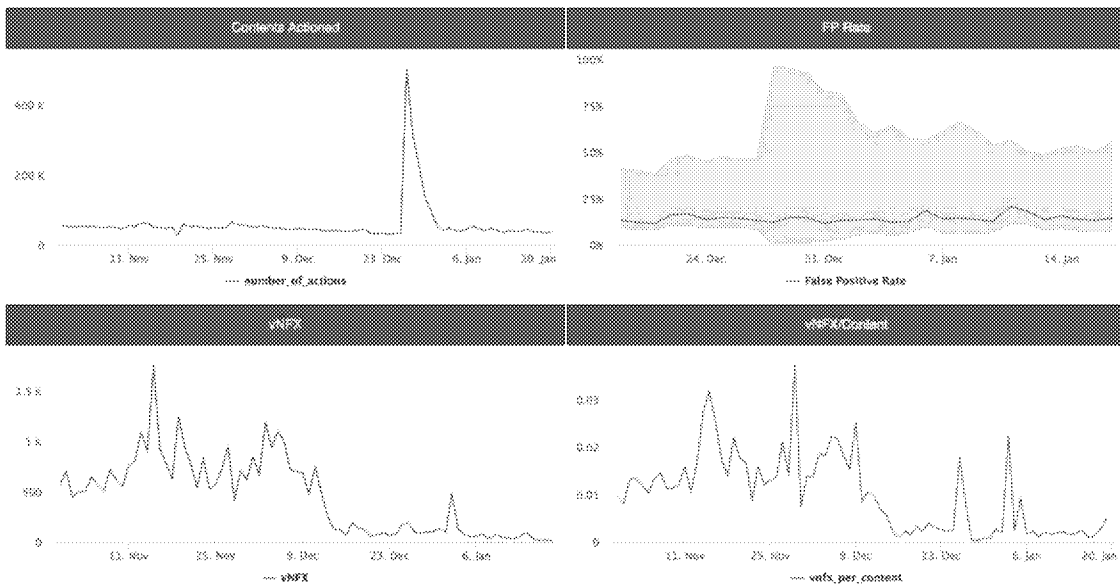
- a. **CEI.** We missed our goal of 1 → 3 on FB.

2. CEI

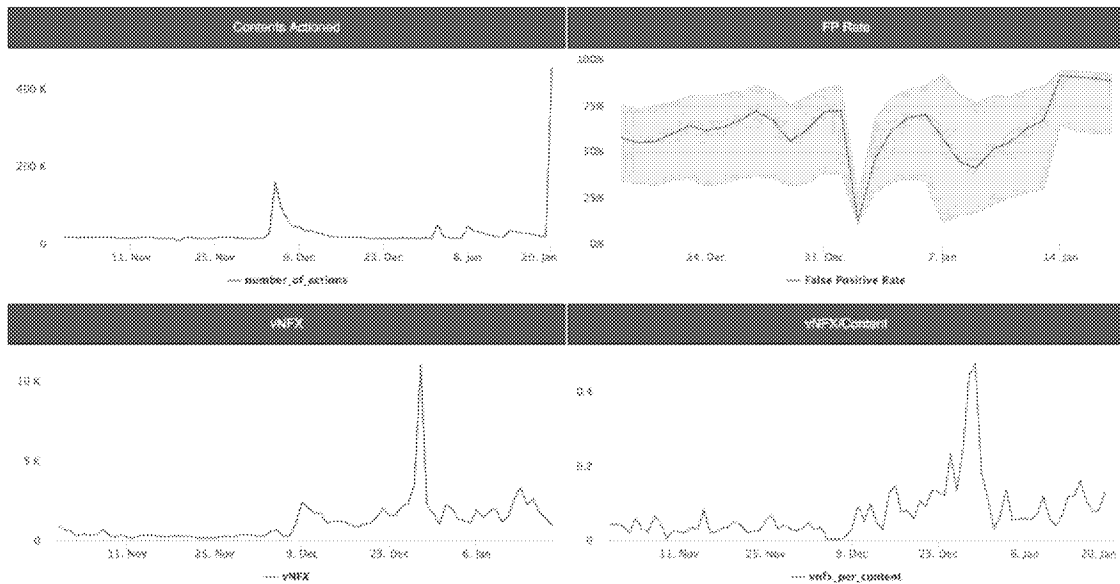
- a. **Rep Attrition.** We've found rep attrition to be higher than expected. As a result, as of the current 66 heads we are allocated, only 50 are filled and 86% have less than 6 months tenure (8 additional in training, 1 resignation in December). Only 6 reps are eligible for banking job review per CO's quality standard.

