

PLAYTIME'S OVER.

■ Child Safety in the Poppy Playtime Chapter 5 Content Ecosystem



CHILD EXPOSURE
RISK: ELEVATED
CONTINUOUS
MONITORING
RECOMMENDED

RISK VECTOR ANALYSIS
CHILD-CENTRIC ECOSYSTEMS
THREAT MODEL DISTRIBUTED
CLASS: PERSISTENT

NODE
17-A

CONTENT
PROPAGATION
MAP

VECTOR NODES
■ 24 / ACTIVE
■ 09 / LATENT
■ 03 / CONTAINED



Version 1.0



September 2026



Greymark Research Ltd | digishieldkids.com



DSK-RPT-PP5-0626

DigiShield Kids

PLAYTIME'S OVER

Child Safety in the Poppy Playtime Chapter 5 Content Ecosystem

Version: 1.0

Lead author: Dr Andrew Langley, Lead Researcher, DigiShield Kids (Greymark Research Ltd)

Second author: James Roycroft, Cybersecurity Lead, DigiShield Kids (Greymark Research Ltd)

DOI: 10.5281/zenodo.21065715

Contents

Executive Summary.....	4
1. Introduction and Research Context.....	6
2. Methodology.....	8
3. Release-Window Volume.....	10
4. Roblox Discovery and Clone Ecosystem.....	11
5. Download-Impersonation Layer.....	13
6. Child-Safety Signal Concentration.....	14
7. Recorded Interaction Rates.....	15
8. Exploratory Tag Co-Occurrence Structure.....	17
9. Concentrated Recorded Views.....	18
10. Policy Implications.....	19
11. Conclusion.....	21
Appendix A: Corpus Statistics.....	22
Appendix B: Methodology Notes.....	23
Appendix C: Evidence and Supplementary File Register.....	27
Publication, Legal, and Data Notes.....	28

Executive Summary

This report examines the Poppy Playtime content ecosystem during the Chapter 5 release window using two research streams: DigiShield Kids' keyword-defined monitoring pipeline, which holds 6,890 YouTube and Roblox records collected between 6 February and 19 March 2026 (Section 2.1 describes the publication dates covered); and a separate open-source monitoring strand that documented download-impersonation and scam indicators from 18 February to 18 March 2026. The corpus is an observational, platform- and query-defined dataset. It is not a probability sample of either platform and does not estimate the prevalence of harm across YouTube, Roblox, or all Poppy Playtime content.

The principal finding is an association. Within the 2,526 records selected for contextual analysis, Poppy Playtime-labelled records were more likely than the heterogeneous non-Poppy comparison group to trigger the screen and the encoded adjudication rule. Tag co-occurrence analysis also placed franchise tags, horror labels, and child-context labels in the same large exploratory cluster. These results identify a concentrated monitoring signal around the franchise release window; they do not establish that the franchise caused the difference, that every flagged item was harmful, or that viewers were children.

Key findings at a glance

- The screening stage flagged 256 of 2,526 analysed records (10.1%). The encoded adjudication rule classified 94 records (3.7%) as schema-positive. The 94 records were not all manually reviewed and are therefore not described as “confirmed” cases.
- In the Poppy Playtime-labelled subset, 160 of 775 records were screen-positive (20.6%) and 74 were schema-positive (9.5%). In the unmatched non-Poppy comparison group, the corresponding rates were 96 of 1,751 (5.5%) and 20 of 1,751 (1.1%). The risk ratios were 3.77 (95% CI 2.97-4.78) for the screen and 8.36 (95% CI 5.14-13.60) for the encoded rule. These are internal associations, not causal or platform-wide prevalence estimates.
- A random sample of 100 screen-positive records produced 27 human-positive judgements (estimated positive predictive value 27%; 95% Wilson CI 19.3%-36.4%). A separate random sample of 100 screen-negative records contained two human-positive cases (sample negative predictive value 98%; 95% CI 93.0%-99.4%). The stratified validation design does not by itself identify classifier sensitivity or justify a quantified “high-recall” claim.
- The dataset holds 239 YouTube video records published on 19 February, the largest single day within the publication window of the volume series. This is a count within the monitoring pipeline's keyword-defined collection frame, not a census of all platform uploads or a count of adjudicated harmful items.
- The Roblox query corpus contained 4,370 experience records. At least 1,196 unique records (27.4%) matched one or more of ten recurring title templates. Template categories overlap, so the individual template counts do not sum to the unique total.
- Screen-positive videos with recoverable engagement metadata had higher recorded like and comment rates than the clean-labelled comparison group. This weakens an explanation based only on low-comment view inflation, but it does not authenticate every view, identify viewer age, or prove that the aggregate view count represents unique people.

- Eight channels (0.5% of 1,721 observed channels) accounted for 80% of the summed recorded views attached to screen-positive videos. The data show concentration of observed view totals; they do not identify recommendation systems as the cause.
- The OSINT log recorded at least 18 distinct Chapter 5 download-impersonation sites or listings during the first 30 days. One observed listing was updated 34 days after launch. No payload testing was conducted, so the report distinguishes deceptive or high-risk indicators from a finding that a file contained malware.
- Checks of Mob Entertainment’s official mobile-store presence identified official Chapters 1-4 but no official Chapter 5 mobile release as at the source-verification date. Listings presenting themselves as an official Chapter 5 mobile product should therefore be treated as unverified unless provenance can be established.¹
- The allegations concerning Google Play and Daigo Game 2020 are allegations in *Mob Entertainment Inc v Google LLC et al*, not findings by this report or a court. The complaint alleges more than 1.1 million downloads and payments of USD 30-95 for links that did not provide the represented game.²

Taken together, the findings support event-window monitoring, discovery-system assessment, and coordinated consumer-protection and online-safety scrutiny. They do not support allegations of criminal liability against any named platform, host, creator, or listing operator.

¹ Official platform listings record the console release on 27 May 2026: Nintendo, <https://www.nintendo.com/us/store/products/poppy-playtime-chapter-5-switch/>; Xbox, <https://www.xbox.com/en-GB/games/store/poppy-playtime-chapter-5/9N0JZ8W66LD8>; PlayStation, <https://store.playstation.com/en-gb/concept/10018882> (accessed 14 August 2026). Mob Entertainment’s official Google Play developer page listed Chapters 1-4 and no Chapter 5 mobile app at the source-check date:

<https://play.google.com/store/apps/developer?id=Mob+Entertainment>. Apple’s Mob Entertainment listings likewise showed official releases through Chapter 4, for example <https://apps.apple.com/gb/app/poppy-playtime-chapter-4/id6746137187>.

² *Mob Entertainment Inc v Google LLC et al*, No. 3:25-cv-00531, U.S. District Court for the Northern District of California, Complaint, filed 15 January 2025. Docket metadata: <https://dockets.justia.com/docket/california/candce/3:2025cv00531/442833>. The complaint text attributes the allegations described in Section 5 to the plaintiff; no allegation is presented as an adjudicated fact.

1. Introduction and Research Context

1.1 The DigiShield Monitoring Pipeline

DigiShield Kids operates a continuous monitoring pipeline designed to identify what it terms “Elsagate-ification”: the migration of child-recognisable characters, formats, and search terms into horror-adjacent, manipulative, deceptive, or otherwise age-inappropriate content ecosystems. The pipeline collects public YouTube metadata and public Roblox experience metadata across a keyword set spanning mascot-horror franchises, adjacent viral content, and child-platform gaming. This report covers records collected between 6 February and 19 March 2026. The volume series in Section 3 uses YouTube records published from 20 January 2026; 753 YouTube records in the corpus were published earlier, back to March 2025; 608 of them are in the analysed set, and none is in the volume series (Section 2.1).

Individual franchise reports are focused extractions from this wider dataset. The comparison group is therefore heterogeneous and internally useful, but it is not a matched control. It differs from the Poppy subset in topic and may also differ in platform composition and Stage 1 selection probability. Results against that comparator are reported as associations within the analysed corpus.

