

DEFINITIONS

Elsagateification

Elsagate was a scandal about disturbing videos aimed at young children. Elsagateification is our name for the recurring process that produced it. Alongside the definition we set out a model of why it happens, which predicts the conditions under which Elsagate-style content will flourish. This note sets out what we mean by Elsagateification, the formula for content production and why it could help platforms and regulators get to grips with inappropriate content in children's online spaces.

What Elsagate was

Between roughly 2016 and 2017, parents began noticing strange videos in the feeds of their young children, on YouTube and on the YouTube Kids app. The videos used characters that children adored: Elsa from *Frozen*, Spider-Man, Peppa Pig, Mickey Mouse. The thumbnails were bright and cheerful, but the content ranged from odd and surreal to frightening and genuinely disturbing.

Some were just crude animations, others used adults in cheap costumes, some featured real children. The characters were injected with needles, buried alive, made pregnant, sat on toilets, trapped, hurt, made to cry. Many had no dialogue, so they worked across every language. Titles were often long strings of search terms (*Elsa Spiderman Baby Injection Learn Colors Finger Family*) written to be found by the recommendation algorithm. Videos ran long, played on autoplay, and were produced in bulk by anonymous channels, sometimes with millions of subscribers.

Users on Reddit gathered examples under the name Elsagate, after the character who appeared most often.¹ In late 2017 the writer James Bridle published a widely shared essay describing the phenomenon, and the *New York Times* reported that the material was reaching children through YouTube Kids.² YouTube responded by removing channels, stripping advertising from large numbers of videos and tightening its policies. The story faded, and Elsagate came to be treated as something that had happened and been dealt with.

From event to process

Despite the post-2017 cleanse, the motifs of Elsagate survived and thrived. The videos were the surface layer of a wider internet-native mechanism which could not be stopped just by removing the offending content. We are proposing the term **Elsagateification** for that underlying process. Elsagate names an outbreak, and outbreaks keep coming.

Commentators now number its successive waves, from 'Elsagate 1.0' in 2016 and 2017 up to an 'Elsagate 3.0' of AI-generated videos reported in 2025.³ *Elsagateification* is the process which produces these outbreaks, and keeps producing them as characters, platforms and formats change around it.

DEFINITION

Elsagateification is the recurring process by which content that carries the surface signals of children's entertainment, but treats its subject in ways unsuitable for the age of the children it reaches, gains and regains a share of children's media environments as characters, platforms and production methods change.

In plain terms: things that look made for children, but are not suitable for the children watching them, keep taking up space in children's media, and keep coming back when they are cleared away.

Reading the definition

Each part of the definition can be checked against content directly, without first assuming why that content exists.

Surface signals of children's entertainment. The cues that tell a child, a parent or a classifier that something is meant for children: familiar characters, a bright cartoon style, nursery formats, child-directed titles and thumbnails. A borrowed character is the strongest of these signals. Original AI creatures and generic cartoon animals can carry the same signals without borrowing anyone.

Unsuitable in treatment, for the age reached. Subject matter alone does not qualify. Good children's programmes deal with doctors, injury, fear and family conflict all the time. What qualifies is treatment: how intense it is, whether distress is shown for its own sake, whether a child is the victim, whether it is sexualised or humiliating, and whether the story leaves the viewer anywhere safe to land. Suitability is always judged against a specific age group.

The children it reaches. Content counts only where there is evidence that children encounter it, such as its appearance in children's feeds, made-for-kids labelling, or recommendation after children's programmes. View counts alone cannot show who is watching.

Gains and regains a share. The measure is proportion. Rising upload counts can simply mean more of everything. The process shows up as a rising share of qualifying content within a consistently sampled children's environment, and as that share recovering after clean-ups.

The definition makes no claim about causes, intent or effects. Why the process happens is a separate question, answered by the model below, which can be revised as evidence arrives without changing what it explains. The definition does not assume that producers set out to harm children, and it requires no coordination between them. It also keeps apart three things

that are often run together: content judged unsuitable for an age group, evidence that children of that age encounter it, and evidence of its effects on them. This note concerns the first two. Showing the third needs a different kind of study.