- b. **Backlog.** This quarter the backlog increased from 845k to 874k. This is due to an increase in the reactive backlog from 106k to 247k. Note, we do have a de-dupe script that regularly runs on the backlog which removed 130k duplicate jobs in Q4. [[HYPERLINK "https://fb.workplace.com/groups/2345294152363949/permalink/2491296521097044/" \h]]
 - c. **Banking.** We missed our banking goals tied to enabling retroaction and increasing bank matches as we uncovered multiple failure points and rep staffing challenges (resiliency concerns) towards the end of the half. [[HYPERLINK "https://fb.workplace.com/notes/[REDACTED]/cei-photo-banking-flow-deep-dive/351889885595389/" \h]]
 - d. **Tiering.** Although we hit our proactive detection goal pertaining to A1-B1 content (industry language for most sexualized + youngest person), throughout the quarter we realized that A1-B1 tiering does not directly map to the type of CEI which, if detected and reported to NCMEC, can potentially save a child. Hence for Q1 and beyond, we propose maintaining our current classifier to "sweep" CEI off our platform while developing a Tier 1 classifier focused on reducing real-world harm. [[HYPERLINK "https://fb.workplace.com/notes/[REDACTED]cei-the-measurement-challenge/331678004283244/" \h], [HYPERLINK "https://fb.workplace.com/notes/[REDACTED]future-directions-of-child-exploitation-images-cei-proactive-detection/363613807756874/" \h]]
3. **NCII**
- a. **Backlog.** Despite changes to escalation guidelines, our current reactive backlog is 1.4M jobs (IG: 940k, FB: 440k) and accumulating 38k jobs/week (IG: 23k/week, FB: 15k/week). We are focusing on reduction and prevention for Q1.
 - b. **Banking.** We established initial bank quality metric but found out that the [HYPERLINK "https://fb.workplace.com/groups/502286320253758/permalink/533480883800968/" \h]. we thus prioritize on bank auditing
 - c. **Recall.** There is little overlap (1.7% for FB, 4.1% for IG) between proactive deletes and reactive deletes implying that our proactive detection may be improved.
4. **N&P**
- a. **CO Published Prevalence.** Although we successfully merged prevalence measurement in the CO workforce using AWD, we missed our aggressive P2 goal of getting DS/IMP/PM alignment to begin goaling/publishing CO measured prevalence in DT by the start of Q1. [[HYPERLINK "https://fburl.com/datainsights/ajy3kg1g" \h]]
 - b. **GT Recall.** Although we hit our GT Prevalence goal of 93.7%, this may have come at the expense of GT recall which dropped from ~80% to ~60%.

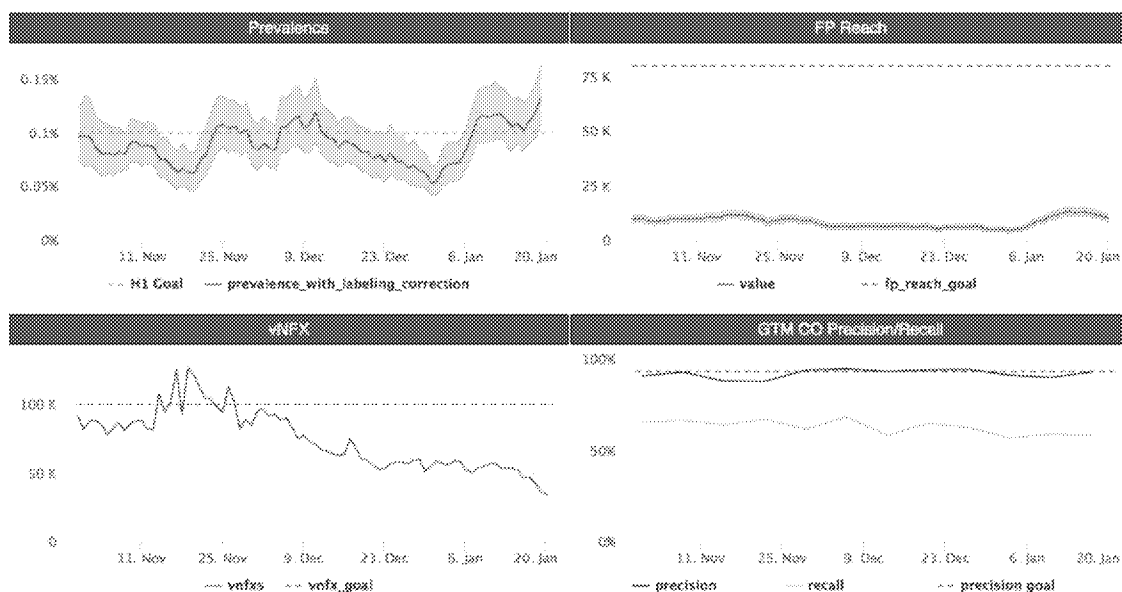
Q4 GOAL METRICS SUMMARY ([HYPERLINK "https://fburl.com/datainsights/oj5b9723" \h])
CEI



NCII



N&P



Q4 Goals STATUS

For more detail information on goal status, see [[HYPERLINK](https://fb.quip.com/cF79AuBajGO6) "https://fb.quip.com/cF79AuBajGO6" \I "SBWACA9C3At" \h]

	Pri	Goal	Actual
1	MA TUR ITY GO ALS		
2	P1	FB NCII from 1-BASIC → 2-MEASURED	COMPLETE
3	P1	IG NCII from 0-RISK → 1-BASIC	COMPLETE
4	P1	FB CEI from 1-BASIC → 3-OPERATIONAL	PARTIAL Hit 2-Measured. Missed 3-Operational Details: [HYPERLINK "https://fb.quip.com/4Cm1m121" \h]
5	P1	IG CEI from 0-RISK → 1-BASIC	COMPLETE
6	P1	FB Prostitution from 0-RISK → 1-BASIC	COMPLETE
7	P1	FB Sexual Solicitation from 0-RISK → 1-BASIC	COMPLETE
8	CHI LD EXP LOI TAT ION IMA		