1.2 The Scam Intelligence Strand

A separate 30-day OSINT operation ran from 18 February to 18 March 2026. It used dated web-search observations, listing capture, and provenance checks to document sites or posts representing Chapter 5 downloads. Counts are minimum observed counts, not estimates of the total ecosystem. The study did not download or execute offered files, make purchases, attribute operators, or assess the knowledge or legal liability of hosting platforms.

1.3 The Mascot Horror Release Window

Poppy Playtime is a horror-puzzle franchise published by Mob Entertainment. Chapter 5 was released for PC on 18 February 2026.³ The base game reached an all-time Steam peak of 52,866 concurrent players that day according to SteamDB; this is a concurrent-player measure, not a sales or audience-age measure.³ The chapter reached supported consoles on 27 May 2026.¹

The franchise combines toy imagery with explicit horror. That aesthetic can be legible and attractive to children while the official game remains horror content. This report studies the surrounding content and discovery ecosystem, not the suitability of the official game for any particular child.

1.4 Research Questions

1. What proportion of records in the analysed monitoring corpus triggered the screen or the encoded adjudication rule?
2. How did the Poppy Playtime-labelled subset compare with the remaining analysed records during the Chapter 5 release window?

³ Mob Entertainment, *Poppy Playtime - Chapter 5*, official Steam listing, release date 18 February 2026: https://store.steampowered.com/app/4100940/Poppy_Playtime__Chapter_5/ (accessed 14 August 2026). SteamDB, *Poppy Playtime Steam Charts*, records an all-time peak of 52,866 concurrent players on 18 February 2026: <https://steamdb.info/app/1721470/charts/> (accessed 14 August 2026). SteamDB is a third-party tracker and is not affiliated with Valve.

3. What upload-volume, discovery, engagement, and download-impersonation patterns were observed around release?
4. What can - and cannot - be inferred from the tag co-occurrence and channel-concentration analyses?

2. Methodology

2.1 Corpus and Collection

The corpus holds 6,890 public records: 2,520 YouTube video records obtained through the YouTube Data API and 4,370 Roblox experience records obtained through public Roblox endpoints. Collection ran from 6 February (Roblox) and 7 February (YouTube) to 19 March 2026. The YouTube records were published between 5 March 2025 and 18 March 2026. The volume series in Section 3 uses only records published from 20 January 2026; the 753 YouTube records published earlier (29.9% of the YouTube set, 612 of them in 2025) belong to the corpus, 608 of them to the analysed set and none to the Poppy-labelled subset, but not to the volume series. The pipeline did not exist before 6 February, so every record published before the first YouTube run on 7 February, 926 in all (the 753 plus 173 published between 20 January and 7 February), was gathered retrospectively by keyword search between 7 and 17 February. Search returns fewer results for older material and at most 50 items per query, so earlier weeks are sampled more thinly than the release week and afterwards. Section 3 and Appendix B set out what this means for the volume series.

The YouTube set mixes formats. Runs between 7 and 17 February collected videos of any length; from 18 February, runs were paired, one of any length and one restricted to videos of 60 seconds or less; from 4 March, all runs were restricted to 60 seconds or less. As a result at least 1,144 YouTube records (45.4%) are longer than 60 seconds and all of them were collected in February. The run settings are inferred from the duration distribution of each run, because the February and March pipeline logs no longer survive (Appendix B).

Throughout the collection window the pipeline applied an ethics filter that excluded videos whose title or description matched a list of terms for very young children (for example “baby”, “toddler”, “preschool”, “nursery rhyme” and “cocomelon”). The filter was switched to log-only at 11:55 UTC on 18 March 2026. Thirty-three YouTube records and about 250 Roblox records in the corpus were collected after that time. The corpus contains 14 Roblox records (0.2%) collected after the switch that the filter would otherwise have excluded, all matched on “baby”, three by title and eleven by descriptions back-filled after collection; none of the 33 YouTube records collected after the switch would have matched. The number of videos excluded is not fully recoverable: 35 exclusions were logged between 4 and 18 March, and no logs survive for February. The effect is conservative for this report’s purpose, since the filter removed some of the most explicitly child-directed material from the corpus, but it is an undisclosed exclusion in the earlier editions and is recorded here.

The unit throughout is a collected record, not a unique person, impression, or viewing session. The keyword frame covered DigiShield Kids’ wider monitoring remit rather than Chapter 5 alone. Collection used public endpoints and public web search; no private communications or authenticated private content were accessed.

The 6,890-record corpus and the 2,526-record analysed subset are not probability samples. Rates describe this collection frame and period only. Public view, like, and comment counts are time-dependent snapshots. Summed views are aggregate recorded views across videos and may include repeat viewing by the same person.

2.2 Two-Stage Analysis

Stage 1 - prioritisation. Metadata rules assigned records to priority analysis, standard analysis, or filtered out. The rules covered research-keyword presence, timing, engagement, upload frequency, and gameplay-spam patterns. Stage 1 selected 2,526 records (36.7%) for contextual analysis.

During quality review, an upload-frequency rule was found to remove release-window material. Seventy-four videos were recovered and re-analysed. This was a post hoc correction made after inspecting the filter's behaviour; it improves coverage but also introduces an analysis decision that should be considered when interpreting release-window estimates.

Stage 2 - screening. The digishield-contextual-v2 classifier is a weighted pattern-matching system with contextual boosting. It assigns metadata-derived themes, risk signals, and scores. A screen-positive item is not thereby a confirmed harm or a finding about the creator's intent.

Encoded adjudication rule. The PP5 schema in Appendix B was converted into a deterministic rule and applied across the analysed corpus. Outputs from that rule are described as **schema-positive**, not "confirmed," because most were not individually adjudicated by a person. Human validation and rule-human agreement are reported separately.

2.3 Poppy Subset and Comparators

The Poppy subset comprised 775 unique analysed records with "poppy playtime" in the title (587 YouTube videos and 188 Roblox experiences). It is a Poppy Playtime-labelled subset observed during the Chapter 5 release window; it is not limited to content exclusively about Chapter 5. The primary internal comparison is the remaining 1,751 analysed records. That comparison group is unmatched and heterogeneous and may include records from both platforms and multiple topic clusters. It also differs in period: all 608 analysed records published before 20 January 2026 fall in the comparison group (31% of its YouTube records) and none in the Poppy subset. The full-corpus comparator includes the Poppy subset itself and is reported descriptively only.

2.4 OSINT Monitoring Strand

The OSINT strand used daily monitoring during 18-22 February, followed by observations on 25 February, 4 March, and 18 March. A site or listing entered the count when it represented or offered a Chapter 5 download without official provenance or used other recorded deceptive indicators. Operators were not attributed from domain appearance alone. Publicly naming a domain as criminal or malicious would require additional evidence; the public report therefore aggregates bespoke domains and retains identifiers in the regulator evidence log.

2.5 Statistical Analysis

Counts and percentages are descriptive. For the Poppy versus non-Poppy comparisons, this revision reports risk ratios with log-scale 95% confidence intervals and two-sided Fisher exact tests. Wilson intervals are used for validation proportions. Engagement rates are per-video ratios (likes/views and comments/views) calculated where the required metadata were recoverable. Mann-Whitney U tests and item-level bootstrap intervals are retained from the analysis output, but videos from the same channel may not be independent; the reported p-values are therefore exploratory and are not adjusted for channel clustering or multiple analytical choices.

2.6 Anonymisation and Research Ethics

Creator-level outputs are pseudonymised for analysis and aggregated in the public report. No inference is made that an individual creator is an adult, a child, or acting with harmful intent. Public metadata may still be personal data. Greymark Research Ltd treats legitimate interests under UK GDPR Article 6(1)(f) as the likely lawful basis for this private-sector research, subject to necessity, balancing, minimisation, retention, and security safeguards; public availability alone does not remove data-protection obligations.⁴

FINDING 1

⁴ Information Commissioner's Office, *What is the legitimate interests basis?*: <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/legitimate-interests/what-is-the-legitimate-interests-basis/>. ICO, *Principles and grounds for processing* (research provisions): <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/the-research-provisions/principles-and-grounds-for-processing/> (accessed 14 August 2026).