Why a new word is needed

Elsagate, as a term, points to a set of videos. Responding to such inappropriate content is simple: find the bad videos, take them down. Every hosting platform has responded roughly that way, and the material has reappeared each time under new characters. The incentive structure, the platform structure and the psychology of how and why children respond to certain content themes doesn't change just because some of the videos are gone. We are arguing for a new term and definition which can locate these structures and hooks in order to tackle the problem.

How it works

Elsagateification has a small number of moving parts, none of which are new on their own. Together they form an engine that keeps producing the same kind of content, whatever the platform and whoever the character of the moment happens to be:

1 The pull: a friendly face in the wrong story

The core formula is to take a face that draws children in, and put it in a scenario that touches something children worry or wonder about. The strongest hook is a character a child already loves and trusts. Cute animals, cartoonish faces and bright, toy-like designs also draw children in, more weakly. Neither half works well alone. A Peppa Pig birthday party has the character but no hook. A cruel-stepmother story with generic cartoon animals has the hook but a much weaker draw. Together they create a pull that is hard for a young child to resist, partly because the scenario feels forbidden/unknown for that character.

2 Why it works on children

The scenarios are not made at random, on the contrary they map onto the ordinary fears and curiosities of childhood: bodies and toilets, doctors and needles, family break-up, being bullied or left out, injury and death. Children are drawn to stories that let them look at these things from a safe distance. In this content the distance is false, and the anxiety is used to hold attention instead of being resolved. Young children are also less able to look away or spot manipulation, so material an adult would reject continues to work on them. Repetition is always key in children's books, films and TV content.

3 Characters cut loose from their stories

A character's own story and world is what makes it safe for children. Bluey is trusted by parents because of how Bluey's family behaves and how Bluey's world works. Once a character is lifted out of their world and dropped into any scenario, the trust travels with it and the safeguards stay behind. Fan art, game mods, AI image tools and knock-off animation make that lifting even more trivial than in 2017. The more out of character the new scenario, the stronger the pull can be for children.

4 The money

Views earn advertising revenue, and children's viewing is long, repetitive and loyal. When YouTube had to remove personalised advertising from children's content after a 2019 privacy settlement, creators of made-for-kids videos produced 18 per cent less of it.⁴ Children's content production follows the money, and this content is no exception.

5 The cost of making it

In 2017 a video still took some effort to produce. Generative AI has reduced that cost to almost nothing. One operator can now produce hundreds of variants in a day, publish them all, and keep making whichever perform. More attempts mean faster discovery of what works on children.

6 The algorithm as editor

Recommendation systems reward whatever holds attention. Over time they shape what gets made as well as what gets seen, because creators copy whatever the feed rewards. Researchers have described the result as an 'algorithmic uncanny', meaning content that feels shaped by machine optimisation more than by any human intention.⁵

7 Staying just under the line

A second pressure works alongside engagement: avoiding removal. The most successful material is transgressive enough to hold a child's attention and ambiguous enough to pass automated checks and a parent's glance. A thumbnail that looks like a cartoon, a title made of search terms, nothing that clearly breaks a rule. The two pressures pull against each other, and the content optimises where both are satisfied.

8 Crossing between platforms

What works in one place can be copied everywhere else. A horror game rated 12+ or 16+ is rebuilt on Roblox within hours, where games can be published, tested and monetised almost instantly. Reaction videos carry it into children's YouTube feeds, clips carry it onto TikTok, and toys and school bags carry it into the playground. Each platform moderates its own space, yet no-one owns the platform drift.

Why it behaves like evolution

Taken together these parts behave like natural selection. Producers generate variation, cheaply and at volume. Money and the algorithm decide what survives. Successful forms spread, and are copied across platforms. The characters are the visible surface that keeps changing; the formula underneath persists and the same motifs keep returning with new faces.