	GE RY (CEI)		
9	P1	Increase CEI realtime proactive escalations to >5k/week with >20% as A1-B1 [HYPERLINK "https://fburl.com/scuba/yi9z9ums" \h], [HYPERLINK "https://fburl.com/daiquery/898zintj" \h]	COMPLETE Dec-2018: 9.4k A1-B1 reported/week with 20.8% action rate
10	P1	Increase CEI MMS A1-B1 bank matches on FB+IG from 3.6K daily to 4.5K daily [HYPERLINK "https://fburl.com/daiquery/68065m0y" \h]	MISSED We identified multiple broken steps, attrition capacity issues. Note HYPERLINK "https://fb.workplace.com/notes/10156189885595301" \h photo-banking-flow-divide/351889885595301
11	P1	Enable CEI bank retroaction and/or improve IG coverage >2x increase on Android Feed/Story photo hits and enable IG DM raven media matches	PARTIAL CEI Retroaction goal missed. IG MMS work hit goal expanding to iOS)
12	P1	Automate detection and actioning of viral content (outrage/humor memes).	PARTIAL Enabled virality-based checkpointing to A1/E photo banks. It caused 70% fewer user disab and we observed significantly lower app rate.
13	P1	Reduce the # of human reviews in takes to bank CEI	MISSED We found that the A1 label at escalation ste of low quality. We are working on the quality control in Q1
14	P1	Reduce CEI backlog by 120K (de-queue, de-dup, regular sweep) [HYPERLINK "https://fburl.com/unidash/o8fk6e6s" \h]	COMPLETE [HYPERLINK "https://fb.facebook.com/oups/2345294152363044/" \h] dequeued fr CEI banking queue.
15	P1	Write Northstar plan for CEI	MISSED

1 6	P2	Launch rep label consistency dash for CEI across proactive and reactive	MISSED Quality audits have launched for all proactive CEI flows. Dashboard completed
1 7	P2	Bank hits quality auditing (counter-goal to reducing # of reviews)	PARTIAL Quality metrics operational and viewable through HYPERLINK "https://our.internmc.fb.com/inter/unidas/shboard/cei_skeleton/_positive_measurements/h]
1 8	P2	Improve quality or add one or more new pieces of high priority information in EXIF (location, time of capture, or device model) as part of NCMEC report	PARTIAL Legal approval received for EXIF retention. Approval for 14 days received from privacy work is pending, expected for H1.
1 9	P2	Ship Convolutional Featurization v2 with intention of using it on NCMEC data	COMPLETE Note link: [HYPERLINK "https://fb.workplace.com/notes/notes-funes-v2/381661675737072]
2 0	P2	Partner with NCMEC on using convolutional featurization to hash their entire dataset and use it to improve our classifiers	MISSED. Complex goal w/legal eng, and NCMEC impact. Taking same goal in Q4. For Q4, the goal was to get approval from NCMEC and sign the contract with vendor team which was completed. Implementation of the downloader tool will be worked on by a third party vendor for most of 2020.
2 1	P2	Partner with NCMEC and Policy team on revamping Downloader Tool	MISSED Resource constrained
2 2	P2	Understand: CEI prevalence in IG direct messages	MISSED
2 3	P2	Understand: quantify where users discover CEI on FB owned properties (ex: xx% FB search)	MISSED Pushed to Q1
2 4	P2	Estimate size of CEI problem on Messenger. (ex: Ven Diagram of CEI on FB/Messenger)	COMPLETE [[HYPERLINK

			"https://fb.workplace.com/groups/192403221787/permalink/224270577954/" \h]]
2 5	P2	Evaluate internal tool CLARA (by core DS team) for automated label consistency measurement	MISSED No DE bandwidth
2 6	Non-Consensual Intimate Imagery (NCII)		
2 7	P1	Improve FB NCII photo classifier from 28% to 40% action rate w/>400 jobs daily Dash: [HYPERLINK "https://fburl.com/scuba/ptzu3dc6" \h] , [HYPERLINK "https://fburl.com/scuba/ypcjnkm0" \h]	COMPLETE 40% action rate with >800 jobs daily.
2 8	P1	Ship NCII detection on FB video contents with >20% action rate Dash: [HYPERLINK "https://fburl.com/unidash/c5yk4pdk" \h]	COMPLETE 37% action rate with >230 daily jobs. Dashboard
2 9	P1	Increase N&P FB reactive escalation action rate from 15% → 30% Increase N&P IG reactive escalation action rate from 1% → 5% Dash: [HYPERLINK "https://our.internmc.facebook.com/intern/unidash/dashboard/careml_ncii/queue_volumesaction_rates/" \h]	PARTIAL Hit IG goal, missed FB escalation rate: 23% IG escalation rate: 6%
3 0	P1	Deliver recommendation w/mocks for what we should build in Q1Y19 for NCII. (Review in-product NCII report flows, reporting barriers, forms and help resources)	COMPLETE Mocks: [HYPERLINK "https://our.intern.facebook.com/intern/px/p/18W..."]
3 1	P1	Develop NCII tiering to inform proactive detection strategy (ex: viral versus non-viral. should severity of action be different?)	MISSED
3 2	P1	Investigate Complex Entities actioned for NCII outside of Safety	Note: [HYPERLINK "https://fb.workplace.com/notes/...generalized-complex-entity-queues-identify-with-low-accuracy/34294287344/" \h]