3. Release-Window Volume

239 YouTube video records published on 19 February 2026

The monitoring dataset holds more YouTube records published on 19 February 2026, the day after Chapter 5's PC release, than on any other day: 239. It holds 162 published on release day and 98 published on 20 February. These daily counts group YouTube records by the calendar date of the stored `published_at` value.

The pipeline's weekly aggregation used fixed Tuesday-to-Monday seven-day blocks beginning on 20 January 2026. The four complete pre-release blocks contained 51, 44, 125, and 469 YouTube records, and the block containing the 18 February release, 17 to 23 February, contained 617. The first three blocks were collected retrospectively: every record in them was gathered between 7 and 17 February, the first two blocks between one and four weeks after publication, the third within days, by keyword search. The fourth block was collected partly as it happened and partly afterwards, and the release block was collected as it happened. Because keyword search returns fewer results for older videos, the low counts in the first three blocks reflect the collection method as well as upload volume. The series therefore supports the narrower claim that the release week contained far more matching uploads than the pipeline could find for any earlier week; it does not measure the rate of upload before the pipeline existed, and the apparent acceleration across the four pre-release blocks should not be read as a trend.

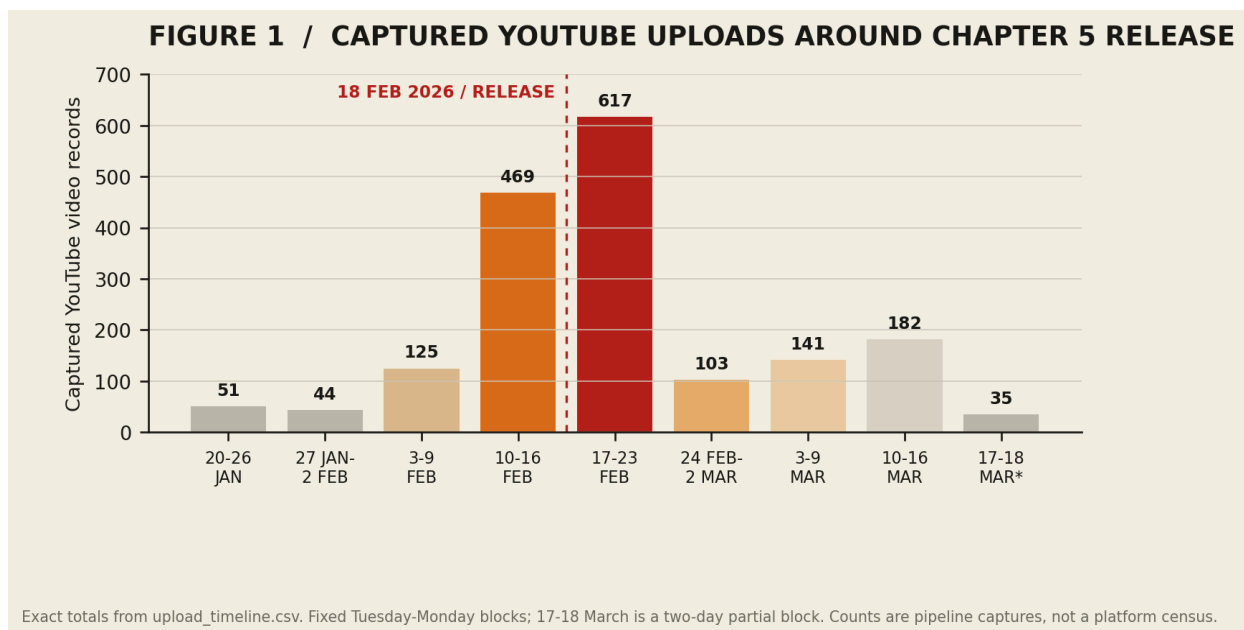


Figure 1. YouTube records by publication week. Blocks before 10 February were collected retrospectively; see Section 3.

The mean daily classifier score was 0.030 on 18 February and 0.036 on 19 February, compared with 0.020 across the 2,526 analysed records. These are descriptive score summaries, not calibrated probabilities of harm.

The OSINT log first recorded Chapter 5 download-impersonation sites or listings within 24 hours of release and at least 12 distinct observations by Day 5. The study did not measure platform moderation throughput or time-to-removal. It therefore supports the narrower conclusion that a large, time-

concentrated review burden formed quickly; it does not establish that a particular moderation system was unable to cope or legally non-compliant.

FINDING 2

4. The Roblox Discovery and Clone Ecosystem

4,370 Roblox experience records in the query corpus; 27.4% matched at least one recurring title template

The Roblox query corpus contained 4,370 collected experience records. The title-template output contained 4,302 usable title rows; 68 collected records shared a title with another record in that output. At least 1,196 unique records - 27.4% of all 4,370 collected records - matched one or more of ten recurring naming templates. The template categories overlap, so the row counts below sum to more than 1,196 and must not be added as though they were mutually exclusive. A re-derivation from the frozen snapshot for this edition gives 1,216 records (27.8%) matching at least one template, or 1,194 distinct titles (27.3%); the figure of 1,196 (27.4%) in the text and in Figure 2 lies between the two and is retained.

Template	Matches	% of all collected records
X Obby	279	6.4%
X Simulator	183	4.2%
Escape X	135	3.1%
Survive X	116	2.7%
X [Chapter N]	113	2.6%
X Tower	110	2.5%
X Horror	102	2.3%
X Run	100	2.3%
X Prison	83	1.9%
X Tycoon	75	1.7%

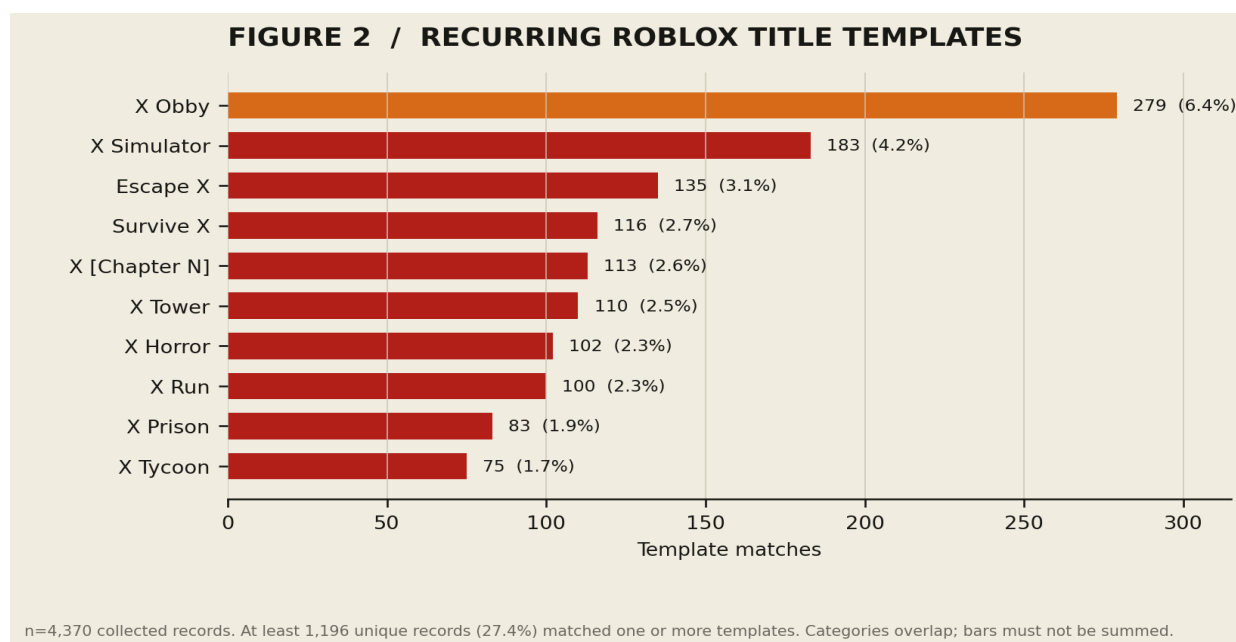


Figure 2. Recurring title-template matches in the Roblox query corpus. Categories overlap.