Later content does not even need to directly copy ElsaGate for this to happen. The same environmental pressures lead different producers to the same solutions, the way unrelated animals evolve similar shapes in similar environments. It is also why removing videos one at a time never ends the problem. Removal can clear a space, but where the pressures remain the space simply refills. This comparison is our working model, its value lies in what it lets us predict and what we can't predict.

The formula

The eight parts can be reduced to five terms. Written as a formula, the mechanism looks like this.

THE ELSAGATEIFICATION FORMULA

$$\text{Elsagateification} = (\text{Character} \times \text{Scenario}) \times \text{Reward} \times \text{Ease} \times \text{Deniability}$$

Character How strongly the face on screen draws a child in. Strongest when it is a character the child already knows and trusts.

Scenario How strongly the situation touches childhood fears and curiosities, and how far it goes out of character.

Reward The money and recommendation reach that pay for making more.

Ease How cheaply the content can be made, varied and copied across platforms.

Deniability How well each step stays below the threshold that would get it removed.

The formula shows the parts and how they combine, but it's not yet a measurement. None of the terms has an agreed scale, so no number should be read from it yet. The empirical work has yet to come.

When it takes hold, and when it should not

Naming a process is only useful if the name predicts something, thus we offer this working hypothesis that *Elsagateification is expected wherever five conditions hold together*.

- 1 An appealing character.** A face that draws children in. Strongest when it is borrowed from a story children already know and trust, which lets it be lifted out of that story. Cute animals and cartoonish faces also work.
- 2 A resonant scenario.** A situation that touches the ordinary fears and curiosities of childhood.
- 3 A reward.** Advertising revenue or recommendation reach that reliably pays for producing more.
- 4 Cheap reproduction.** The content can be made, varied and copied faster than it can be gated.
- 5 Defensible steps.** No single item or action crosses a threshold that any one system is designed to catch.

Where all five hold, the process is predicted to take hold. Where one is genuinely absent, it should not. That is the claim we intend to test, with data and observation.

Why it matters for regulation

The UK Online Safety Act already frames platform duties around systems and processes rather than individual pieces of content. Elsagateification is a systems-level phenomenon by construction, so the concept and the duty point in the same direction. The gap is in enforcement practice, which still largely operates one item at a time. Treating the pipeline instead as the unit of harm changes what questions a regulator can ask. Alongside *is this video permitted?*, a regulator can ask how quickly does a rated property appear unrated on a children's platform, and whose responsibility that lag is. Which recommendation behaviours carry age-rated material to under-thirteens? Where does licensed merchandise sit in a chain that a platform disclaims?

WHAT WE ARE AND ARE NOT CLAIMING

Elsagate is not our term. It was named by the online communities that documented the original content in 2017, and the credit is theirs.

Elsagateification is the contribution we are offering: the process concept, the model of how it works, and the five conditions above.

Where this goes next

This note is the reference definition. A preprint will follow with the full theoretical treatment, its place in the existing literature, and formal falsification criteria. An empirical paper will come after that.

SOURCES

1. <https://www.reddit.com/r/ElsaGate/>
2. June, L. (2017). YouTube has a fake Peppa Pig problem. *The Outline*, March 2017; Bridle, J. (2017). Something is wrong on the internet. Medium, November 2017; Maheshwari, S. (2017). On YouTube Kids, startling videos slip past filters. *New York Times*, November 2017.
3. Holliday, L. (2025). Dozens of YouTube channels are showing AI-generated cartoon gore and fetish content. *WIRED*, 2 May 2025. On the 'Elsagate 3.0' label, see for example Raymundo (2025). Elsagate 3.0 is worse than we thought. YouTube, February 2025.
4. Johnson, G. A., Lin, T., Zhong, L. and Cooper, J. C. (2026). COPPAcalypse? The YouTube settlement's impact on kids' content. *Management Science*. doi:10.1287/mnsc.2024.05295
5. Balanzategui, J. Disturbing children's YouTube genres and the algorithmic uncanny. (2021). *New Media & Society*. 25(12), 3521-3542. doi:10.1177/14614448211049264

Cite as: DigiShield Kids (2026). *Elsagateification: A Definition*. Concept note v1.1. Available at digishieldkids.com.