3 3	P1	Increase SRT jobs closed/day from 12k/week → 30k/week across FB+IG proactive and reactive queues Dash: [HYPERLINK "https://fburl.com/daiquery/kwpj8rb6" \h]	COMPLETE Weekly closes: ~32k Week of 21-Jan-2019
3 4	P1	Write Northstar plan for NCII	MISSED
3 5	P2	Improve IG NCII photo classifier from 21% to 40% action rate w/>400 jobs daily [HYPERLINK "https://fburl.com/scuba/u5bqftli" \h], [HYPERLINK "https://fburl.com/scuba/a8ra7x8g" \h]	COMPLETE 46% action rate with 700 daily jobs. [HYPERLINK "https://fb.facebook.com/notes/yuan-li/restore-the-recent-volume-drop-in-ncii-ig-photo-proactive-report/367684577310" \h]
3 6	P2	Ship NCII detection on IG video contents with >20% action rate	COMPLETE ~25% action rate [HYPERLINK "https://fburl.com/daiquery/4ueypwqh" \h]
3 7	P2	Bank context as part of NCII banks to give us confidence for future retroaction & fanouts	MISSED
3 8	P2	Support Appeals for NCII	PARTIAL NCII Profile appeals shipped. NCII Content appeals NOT shipped.
3 9	P2	Cross app disables FB→ IG	MISSED Need IG PXFN/Messaging alignment
4 0	P2	Run Proactive NCII detection on Messenger in Dark Mode	MISSED
4 1	P2	Qualitative understand on non-reporters and non-US victims (barriers to reporting)	MISSED Survey designed, held in PXFN
4 2	P2	More fully understand NCST	COMPLETE Link: [HYPERLINK "https://fb.workplace.com/notes/ncst-shallow-diversity-summary/35263936528/" \h] Link: [HYPERLINK "https://fb-my.sharepoint.com/p/EeSFM6C" \h]

			5KksQMZvzEeyoB88 F6_F15HBNdg0ndfHF =7pCa3W" \h]
4 3	Nud y (N& P)		
4 4 4 5 4 6 4 7	P1	Maintain N&P v(NFX+FRX) below 100k/day weekly avg.	COMPLETE
	P1	Maintain N&P FP Reach below 50k/day weekly avg	COMPLETE
	P1	Improve N&P Precision [Baseline - 91%] by 30% to 93.7%	COMPLETE
	P1	Investigate and ship one or more improvement to reduce reactive reports	COMPLETE Reduction in [HYPERLINK "https://fburl.com/unic peggwoy" \h]ts from to 50k/day (Pages Dis Reduction in FB proa linkshares daily target ~20% (110k/day)
4 8	P1	Replacement of 'delete_nudity' Double Review (DR) process with Dynamic Multi Review (DMR)	MISSED Experiment complete Results show DR → I get us 47% (a decrea ~36 hours/day for non marketized reactive N [[HYPERLINK "https://fb.workplace.c notes replacing-dou review-with-dmr-for- np/232897547628177]]
4 9	P1	Merge 100% of N&P reactive and proactive workforce	Goal to carryover in C COMPLETE [HYPERLINK "https://fb.facebook.co oups/3440007361030 ermalink/4991319339 4/" \h]
5 0	P1	Identfy and reduce/disable two or more low action rate or duplicative reporting sources. (CareTools Lockdown as input).	COMPLETE Link: [HYPERLINK "https://fb.facebook.c

5	P1	Reduce CO rep reviews by turning off Pages Disaggregation review model	oups/2345294152363 permalink/245135653 710/" \h] Link: [HYPERLINK "https://fb.facebook.co oups/3236261844934 ermalink/9296002672 4/" \h] COMPLETE Net savings of ~850 hours/week across N and market teams. [HYPERLINK "https://fb.facebook.co otes. roll-back- disaggregation-of-pag reported-for- pornography/259306- 77091/" \h].
5 2	P1	Ship and regularly run SRT deduplication fanout on N&P queues	COMPLETE Shipped the Ignore fa (offline version) on N proactive queues run every 15 mins.
5 3	P2	Complete transition of N&P prevalence flows from PDO to CO	COMPLETE We have PDO and C measured prevalence measured in parrallel today
5 4	P2	Publish a CO based prevalence metric	MISSED. Need consensus from PM, IMP after analysi risks [funnel loss/reca We are ready to publi CO produced prevale metric anytime.
5 5	P2	Merge N&P scaled review with measurement (prevalence+FP reach) flows.	COMPLETE CO prevalence jobs merged into AWD
5 6	P2	Reduce visual model False Positives in 4 N&P categories: art paintings, digital art, nudity, sculptures and breast feeding images	MISSED De-prioritized
5 7	P2	Establish a repeatable and structured RCA process for N&P prevalence	MISSED De-prioritized. GV tak lead.