Exact title duplication was 2.7%. Template variation is consistent with genre conventions and search optimisation, but the dataset does not establish why a creator chose a title or whether a title was intended to avoid detection.

Official provenance and potential confusion

The Roblox listing for *Poppy Playtime Forever* identifies the experience as “By Poppy Playtime on Roblox” and carries copyright notices for Jazwares and Mob Entertainment.⁵ It was the only Roblox experience in this research for which official franchise provenance was verified. Other experiences are described as **apparently unofficial** where no official provenance was identified; that is not a legal finding of infringement.

During the monitoring window, users searching the franchise could encounter multiple Chapter 5-labelled results even though the official Chapter 5 game was not available through Roblox. Some experiences offered Robux-based game passes or group prompts. The report does not infer deception from monetisation alone.

Roblox announced age-based Roblox Kids and Roblox Select accounts after the monitoring window, with roll-out planned for June 2026, alongside content-maturity controls and age checks.⁶ Those later controls were not evaluated by this January-March dataset. Current-policy conclusions should therefore be based on a fresh audit rather than treating the sampled discovery environment as unchanged.

FINDING 3

⁵ Roblox, *Poppy Playtime Forever* official experience listing: <https://www.roblox.com/games/13520190673/Poppy-Playtime-Forever> (accessed 14 August 2026). The listing identifies the experience as “By Poppy Playtime on Roblox” and displays copyright notices for Jazwares and Mob Entertainment.

⁶ Roblox, *Roblox Introduces New Age-Based Accounts and Expanded Parental Controls for Users Under 16*, 13 April 2026: <https://about.roblox.com/newsroom/2026/04/introducing-roblox-kids-and-select-accounts>. Roblox, *Content Maturity Labels*: <https://en.help.roblox.com/hc/en-us/articles/8862768451604-Content-Maturity-Labels> (accessed 14 August 2026). The announcement describes Roblox Kids and Roblox Select accounts, age checks, age-based content access, content ratings, and expanded parental controls.

5. The Download-Impersonation Layer

18+ distinct sites or listings observed within 30 days; maintenance observed after the formal monitoring window

The OSINT log recorded at least 18 distinct sites or listings representing Chapter 5 downloads during the first 30 days after release. This is a minimum observation count. It is not a count of unique operators, verified victims, successful downloads, or sites active simultaneously.

No official Chapter 5 mobile release identified

Mob Entertainment's official Google Play developer page listed Chapters 1-4 but not Chapter 5 when rechecked on 14 August 2026; Apple's official listings likewise identified Mob Entertainment releases through Chapter 4.¹ A listing represented as an official Chapter 5 mobile product therefore lacked verified official provenance at the source-check date. This conclusion should be rechecked if the report is republished later.

The monitored listings used indicators including official-version claims without provenance, instructions to disable device protections, fabricated reviews or safety claims, and download or payment flows that did not match the official release channels. The study did not execute files or perform malware analysis. Accordingly, it does not label an untested payload as malware.

The Mob Entertainment complaint

In January 2025, Mob Entertainment filed *Mob Entertainment Inc v Google LLC et al* in the Northern District of California. The complaint concerns apps described by the plaintiff as unauthorised versions of Chapters 3 and 4. It alleges that the apps accumulated more than 1.1 million downloads, requested payments of USD 30-95, and then supplied a dead link rather than the represented game. It also alleges removal and later reappearance following notices.² These are pleaded allegations. This report makes no finding on their truth, Google's knowledge or liability, the developer's liability, or the ultimate merits of the case.

Evidence-led triage

Priority	Observed listing type	Reason for review
Immediate	Purported Chapter 5 mobile download	No official mobile provenance identified; some listings instructed users to disable protections
High	Purported PC repack on a third-party publishing service	Unverified provenance or protection-disabling instruction
Contextual	Piracy or aggregator listing	Legal, security, and provenance risks vary; no lower-risk assumption is made
Contextual	Fan-made experience	Potential source confusion; fan status alone is not evidence of fraud or infringement

The priority labels are an internal research triage device, not a criminal, legal, or malware classification.

Observed development over time

The log increased from one observed site or listing in the first day to at least 12 by Day 5, approximately 16 by Week 2, and at least 18 by Day 30. Later pages in the sample tended to contain more elaborate

search-optimised guides. The data do not establish that the same operators became more sophisticated over time. One observed listing carried an update dated 24 March 2026 - 34 days, or just under five weeks, after launch. The earlier description of this as “six weeks” was incorrect.

Public reporting aggregates purpose-built domains to reduce unnecessary legal and safety risk. Exact URLs, capture dates, screenshots, and indicators remain in the regulator evidence log. Mainstream hosting or publishing services are mentioned only as locations where user-generated posts were observed; no inference is made about operator knowledge, endorsement, or liability.

The console release as a second observation window

Chapter 5 reached supported consoles on 27 May 2026.¹ That later release provides a pre-specified opportunity to test whether similar query and impersonation patterns recur. The January-March evidence does not itself establish what happened after 27 May, and no claim is made here about the age profile of console players.

FINDING 4

6. Child-Safety Signal Concentration

8.36x schema-positive rate in the Poppy-labelled subset versus the unmatched non-Poppy comparison group (95% CI 5.14-13.60)

Across 2,526 analysed records, the screen flagged 256 (10.1%). The encoded adjudication rule classified 94 records (3.7%) as schema-positive. Summed public view counts attached to the 256 screen-positive records totalled 121,829,240 at the relevant metadata snapshots. That figure is aggregate recorded views, not unique viewers, verified humans, or verified children.

Comparison	Poppy-labelled subset (n=775)	Non-Poppy group (n=1,751)	Risk ratio (95% CI)
Screen-positive	160 (20.6%)	96 (5.5%)	3.77 (2.97-4.78)
Schema-positive	74 (9.5%)	20 (1.1%)	8.36 (5.14-13.60)

The comparison group also differs from the Poppy subset in period (Section 2.3); the risk ratios are not adjusted for this.

Both two-sided Fisher exact tests returned $p < 0.001$. Because the collection is not a probability sample and the comparison group is unmatched, the confidence intervals quantify sampling-style uncertainty around the observed 2x2 tables; they do not remove selection, platform, keyword, or measurement bias.

Against the full analysed corpus, the corresponding descriptive ratios were 2.04 for the screen and 2.57 for the encoded rule. Those ratios are not independent comparisons because the full corpus contains the Poppy subset.

The result supports a strong association within this analysed corpus: Poppy Playtime-labelled records selected for Stage 2 were more likely to satisfy both classification tiers than the remaining analysed records. It does not show that the franchise caused the difference, that 9.5% of all Poppy Playtime

videos on YouTube are harmful, or that the comparison would remain unchanged after matching by platform, date, channel, format, language, and Stage 1 inclusion probability.

The screen's flag categories overlap. Counts of individual flags therefore exceed the number of unique screen-positive records:

Screen flag	Count
horror_game_for_kids	154
horror+child_context	148
fear_clickbait	25
shock_clickbait	14
emotional_manipulation_kids	6
predatory_free_currency	2

The highest-view schema-positive items included endings, rescue or distress narratives, lore explainers, and short-form character content. Titles and channel identifiers are retained for regulator disclosure. The report does not infer creator intent from these formats.