Q1 Themes, Goals and Projects

Reminder: [[HYPERLINK "https://fb.quip.com/XZBZAQlg1i7L"](https://fb.quip.com/XZBZAQlg1i7L) \h]

What do the priorities mean? Check out [[HYPERLINK](https://www.facebook.com/notes/guy-rosen/my-3-ps-of-priorities/1235429566505768/)

"https://www.facebook.com/notes/guy-rosen/my-3-ps-of-priorities/1235429566505768/" \h]

CEI: Q1 Summary

1. **Banking:** Bank more content and drive more deletes per content banked.
 - a. Data/why: MMS accounts for 95% of contents deleted. Monthly contents deleted have been steadily declining an average of 300k per month throughout H2 2018. We believe this is due to inefficiencies in our banking process. [[HYPERLINK "https://fburl.com/daiquery/tj3cl9yj"](https://fburl.com/daiquery/tj3cl9yj) \h]
 - b. Key goal(s): Increase banking jobs closed from 60 to 600 jobs per week. Increase # of content deletes from newly banked clusters from 50k to 100k per week.
2. **Tier 1 CEI:** Re-define Tier 1 CEI targeting reduction in real world harm and build a classifier prototype.
 - a. Data/why: For proactive detection, we want to use limited review resource on the most important contents - those linked to real world harm. By restricting to A1/B1 labels, we still see the majority are [[HYPERLINK "https://fb.workplace.com/notes/future-directions-of-child-exploitation-images-cei-proactive-detection/363613807756874/"](https://fb.workplace.com/notes/future-directions-of-child-exploitation-images-cei-proactive-detection/363613807756874/) \h]. We would like to define our targets better by considering factors including the content itself (sex act/abuse), user intent, and distribution pattern.
 - b. Key goal(s): Define more targeted Tier 1 concept. Build Tier 1 CEI classifier prototype using a relatively simple definition. Potentially use NCMEC dataset to help bootstrap.

CEI: Q1 Goals

	Area/Theme	Priorit y*	Goal	Goal Own
1	Problem Maturity			
2	Messenger	P1	Get CEI from 0-RISK to 1-BASIC on Messenger	
3	Detection & Actioning			
4	Banking	P1	Increase deletes from recently (last 28 days) banked clusters from 50k to 100k per week	
5		P1	Increase # of banking jobs from 1800 → 4000 jobs/week	
6		P2	Bank quality control - Finish bank FP metric (DS) - Automate process to remove mis-banked clusters	
7	Tier 1 CEI	P1	Re-define & build Tier 1 CEI classifier prototype: target Q2 launch	

8	Messenger	P1	Infra readiness and system improvement for CEI in Messenger -review flow, share root actioning	
9	Video	P2	Improve CEI FB Video classifier to >25% report to NCMEC rate	
10	Automate Severe Meme Actioning	P2	-Apply escalating feature limits in partnership with Actor Experience -Ship virality tracker for FB video -Ship virality tracker for IG photo	
11	Prevalence	P2	Integrate Semantic Search Engine with CEI classifiers on 100% of FB video traffic	
12	Operations			
13	Weekly Business Review Dashboard Efficiency	P1	Problem Area operational dashboard to facilitate weekly CIPCO business review meetings	
14	Banking quality	P2	Save 200 CTR hours/week through process Ops process improvements	
15	Banking quality	P2	Develop a quality baseline for existing CEI banked content (A1-SM) and improve the accuracy level to 99%	
16	XFN Surfaces			
17	Group Images (WA / MSGR)	P2	Run CEI classification on WhatsApp group images or Messenger Requests (offline/sizing)	
18	Messenger	P2	Run one time experiment of reviewing images in violating CEI conversations	
19	Experimentation	P2	Add CEI MMS Metrics to Deltoid to view impact of experiments	
20	Industry Partnerships			
21	NCMEC	P1	Get NCMEC CEI dataset ingested into FB data corpus	
22		P1	Retrain and deploy X-Ray OC leveraging NCMEC CEI dataset	
23		P2	Pass flag for viral content for NCMEC to assist with prioritization	
24		P2	Enable manual submissions of NCMEC reports.	
25		P2	Establish repeatable process to regularly refresh dataset with new NCMEC images & videos	
26	Microsoft	P2	Establish CEI partnership w/Microsoft -Microsoft trains CEI classifier using our CF images and videos -Microsoft starts to CF images/videos of CEI they discover and share with FB	

2 7	Tech Coalition	P2	Create a standard for sharing sensitive tensor information for training our CEI classifier and extendable to other safety areas
--------	----------------	----	---

CEI: q1 Non-Goals

1. One off prevalence estimate on IG DMs to understand the severity of the issue on Direct Messages
2. Improve labelling of CEI complex objects
3. Trim video clips to violating segment before adding to MMS bank

NCII: Q1 Summary

1. **Backlog:** Reduce FB and IG reactive backlogs and incoming volumes.
 - a. Data/why: Our current reactive backlog is 1.4M jobs (IG: 940k, FB: 440k) and the backlog is accumulating 38k jobs/week (IG: 23k/week, FB: 15k/week) . Because 68% of the reactive backlog is composed of IG jobs, we plan to focus more of our efforts on reducing the IG backlog. [[[HYPERLINK "https://fburl.com/daiquery/2inaic0x" \h](https://fburl.com/daiquery/2inaic0x)]]
 - b. Key goals: Reduce IG reactive backlog and incoming volume by 20%, FB by 10%.
2. **Banking:** Bank context as part of NCII banking process for simple objects.
 - a. Data/why: Without the context we cannot confidently assess the quality of MMS clusters/NCII banks overall. The MMS FP Rate is estimated to be 50-70%, but it is unknown how much of this is due to poor bank quality vs. missing context. Until we feel confidently in the quality of our banks we cannot we cannot responsibly increase bank coverage and enable retroaction. [[[HYPERLINK "https://our.internmc.facebook.com/intern/unidash/dashboard/ncii_dashboards/fpmeasurement/?edit=false" \h](https://our.internmc.facebook.com/intern/unidash/dashboard/ncii_dashboards/fpmeasurement/?edit=false)]]
 - b. Key goals: Bank context with simple objects to unlock quality audits and retroaction in Q2.
3. **Proactive Recall:** Improve classifier for FB and IG.
 - a. Data/why: Our NCII proactive deletions have little overlap (1.7% for FB, 4.1% for IG) with reactive deletions suggesting an opportunity for classifier improvement. Higher recall will help 1) uncover more previously unseen cases and 2) allow us to eventually auto-close low scoring reactive reports [[[HYPERLINK "https://fburl.com/daiquery/vbuk81sr" \h](https://fburl.com/daiquery/vbuk81sr)], [[HYPERLINK "https://fburl.com/daiquery/lksdlhcx" \h](https://fburl.com/daiquery/lksdlhcx)]]
 - b. Key goals: Improve NCII classifier recall to 45% for Facebook and to 65% for Instagram.

NCII: Q1 Goals

Area/Theme	Priorit y*	Goal	Goal Ov
------------	---------------	------	---------

1	Problem Maturity			
2	NCII Messenger	P1	Get NCII from 0-RISK to 1-BASIC on Messenger	
3	Backlog			
4	Reduce & Prevent	P1	Reduce IG reactive backlog and incoming volume by 20%	
5		P2	Reduce FB reactive backlog and incoming volume by 10%	
6	Detection & Actioning			
7	Recall	P1	Improve FB NCII classifier recall: - FB 33% → 45% - IG 56% → 65%	
8	Banking	P1	Bank context with simple objects (enables quality audits/retroaction)	
9		P2	Bank contents from Messenger and complex object review flows	
10	Operations			
11	Weekly Business Review Dashboard	P1	Problem Area operational dashboard to facilitate weekly CIPCO business review meetings	
12	Complex Objects	P2	Ship ability to flag/report additional violating content during Profile review	
13	Content Appeals	P2	Ship NCII Content Appeals for FB	
14	Reporting Experience			
15	VHS UX Experiment	P1	-Test new NCII NFX/FRX flow in Q1 -Measure impact/results -Devlop mocks for what to build/ship in Q2 based on results	
16	Understand			
17		P2	NCII International - Do we need different tools, products or policies?	

NCII: q1 Non-Goals

1. Implement tiering to refine our detection.
2. Add content reporting to Messenger.
3. Cross-app disables for FB → IG.

N&P: Q1 SUMMARY

1. **GT Recall:** Although we met our goal of driving GT accuracy to 93.7%, recall plunged to ~58%.
 - a. Data/why: Need high recall to ensure our CO actioning and metrics represent the full scope of N&P.
 - b. Key goals: Maintain accuracy >93% and prevalence <11.5 bps while increasing Recall >80%.
2. **Cost Reduction:** Complete 50% reduction in CO workforce spend for N&P (\$16.8M/year savings)
 - a. Data/why: CO budget constrained/flat for FY2019. 61% of N&P jobs closed d not impact prevalence [[HYPERLINK "https://fb.quip.com/f7g1A2snYEu" \h](https://fb.quip.com/f7g1A2snYEu)].
 - b. Key goals: Decrease hours/week by 50% from 24k per week to 12k per week while maintaining prevalence, accuracy goals.
3. **Semantic Search Engine:** Semantic search engine (SSE) is a new video index matching service built by VUE
 - a. Data/why: In Q4, IntegrityCV built the infra to leverage real-time SSE for detecting similar N&P videos in production and tested it on 0.5% FB VOD traffic. Preliminary results are positive.
 - i. Notes: [[HYPERLINK "https://fb.workplace.com/notes/semantic-search-engine-for-np-videos/280631695935209/" \h](https://fb.workplace.com/notes/semantic-search-engine-for-np-videos/280631695935209/)], [[HYPERLINK "https://fb.workplace.com/notes/semantic-search-engine-h2-lookback-h1-lookahead/336883000252615/" \h](https://fb.workplace.com/notes/semantic-search-engine-h2-lookback-h1-lookahead/336883000252615/)]
 - b. Key goals: Increase video auto-actioning by 10% with SSE matching for N&P (and potentially integrate into CEI classifier as P2).