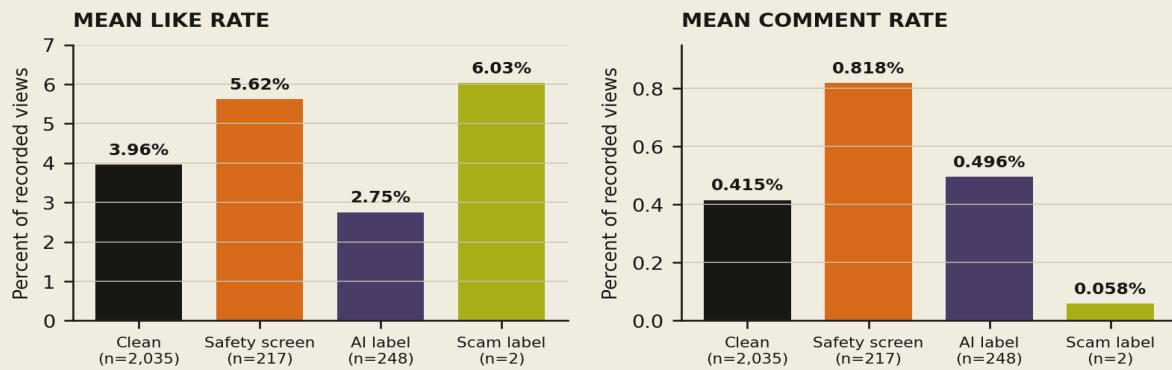
FINDING 5

7. Recorded Interaction Rates

4.4x median comment-rate ratio for schema-positive records versus clean-labelled records

Engagement metadata were recoverable for 2,502 analysed records. Rates are per-video likes divided by recorded views and comments divided by recorded views. The descriptive groups overlap except for the clean-labelled baseline; AI- and scam-labelled rows should not be interpreted as mutually exclusive categories.

Group	n	Mean like rate	Mean comment rate	Descriptive comparison with clean
Clean-labelled	2,	0.0396	0.00415	baseline
Child-safety screen-positive	21	0.0562	0.00818	+42% likes; +97% comments
AI-labelled	24	0.0275	0.00496	-31% likes; +20% comments
Scam-labelled	2	0.0603	0.00058	n too small for inference

FIGURE 3 / RECORDED INTERACTION RATES BY ANALYSIS LABEL

Per-video rates; groups overlap except the clean-labelled baseline. Rates do not authenticate traffic or identify viewer age. Scam-labelled n=2: no inference.

Figure 3. Mean recorded interaction rates for records with recoverable metadata.

The child-safety screen-positive group also had higher medians: comment rate 0.00126 versus 0.00034 and like rate 0.0252 versus 0.0159. The reported two-sided Mann-Whitney test for comment rate was $U=283,241$, $p<0.001$, corresponding to a rank-biserial effect of approximately 0.283; the item-level bootstrap 95% interval for the median difference was 0.0006-0.0015.

For the 94 schema-positive records, the median comment rate was 0.00152 and median like rate 0.0287. Against the same clean baseline, the reported comment-rate test was $U=132,152$, $p<0.001$ (rank-biserial effect approximately 0.382), with an item-level bootstrap 95% interval for the median difference of 0.0005-0.0028. A sensitivity analysis excluding zero-comment videos retained the direction of the result.

These patterns show more recorded interaction per recorded view in the flagged groups. They are inconsistent with an explanation based solely on a specific low-comment anomaly pattern, but they do not prove that every view was human. Automated accounts can also like or comment; disabled comments cannot be distinguished from zero comments in the metadata; repeated videos from the same channel reduce independence; and Stage 1 and the adjudication rule use engagement or reach information. The engagement analysis is therefore not an independent validation of the safety labels.

A heuristic identified 64 videos with more than 100,000 recorded views and fewer than 10 recorded comments. They totalled 93 million recorded views and 223 comments. The rule is described as a **low-comment anomaly screen**, not a bot detector, because no ground-truth traffic data were available.

The defensible policy inference is limited: the observed safety signal cannot be dismissed simply because some unrelated records show low-comment view anomalies. The data do not identify viewer age and do not convert aggregate views into a count of exposed children.

FINDING 6

8. Exploratory Tag Co-Occurrence Structure

148 tags assigned to the largest horror/child-context cluster under the specified procedure

The analysis constructed 6,332 co-occurrence edges among 751 tags occurring at least five times. The input combined platform-supplied YouTube tags with classifier-assigned theme and risk labels. A seed-expansion procedure assigned tags to eight labelled clusters. This is an exploratory descriptive model; the labels were applied by the researchers and are not independently validated latent classes.

Cluster	Research label	Tags
1	Poppy Playtime horror/child-context	148
2	AI cat / animated content	60
3	AI girlfriend / romance	49
4	Roblox / emerging IP	35
5	Smiling Critters / Chapter 3	33
6	AI companion reviews	25
7	Poppy Playtime gameplay / walkthrough	22
8	Family-friendly Roblox	22

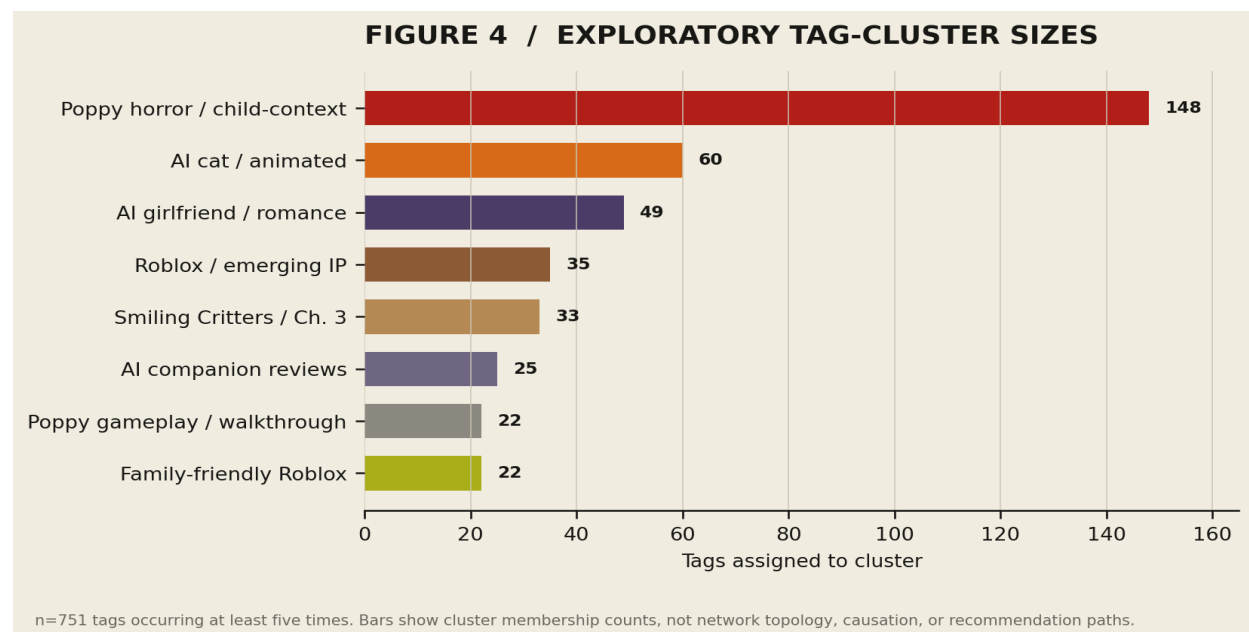


Figure 4. Exploratory tag-network cluster sizes. Cluster labels are researcher interpretations.

Cluster 1 contained Poppy Playtime franchise tags, horror-related labels, and child-context labels. Cluster 8 formed separately under the same procedure. This supports the narrower claim that those fields repeatedly co-occurred in the analysed metadata and were grouped together by the chosen algorithm. It does not by itself prove platform recommendation, audience age, harm, or causal integration.

The Poppy-character x horror/child grouping appeared in 311 videos (11.1% of the relevant network input) across 433 unique tag pairs, with summed edge weight 5,064. Horror x child terms appeared in 282 videos, and scam x platform terms in 63. Summed edge weight is a descriptive co-occurrence count, not an effect size or independent confirmation of the classifier.

FINDING 7

9. Concentrated Recorded Views, Distributed Production

0.5% of observed channels accounted for 80% of summed views on screen-positive videos

The analysed records came from 1,721 channels. Eight channels (0.5%) accounted for 80% of summed recorded views attached to screen-positive videos, while 86 channels (5.0%) accounted for 80% of the screen-positive video count.

This is a Pareto description of the observed dataset. It does not show why those videos received views. Search, subscriptions, external traffic, channel size, publication timing, and recommendation systems are all possible contributors. Accordingly, the result supports prioritising a small number of high-view nodes for platform review, but it does not establish that recommendation systems caused the concentration or that creator-level action would be ineffective.