N&P: Q1 Goals

	Area/Theme	Priorit y*	Goal	Goal Own
1	Top Line Metrics			
2	GT Accuracy	P1	Maintain GT precision at or above 93% weekly avg	[HYPERLINK "https://fb.workplace.com/production.php?id=1000044244914"]
3	GT Recall	P1	Increase GT recall to >80% weekly avg to unlock CO Prev goaling in Q2	[HYPERLINK "https://fb.workplace.com/production.php?id=1000044244914"]

4	False Positive Reach	P1	Maintain FP Reach below 80k/day weekly avg	[HYPERLIN "https://fb.w ace.com/pro php?id=100 932072913"] & [REDACTED]
5	Prevalence	P1	Maintain data transparency reported prevalence of <11.5 bps	[HYPERLIN "mailto: [REDACTED]@fb.com] & [REDACTED]
6		P1	Stabilize Prevalence and FP Reach (Sigma → PHP migration)	[HYPERLIN "https://fb.w ace.com/pro php?id=100 2477325978 br_rs" \h]
7		P2	Transition prevalence metric ownership from PM to CO Ops	[HYPERLIN "https://fb.w ace.com/pro php?id=100 228861284"
8	Efficiency			
9	Cost Reduction	P1	Decrease hours/week by 50% from 24k per week to 12k per week, saving \$16.8M/year	[HYPERLIN "mailto:apro [REDACTED]@fb.com] & [REDACTED]
10	Proactive Detection			
11	Prevalence	P1	Increase video auto-actioning by 10% using Semantic Search Engine matching	[HYPERLIN "https://fb.w ace.com/pro php?id=100 651424345"
12	Borderline	P2	Refresh 2+/3+ image classifier with new Xray OC signals	[HYPERLIN "https://fb.w ace.com/pro php?id=100 231412917"
13	Unification of Workforce & Protocol			
14	Prevalence	P2	Roll out X-QPri across all Nudity enforcement queues -Remove 100% of SLAs for reactive -Remove 100% of Targets for proactive	[HYPERLIN "https://fb.w ace.com/pro php?id=100 904696596"

1	Policy/XFN			
5				
1	GT Accuracy	P2	Ship 2 policy updates that reduce enforcement ambiguity	[HYPERLIN
6				"https://fb.w
				ace.com/pro
				php?id=100
				044244914"

N&P: Non-Goals

1. Take over IG violating N&P proactive work
2. IG maturity 0 → 1 for Prostitution and Sexual Solicitation
3. Auto-apply warning screens instead of quarantining content to improve perception/trust with users.
4. Take over IG Sexual Solicitation proactive work; Scan IG bio using SS classifier
5. Establish a quality measurement framework for transfers (T21079619) to increase action rate.

Q1 Execution

TEAM

	A	B	C
1	Role	People	Burning needs
2	PM		
3	Eng		
4	TPM		
5	DS		
6	DE		
7	Design		
8	Research		
9	Content Strategy		
1	PMM		
0			
1	CO		
1			
1	Policy	Antigone Davis,	
2			
1	IG XFN		
3			

1	Legal	
4		

XFN Needs & Dependencies

Product XFN

Dependencies we've taken on other teams:

Team	Description	Priority	Status
IG Site Integrity	Implement Matching by Hashing Before Modification(MHM) on IG iOS	P1	Agreed T390289
Support UI	NCII FRX/NFX experiment	P1	Agreed
CO	CEI Reviewer resources for IG Video	P2	TBD T393655
SDI	Implement online MHM (Matching by Hashing before Modification) matching in MMS infra utilizing PDQ ([HYPERLINK "https://fb.workplace.com/notes/chih-chiang-wei/instagram-photo-matching-by-hashing-before-modification-on-cei-ct-ncii/223129065233390/" \h])	P2	Agreed T377921
CO	Make CEI checkpoint escalate	P2	Agreed T336708
SRT infra	Improve AWD for CEI rep resiliency	P1	TBD
SRT	Build appeal review flow for NCII (currently only have profile appeal, need content appeal too)	P2	TBD
infra/CareTools			
SDI	Delete the root media when a Messenger share is deleted	P2	(May or require S help) T367759
NCMEC/Outreach	Enable CEI retroaction; coordinate with NCMEC on increasing number of reports	P2	T339932
Detection and Delivery	Take over violating N&P classifier pipeline from DnD	P2	Agreed
Public Safety Engineering	Prioritize ingested hashes based on the number of hits they generate	P2	TBD
Cluster Management	Optimize grouped review flow for CEI (don't dequeue multi-review jobs which have been reviewed once)	P2	TBD T389436

Dependencies other teams have taken on us:

Team	Description	Priority	Status
WhatsApp	Run proactive CEI classifier on group images	P2	Accepted T341747
Messenger Integrity	Run proactive CEI and N&P classifiers on image Messenger requests	P2	TBD (en concerns T377639
CO	Fix Messenger CEI escalations job status so that the accurate job count is shown in ops reporting	P2	Accepted T391715

Stories Integrity	Stop proactively reporting expired stories media	P2	Accepted T378173
Detection and Delivery	Investigate rises in prevalence around the time of the violating N&P classifier migration	P1	Accepted e [HYPER "https://fl m/ei2zlf