Channels remain unnamed in the public report. The analysis concerns distribution, not an allegation against individual creators.

10. Policy Implications

The recommendations below distinguish current legal duties from proposals. They do not offer a legal opinion on any named service's compliance.

Assess predictable event windows within existing risk duties

Ofcom's framework already distinguishes the children's access assessment - the threshold exercise for in-scope user-to-user and search services - from the children's risk assessment that follows where a service is likely to be accessed by children.⁷ Ofcom's Children's Risk Profiles identify recommender systems as a risk factor, and the Protection of Children Codes include recommender-system measures.⁷

The additional policy proposal from this research is that services and regulators should test whether predictable franchise or release windows materially change the incidence and distribution of relevant content, and document that evidence in risk-assessment reviews. This is not a statement that current law expressly mandates a separate "franchise event" assessment. Ofcom's June 2026 crisis-protocol measure addresses exceptional public-safety crises, a narrower category than ordinary entertainment-release spikes.⁷

Use cross-platform evidence without blurring legal remits

The monitoring data connect search demand, video publication, Roblox discovery, and download-impersonation listings. Each service remains responsible only within the legal framework applicable to it. Ofcom, the CMA, trading standards, app-store operators, rights-holders, and independent researchers have different powers and evidence. A practical response would use a common event-window evidence format - timestamps, provenance, search terms, listing status, and action history - so that evidence can move between those bodies without implying that one regulator controls the entire chain.

Re-audit Roblox after its 2026 age-control changes

The sampled Roblox query environment predates the roll-out of Roblox Kids and Roblox Select accounts.⁶ The report therefore recommends a like-for-like post-implementation audit of search results, content-maturity labels, official-provenance cues, and Robux prompts. Ofcom has publicly identified Roblox among major platforms in its recommender-system monitoring work; this makes discovery-layer evidence relevant, but this report does not make a compliance finding.⁷

Treat false official-version claims as a consumer-protection issue

Misleading commercial claims and deceptive app or download representations can engage UK consumer-protection law. The CMA has direct consumer-enforcement powers under the Digital Markets, Competition and Consumers Act 2024, while other enforcers retain roles.⁸ Whether a particular

⁷ Ofcom, *Children's access assessment duties under the Online Safety Act*, updated 29 June 2026: <https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/childrens-access-assessment-duties-under-the-online-safety-act>. Ofcom, *Protection of children duties under the Online Safety Act*: <https://www.ofcom.org.uk/online-safety/protecting-children/protection-of-children-duties-under-the-online-safety-act>. Ofcom, *Statement: Protecting children from harms online*, including recommender-system measures: <https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/statement-protecting-children-from-harms-online>. Ofcom, *Statement: Crisis response protocol*, 9 June 2026: <https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/statement-crisis-response-protocol> (all accessed 14 August 2026).

⁸ Competition and Markets Authority, *Unfair commercial practices* (CMA207), 18 November 2025: <https://www.gov.uk/government/publications/unfair-commercial-practices-cma207/unfair-commercial-practices>. Digital Markets, Competition and Consumers Act 2024: <https://www.legislation.gov.uk/ukpga/2024/13>. CMA and Ofcom, *Online safety and competition in digital markets*, 14 July 2022: <https://www.gov.uk/government/publications/cma-ofcom-joint-statement-on-online-safety-and->

listing or app-store practice infringes the law depends on evidence and is for the competent authority or court. The report recommends a referral protocol for child-recognisable franchise impersonation, especially where payment is requested or device protections are discouraged.

Preserve evidence before public attribution

For each high-risk listing, the minimum evidential package should include capture time, URL, represented product, provenance check, payment or download flow, safety instructions, hosting context, and later status. Public naming should follow legal review and, where appropriate, a right-of-reply or responsible-disclosure process. Aggregated public reporting can coexist with named regulator disclosure.

11. Conclusion

During the Chapter 5 release window, the DigiShield monitoring corpus showed a concentrated Poppy Playtime-related safety signal, a release-week concentration of matching uploads, recurring Roblox title templates, concentrated recorded views, and rapid appearance of unverified download listings. The strongest numerical comparison is between the 775 Poppy-labelled records and the 1,751 unmatched remaining analysed records: screen-positive risk ratio 3.77 (95% CI 2.97-4.78) and schema-positive risk ratio 8.36 (95% CI 5.14-13.60).

Those figures are meaningful but bounded. The corpus is keyword-defined, Stage 1 selected only 36.7% of collected records for contextual analysis, the comparison group is heterogeneous in topic, platform and period, and the 94 schema-positive records were produced by an encoded rule rather than universal manual adjudication. Aggregate views are not unique viewers; interaction metrics do not authenticate traffic or identify children; the network analysis is exploratory; and OSINT indicators do not establish criminal liability or malware content.

Within those limits, the evidence supports a system-level research proposition: predictable entertainment releases can create short, cross-service windows in which age-inappropriate framing, source confusion, and deceptive download representations become easier to encounter. The appropriate response is repeatable event-window measurement, provenance cues, discovery-layer assessment, evidence preservation, and regulator-ready disclosure - accompanied by language that remains as careful as the evidence.

Appendix A: Corpus Statistics

Metric	Value
Total collected records	6,890
YouTube video records	2,520
YouTube records published before the first YouTube run (collected retrospectively 7-17 February)	926 (36.7%): 753 published before 20 January, 173 from 20 January
YouTube records longer than 60 seconds (all collected in February)	at least 1,144 (45.4%)
Roblox experience records	4,370
Publication dates of YouTube records	5 March 2025-18 March 2026 (753 records before 20 January 2026)
Collection runs	6 February-19 March 2026 (YouTube 7 February-18 March)
Records submitted to Stage 2	2,526
Records filtered at Stage 1	4,364
Videos recovered in post hoc filter review	74
Screen-positive records	256/2,526 (10.1%; 95% Wilson CI 9.0%-11.4%)
Schema-positive records	94/2,526 (3.7%; 95% Wilson CI 3.1%-4.5%)
Aggregate recorded views on screen-positive records	121,829,240 (not unique viewers)
Poppy-labelled analysed subset	775 unique records (587 YouTube videos, 188 Roblox experiences)
Poppy screen-positive	160/775 (20.6%; 95% CI 17.9%-23.6%)
Poppy schema-positive	74/775 (9.5%; 95% CI 7.7%-11.8%)
Non-Poppy analysed group	1,751 heterogeneous records
Non-Poppy screen-positive	96/1,751 (5.5%; 95% CI 4.5%-6.6%)
Non-Poppy schema-positive	20/1,751 (1.1%; 95% CI 0.7%-1.8%)
Poppy vs non-Poppy screen risk ratio	3.77 (95% CI 2.97-4.78)
Poppy vs non-Poppy schema-positive risk ratio	8.36 (95% CI 5.14-13.60)
OSINT monitoring period	18 February-18 March 2026
Distinct download-impersonation sites/listings observed	at least 18

Confidence intervals do not correct for the non-probability collection frame, unmatched groups, platform mixture, selection into Stage 2, or classification error.

Appendix B: Methodology Notes

Pipeline keyword scope

The monitoring pipeline keyword set covers the full Elsagate-ification research remit, not PP5 specifically. Keyword clusters include mascot horror franchises and horror-adjacent content. The heterogeneity of the corpus is by design and enables genuine concentration analysis.

Collection timing and retrospective capture

The collection database was created on 6 February 2026. Records published before the first YouTube run on 7 February were gathered retrospectively by keyword search between 7 and 17 February; the 926 YouTube records of this kind include 753 published before 20 January 2026 and 173 published between 20 January and 7 February. For records first stored on or after 4 March 2026, the pipeline logs recovered in September 2026 give the true first-collection time: median 3.7 days after publication, 90th percentile 7.0 days. For the 2,047 records stored before 4 March the database holds only the most recent time the record was seen, because the pipeline overwrote the collection timestamp on each re-collection until 14 September 2026; for those records the lag between publication and collection is bounded below by the first run after publication and above by the last-seen time. All database timestamps are UTC, which coincides with UK time for the whole collection window.