CO Resource REQUIREMENTS

Team	Description & Impact	OS HC	Contract or HC	PDI HC	Eng Commit	C R s T 7
CEI	Add FTCs. Capability to action all incoming jobs and decrease backlog volumes while H1 efficiency goals are implemented	-	50	-	-	T 8
N&P	Reduce N&P IG proactive reporting daily targets	-	-	-	-	T 2
	Reduce N&P FB proactive reporting daily targets	-	-	-	-	T 1
	Change the SLAs for N&P reactive queues	-	-	-	-	T 3
	Cluster Review in N&P Proactive Photo	-	-	-	-	T 7
	Shift reports created by FRX Proactive Reporter from the reactive workflow to proactive flows	-	-	-	-	-
	GT for CO Prevalence labelling	-	TBD	-	-	-

content Policy/LEGAL DEPENDENCIES and goals

1. Move from FTC to OS resources for NCII and CEI on FB and IG.
2. Stop reporting viral CEI (humor or outrage) and animated/hand-drawn/cartoon CEI to NCMEC.
3. Remove human review requirement for MMS matches and high threshold scoring CEI.
4. Approval to use Facer for face recognition on potential NCII.
5. Approval to store banking context for NCII.
6. NCII policy alignment/updates for actioning complex objects (pages, groups)

Open Questions / Topics for Discussion

1. **CO Prevalence and the impact on GT Recall.** This quarter N&P is taking new goals on DT facing prevalence and GT Recall as we migrate Prevalence measurement from PDO to CO. CO measured prevalence is exposed to CO recall impacting the accuracy of prevalence. Arguably, PDO measured prevalence has this same problem, but we do not know for certain since we have

no accuracy/recall measurement for PDO. Regardless, CO may be incentivized to drive up precision at the cost of recall thus impacting the recall of CO prevalence. To safeguard against this, we are taking GT recall, accuracy and prevalence goals simultaneously. We are interested in feedback on this plan as it will impact every problem area that transitions Prevalence to CO (ex: Hate and GV).

- a. If faced with GT Precision vs Recall — prioritize recall for prevalence measurement?
2. **Self-funding.** In Q1, N&P is committed to a 50% reduction in CO capacity while shifting prevalence measurement to CO. At one point ([[HYPERLINK "https://fb.workplace.com/groups/146401082647140/permalink/306727786614468/" \h \]](https://fb.workplace.com/groups/146401082647140/permalink/306727786614468/)) problem teams were told we may get up to 50% of our resources back. However, the guidelines were that these resources must be used in the same work type and market (ex: ESLA complex objects savings must be used on ESLA complex objects). Are these guidelines still true today? For Graphic Content, we like to shift the \$ savings we incur for N&P to increasing our spend in NCII and CEI and fund GT recall for CO Prevalence.
3. **Are we aligned on moving away from SLA goals for reactive?** This will unlock true cross Q-Pri for N&P, allow us to truly prioritize prevalence reduction across reactive/proactive in N&P, and allow us to prioritize CEI/NCII based on action rate/vNFX and/or other metrics.
4. **The team is pressing forward with Messenger on all fronts.** Is CI united on this remaining a top priority if E-2-E is definitely coming? Is there a realistic change that we can prevent E-2-E from occurring? If so, how can we best focus our efforts to do that?
5. **We are focusing and prioritizing our efforts on CEI/NCII and maintain with N&P.** This means we let Prostitution, Sexual Solicitation, Creep Shots and Non Consensual Sexual Touching (NCST) go essentially unfunded. Does this sound correct and is aligned with FB priorities overall? We believe so but would like to get feedback.
6. **NRX/FRX for NCII and CEI (Feedback vs Report).** Leave feedback does not match the severity for CEI/NCII. Is there a plan/appetite to change this in the future?

End-2-End Encryption

how big a deal is E2E for us?

1. **Is MSGR E-2-E Encryption a concern for CEI?**
 - a. Yes. Messenger is the platform where the large majority of CEI is shared, accounting for 80% of total H2 2018 deletions (FB & Messenger).

[[
[HYPERLINK "https://fburl.com/daiquery/u7jbmw1a" \h \]](https://fburl.com/daiquery/u7jbmw1a), [[HYPERLINK "https://fburl.com/daiquery/v4jvldse" \h \]](https://fburl.com/daiquery/v4jvldse)]

[REDACTED]

2. **Is MSGR E-2-E Encryption a concern for NCII?**

- a. Yes. Messenger is the platform where the large majority of NCII is shared, accounting for 90% of overall deletions (FB & Messenger).

[REDACTED] [[HYPERLINK "https://fburl.com/daiquery/kn1ou151" \h]]

3. **What are we doing in Q1 and H1 for E-2-E?**

- a. We are going from 0 → 1 maturity on Messenger in Q1 and planning to go from 1 → 2 in Q2 for CEI and NCII.
- b. We have a Q1 P1 goal to develop a "Tier 1" CEI classifier in Q1 -- this will target CEI which, if reported to NCMEC quickly, is the most likely to save a child (i.e. CEI we've never seen before on our platform + involves sex act/abuse w/child). When ready, we plan to deploy this in Q2 on Messenger.
- c. We have a Q1 P1 goal to build out the infrastructure so we can "flip a switch" to start running our Tier 1 CEI classifier on every Messenger image/video once it's complete.
- d. We have preliminary PXFN/Legal/PM approval to roll-out the CEI Tier 1 classifier on Messenger in Q2.
- e. We are working towards understanding discoverability and finding opportunities to prevent violating content seekers from being able to find CEI or connect with other CEI violators.