Content-format mix

Of the 2,520 YouTube records, at least 1,144 (45.4%) are longer than 60 seconds. The long-form records were all collected in February: runs between 7 and 17 February collected any duration, runs from 18 February were paired between any-duration and 60-seconds-or-less, and runs from 4 March collected only videos of 60 seconds or less. These settings are inferred from the duration distribution of each collection run, since the pipeline logs for February and March were deleted by a log-retention setting in September 2026. The engagement comparisons in Section 7, the reach concentration in Section 9 and the low-comment anomaly screen therefore mix short-form and long-form records, whose baseline interaction rates differ. The report does not split them; the supplementary files carry a duration field for readers who wish to.

Child-directed content filter

The pipeline's ethics module applied a keyword filter for content aimed at very young children throughout the collection window, excluding videos whose title or description matched terms such as "baby", "toddler", "preschool", "nursery rhyme" and "cocomelon". The filter was set to log-only at 11:55 UTC on 18 March 2026; 33 YouTube records and about 250 Roblox records were collected after that point. The corpus contains 14 Roblox records (0.2%) collected after the switch that the filter would otherwise have excluded, all matched on "baby", three by title and eleven by descriptions back-filled after collection; none of the 33 YouTube records collected after the switch would have matched. Thirty-five exclusions were logged between 4 and 18 March; the February count is unrecoverable. The exclusion works against the report's central association, since it removed some of the most explicitly child-directed items before analysis, and the Poppy-labelled rates in Section 6 are if anything understated relative to an unfiltered corpus. Editions 1.0 and 1.1 did not disclose the filter.

Filter rules (Stage 1)

Content farm detection (10+ uploads same day, multiple themes); low engagement threshold (below 1,000 views, over 30 days old, under 60 seconds duration); gameplay spam (PART #, EPISODE #, FULL GAMEPLAY, WALKTHROUGH). Priority flags: research keywords (scam, free robux, hack, generator); high engagement (over 100K views); Poppy Playtime character terms; within 14 days of release date; Roblox free games.

Null results: reading level and sentiment

Flesch-Kincaid grade level analysis found no statistically significant difference between child safety flagged and clean content titles ($p=0.704$, effect size $r=0.014$). VADER sentiment analysis found no statistically significant difference in title sentiment (title negative-sentiment score, $p = 0.852$). Description-layer sentiment showed higher negative scores in child safety flagged content, suggesting emotional loading occurs below the title level. These null results constrain interpretation: child-targeting is not detectable through title readability or sentiment, pointing to character-based, algorithmic, and description-layer mechanisms rather than lexical ones.

Roblox timestamp limitation

Roblox published_at timestamps reflect collection date rather than game creation date, and the Roblox view_count field records the number of players online at the moment of the search, not cumulative visits. Temporal claims are restricted to YouTube data throughout, and no Roblox popularity ranking is reported. Roblox findings are reported as aggregate ecosystem characteristics without temporal sequencing.

AI classification note

The digishield-contextual-v2 classifier uses weighted pattern matching with contextual boosting. Its flags and risk scores are metadata-derived screening signals, not findings of harm. Validation is reported below. The available stratified samples estimate positive and negative predictive values within those samples but do not establish full-corpus sensitivity or recall. Throughout this report, screen-positive and schema-positive results are kept separate from human adjudication.

Classifier and Rule Validation

Validation used two stratified random samples (seed 20260610): 100 screen-positive records and 100 screen-negative records. A separate highest-view sample and deliberately chosen edge cases were descriptive stress tests and are not combined with the random-sample estimates.

Validation sample	Human adjudication result	Correct interpretation
100 randomly sampled screen-positive records	27 positive; 73 negative	Estimated screen positive predictive value 27% (95% Wilson CI 19.3%-36.4%)
100 randomly sampled screen-negative records	98 negative; 2 positive	Sample negative predictive value 98% (95% Wilson CI 93.0%-99.4%); false-omission proportion 2%
25 highest-view screen-positive records	11 human-positive	Purposive sample; not an unbiased precision estimate
10 selected edge cases	qualitative error patterns	Diagnostic only; not an accuracy estimate

Because sampling was stratified on the classifier output, the two percentages above do not directly provide sensitivity, specificity, or recall for the full analysed corpus. Describing the screen as “high recall” is a design intention, not a demonstrated performance statistic.

The encoded adjudication rule was compared with the human judgement in the 100-item screen-positive validation sample. It agreed on 80 records and disagreed on 20: 14 rule-positive/human-negative and 6 rule-negative/human-positive. Treating the human judgement as the reference for this limited sample gives 77.8% sensitivity, 80.8% specificity, 60.0% positive predictive value, 90.8% negative predictive value, and Cohen's kappa 0.536. This is moderate agreement, not interchangeability. No independent second-human adjudication or inter-rater statistic was available.

The encoded rule returned 94 schema-positive records across the analysed corpus. Those outputs inherit rule error and should not be described as 94 manually confirmed cases. The screen and rule also contain reach- or engagement-related features, so later engagement comparisons are not independent validation.

Statistical Methods and Reproducibility Boundary

Risk-ratio intervals use the log (Katz) method; validation-proportion intervals use Wilson's method; 2x2 p-values use two-sided Fisher exact tests. Engagement tests use two-sided Mann-Whitney U statistics and 10,000-resample item-level bootstrap intervals as recorded in the analysis output. The bootstrap did not cluster by channel. Exact software versions, zero-denominator handling, and the raw analysis rows must accompany any claim of computational reproducibility.

The supplementary CSV files listed in Appendix C were not embedded in the submitted DOCX used for this revision. Arithmetic and statistical framing were checked from the reported counts, but the item-level analyses could not be independently rerun in this document pass.

Network Construction and Limits

The tag network combines platform-supplied YouTube tags (unprefixed), classifier themes (theme:), and classifier risk labels (flag:). Title-derived keywords were not included. Association between separately sourced fields reduces direct rule tautology, but theme and flag nodes remain derived from the same record metadata and are not independent observations. Cluster labels are researcher interpretations. The network supports descriptive co-occurrence, not causation, recommendation-path inference, or prevalence. Cluster membership in the seed-expansion procedure depends on set-iteration order and did not reproduce exactly when re-run for this edition (a deterministic re-run gives largest-cluster sizes of 140, 49, 46, 40, 34, 33, 33 and 25 against the 148, 60, 49, 35, 33, 25, 22 and 22 reported); the cluster sizes should be read as indicative and the labels as researcher interpretation.

Flag Counts, Overlaps, and Engagement Subset

Flag categories are not mutually exclusive. The 256 screen-positive records are unique items with at least one child-safety flag. Engagement analysis used 2,502 records with recoverable view, like, and comment metadata; 217 were child-safety screen-positive. Missing engagement data can produce selection bias if missingness is related to upload age, removal, platform, or risk.

OSINT Limitations

Site and listing counts are minimum observed counts. The study did not execute files, make purchases, identify victims, estimate financial loss, or attribute common control. Search indexing, localisation, personalisation, and later deletion affect observability. Domain or platform appearance alone does not establish operator identity, intent, knowledge, endorsement, criminality, or legal liability.

DigiShield Kids adjudication schema (PP5)

This schema defines what constitutes a schema-positive child-safety concern within the PP5 dataset, as distinct from a screen-level flag. A verdict of relevant means the screen flag is justified on age and suitability grounds; not-relevant means it is a false positive. The question is not “is this content harmful?” in the abstract, but “does this content warrant a child-safety concern based on its format, framing, and audience signals?” The schema was applied to the manually reviewed samples and then encoded for corpus-wide application. The flag hierarchy and false-positive patterns are general; the format-signal list is specific to this franchise ecosystem and would need revision for a different one.

Reach gate. Items with zero recorded audience are presumptively not-relevant. Items at 1 to 500 require a Tier A flag to be relevant. Items at 500 or above proceed to content analysis. Reach at or above 100,000 lowers the threshold on borderline flag combinations; reach at or above one million makes a single strong flag sufficient. Reach is a modifier on genuine concern, not a source of concern in itself.

Flag-type hierarchy. Tier A (strong, one flag can support relevant): `emotional_manipulation_kids`; or `horror+child_context` and `horror_game_for_kids` triggered together. Tier B (moderate, supports relevant when combined with child-targeting signals in the content): `horror+child_context` alone; `horror_game_for_kids` alone at risk score at or above 0.1; `fear_clickbait` with a second flag. Tier C (weak, presumptively false positive): `horror_game_for_kids` alone at risk score below 0.1; `fear_clickbait` alone; `shock_clickbait` alone; `emotional_manipulation` without the `_kids` variant.

Content-format signals, independent of the flags. Presumptively child-targeted: comic or story dub; toys-react or toy-unboxing hybrid; lore explainer for a children’s IP; character death or survival emotional title; Roblox obby; tea-party or cute-plus-horror hybrid. Neutral: standard walkthrough; reaction compilation; lore theory; speedrun or showcase; behind-the-scenes; non-English gameplay. Adult-skewing or general-audience: AI-presidents or AI-politicians meme format; general-audience franchise crossover; documentary or journalism; creator vlog.

Known false-positive patterns, marked not-relevant without further analysis: sexualization flags on craft, drawing, clay, fan-art, academic, food, or obby content; `fear_clickbait` on food, foraging, health journalism, nature, film trailers, or academic content; `horror_game_for_kids` on Nextbots-in-Playground or Gmod sandbox content (physics sandboxes with comedy intent, not horror for kids); `horror+child_context` on “brainrot” meme content (absurdist internet-meme genre, not horror); and `horror+child_context` on Gmod survival or escape content (third-party sandbox, not a children’s IP product).

Combining the signals. A relevant verdict requires one of: a Tier A flag, or Tier B plus a presumptively child-targeted format, at reach of 500 or above (clear case); or Tier B plus a neutral format at reach of 100,000 or above, or a Tier A flag at reach of 100 to 500 (borderline case). Otherwise the verdict is not-relevant.

What does not constitute child-targeting evidence: a single capitalised adjective in the title; an exclamation mark or fire emoji; a character name in the title; a plot-summary title; or a high view count on its own.

Schema limitation identified on application. The Tier C placement of `fear_clickbait` alone caused two items concerning AI chatbots embedded in children’s products to be missed, one of them at 1.1 million

views and confirmed relevant on human review. The current taxonomy has no category for AI-exploitation-of-children content. This is logged as a taxonomy gap for the next pipeline iteration and is treated in separate DigiShield Kids work.

Appendix C: Evidence and Supplementary File Register

The following files are retained in the project evidence bundle. They are not embedded in this DOCX and should not be described as publicly deposited unless persistent links and file checksums are added before publication.

- `tag_network.csv` - co-occurrence edge list (6,332 edges)
- `tag_cooccurrence.csv` - high-weight co-occurring tag pairs
- `channel_pareto.csv` - 1,721 pseudonymised channels ranked by screen-positive counts and recorded views
- `engagement_forensics.csv` - per-record engagement calculations and anomaly fields
- `upload_timeline.csv` - daily counts of YouTube records by publication date; Roblox timestamps excluded from temporal inference
- `roblox_title_templates.csv` - 4,302 usable title rows with template fields
- `poppy_playtime_corpus.csv` - 840 collected rows covering 832 unique videos; 775 unique Stage 2 records

Two of the 2,526 analysed records (analysis identifiers 1588 and 2494) refer to a YouTube video and a Roblox experience that were removed from the live database in July 2026 after their thumbnails could not be retrieved. The counts in this report are as at analysis on 18-19 March 2026 and include both; the supplementary files include the YouTube record's analysis row and March metadata; the Roblox record appears in none of them, and the live database no longer holds either.

- `sexualization_investigation.csv` - 50-record diagnostic sample; not cited as a report finding

Regulator disclosure should use a minimised version with a data dictionary, generation date, code version, pseudonymisation method, and SHA-256 checksum for each file. Public release requires a separate privacy and platform-terms review.

Publication, Legal, and Data Notes

Publication record

Edition: Version 1.0, 09/10/2026

Methods correction pass: This edition corrects the description of the corpus: collection ran from 6 February to 19 March 2026, not from 20 January to 18 March, and 753 of the YouTube records were published before 20 January 2026, the start of the volume series. It discloses that records published before 7 February were gathered retrospectively, discloses the mixed short-form and long-form composition of the YouTube set, discloses the child-directed content filter active during collection, notes two analysed records whose source metadata was later removed, and corrects the description of the Poppy-labelled subset, which comprises 587 YouTube videos and 188 Roblox experiences, not 775 YouTube videos. Reproduction of the headline figures from the frozen March 2026 snapshot for this edition confirmed 38 of 55 checks exactly and led to four corrections: the clean-labelled engagement group is $n = 2,037$ (previously 2,035); the flagged comment-rate median is 0.00126 (previously 0.00127); the low-comment anomaly screen's 64 videos total 93 million views and 223 comments (previously 96 million and 43); and the analysed-corpus mean classifier score is 0.020 over 2,526 records (previously 0.017, population unstated). One unreproducible statement (five channels with every video flagged) is removed, and the Roblox template share and tag-cluster sizes are caveated. No risk ratio, confidence interval or validation figure changes. The corrections follow an audit of the collection pipeline in September 2026.

Original report date: 26 June 2026

Legal, statistical, and source-verification pass: 14 August 2026

Entity: Greymark Research Ltd, publishing through DigiShield Kids

DOI shown in the source document: 10.5281/zenodo.21065715

This revision preserves the original observation windows. Later source checks are identified where they update product availability, platform controls, or legal context.

The severity words and micro-labels used on the cover are editorial design elements. They do not create a separate legal classification or statistical finding beyond the qualified conclusions in the report body.

Copyright and permitted use

Copyright © 2026 Greymark Research Ltd. Schools, safeguarding professionals, regulators, researchers, and community organisations may reproduce and distribute the report **unchanged and in full** for non-commercial purposes with clear attribution. Commercial reuse, sale, or adaptation requires prior written permission. Nothing in this notice restricts quotation, criticism, review, or other uses permitted by law.

Third-party names and trademarks are used descriptively to identify the subjects of research. No affiliation, sponsorship, or endorsement is claimed. The report does not reproduce third-party screenshots in the public edition. Where third-party material is later added, the legal basis and sufficient acknowledgement should be assessed item by item rather than relying on a blanket fair-dealing statement.

Legal characterisation

Terms such as *unverified*, *impersonation indicator*, *deceptive indicator*, and *scam classification* describe the project's research criteria. They are not findings of criminal liability. Allegations from litigation are attributed to the pleading. Fan-made or unofficial content is not treated as unlawful merely because it lacks official provenance. The report does not determine the knowledge, intent, endorsement, or liability of hosting platforms, app stores, creators, domain registrants, or listing operators.

Data protection and research scope

The research used public metadata and content through public APIs and web search. No private communications, private accounts, or authenticated private content were accessed. Publicly visible account or channel data can still be personal data. Creator identifiers were pseudonymised for analysis and omitted from the public report. Greymark Research Ltd's likely lawful basis for this private-sector research is legitimate interests under UK GDPR Article 6(1)(f), supported by purpose, necessity, and balancing assessments and by data-minimisation, retention, access-control, and security safeguards.⁴

Responsible disclosure

The public edition aggregates bespoke download domains. A regulator evidence bundle may contain identifiers, dated captures, and provenance checks under controlled disclosure. Findings are intended for consideration by Ofcom, DSIT, consumer-protection bodies, app-store operators, platforms, and relevant industry organisations. This wording does not state that any body has accepted the findings or opened an investigation.

About DigiShield Kids

DigiShield Kids is an independent child online-safety research initiative based in North Yorkshire and operated by Greymark Research Ltd. It uses open-source research to study online content, games, AI services, and cross-platform risk. Professional-development workshops, briefings, and bespoke research are available through **digishieldkids.com** or **info@digishieldkids.com**.

Greymark Research Ltd | Popeshead Court Offices, Peter Lane, York YO1 8SU | digishieldkids.com