

Companion for Kids

A Private, Offline AI Companion for Children

Version: 1.0 **Built on:** Companion v2.0 (Evidence-Based Revision) — the eldercare platform **Status:** Concept extension. No code written yet for this audience.

Executive Summary

Companion for Kids is an extension of the Companion idea — originally built as a privacy-first AI companion for older adults — now aimed at children.

The underlying insight that made the eldercare product work carries over directly: a small, offline, on-device language model doesn't need to be capable of high-level reasoning to be a good companion for someone who isn't asking it for high-level reasoning. Older adults mostly want to be listened to, remembered, and gently kept company. Children mostly want to be listened to, remembered, and delighted by a familiar voice. Neither audience is trying to get research-grade answers out of the device — and that's exactly the case a small, cheap, entirely local model is well suited to.

Companion for Kids runs on **the same codebase, the same Gemma-class model, and the same staged roadmap** as the eldercare product (Stages A–E, described below). What changes is the personality layer: instead of a single companion persona, the platform supports a library of **original, in-house characters** — each with its own voice, its own memories, and its own way of talking to a child — plus a substantially expanded parental control system, because the buyer and the user are different people with different needs.

This document assumes familiarity with the eldercare white paper (`Companion_White_Paper_v2_0.md`) and calls out only what's different for this audience. Where this document is silent, the eldercare spec applies.

A note on characters

An earlier draft of this concept proposed letting children's Companion units impersonate existing franchise characters — a Frozen-style ice princess, a band of small woodland miners, a wisecracking cartoon rabbit, and so on. We're not doing that, and it's worth being direct about why: those characters are owned by Disney, Warner Bros., and other studios, and a commercial device that ships pretending to be them is copyright and

trademark infringement, not a gray area. It's also a fragile foundation for a product — a licensing dispute or a takedown request could gut the feature overnight, and you'd own none of the goodwill you'd built.

The fix costs almost nothing and gains something real: **build original characters that hit the same emotional notes**. A warm older-sister figure with a gentle, protective voice. A cheerful band of small, industrious forest friends. A quick-witted woodland trickster with a soft heart. Kids don't need the Disney logo to love a character — they need a consistent voice, a warm personality, and memories that make the character feel like *theirs*. Built in-house, Companion owns every character outright, can develop them without asking anyone's permission, and can eventually spin them into their own book, show, or merchandise line if they take off. That upside doesn't exist with borrowed IP.

The character library described below is written on that basis.

System Design Principles

Carried forward from the eldercare product, unchanged:

1. **100% offline in normal operation.**
2. **On-device processing only.** No raw audio, text, or embeddings leave the device.
3. **Low-power embedded hardware.**
4. **Privacy-by-design**, including the negative space — no camera.
5. **Fail-safe simplicity.**

Added for this audience:

1. **The parent is the account owner; the child is the user.** Every design decision that could conflict between "what the child wants" and "what the parent needs to see" resolves in the parent's favor.
 2. **The companion never asks a child to keep a secret from a parent**, never discourages a child from telling a parent something, and never frames itself as a replacement for parental attention. This mirrors the eldercare "encourages family contact, doesn't replace it" principle, adapted for a child's relationship to their own guardians.
 3. **Nothing is collected that isn't shown to the parent.** The parent dashboard (below) is not a partial view — if the device knows it, the dashboard can show it.
-

Why a Low-Capability LLM Fits This Audience

This is the same argument the eldercare paper makes about latency and model size, applied to a different population:

- **Neither audience needs frontier reasoning.** A child asking their companion character about their day, or an older adult reminiscing about a trip to Paris, is not solving a math problem. A 2–4B parameter model, quantized to Q4, handles warm, memory-aware conversation well — the eldercare prototype already demonstrated this.
 - **Smaller models are easier to keep safe.** A model with a narrower range of capability has a narrower range of ways to go wrong. This matters more for children than for any other audience the company will ever serve.
 - **Cost and power stay low**, which matters for a toy price point far more than it matters for eldercare hardware.
 - **The same time-to-first-audio latency architecture applies.** Children are, if anything, more forgiving of a brief "thinking" sound than adults — a short pre-rendered acknowledgment ("Hmm, let me think!") before the first sentence streams in reads as *in character* for a wisecracking companion, not as lag.
-

Character System

Architecture

This is not a new mechanism built for kids — it's the same persona library already running and validated on the eldercare product. The eldercare Companion isn't limited to one fixed companion voice: its

`PersonalityStore` loads any well-formed personality file at runtime and swaps it without restarting the model, and this has already been demonstrated past the single-default case with a fully structured persona for a historical figure (temperament, documented speech patterns, values, a defined behavioral stance) loaded through the exact same pipeline. Companion for Kids reuses that persona library wholesale: a loadable personality description plus a stacking preference store, **no fine-tuning required**. Each character is a personality file, not a different model — the underlying Gemma-class model and the memory system are shared across every character in the library, exactly as they are shared across eldercare personas.

What's new for this audience is the **character library's content** (original characters designed for kids, per the note above) and **per-character memory partitioning** — the eldercare product's single historical persona didn't need a separate memory space from the default companion, but a child talking to four different characters does.

The initial character library (original, in-house)

Character	Archetype	Voice & tone	Notes
Skye	Gentle ice-realm princess	Warm, protective, a little wistful; sings when happy	Original character; visual design and backstory owned outright
The Timberfolk	Seven small forest craftsmen	Cheerful, busy, encouraging about chores and effort	A found-family ensemble persona; can "hand off" between named members
Fenn	Quick-witted forest trickster	Playful, teasing, ultimately kind	A rabbit-adjacent trickster archetype, redesigned from scratch
Nightlight	Gentle bedtime companion	Slow, soft, calming; specializes in wind-down routines	Purpose-built rather than archetype-derived; useful as a default

Each character carries its **own memory partition** — the same episodic/semantic memory system from the eldercare architecture, keyed per-character rather than per-user. A child talking to Skye about their lost tooth and to Fenn about their soccer game gets two consistent, separately-remembered relationships, not one blended personality. A child can also switch which character they're talking to at any time; switching characters does not erase memory, it changes which memory partition is active.

Content boundaries (built into every character file)

- No character ever role-plays violence, romance, or content outside a general-audience range, regardless of what a child asks for in-character.
- No character discusses topics outside its designed scope (medical, legal, or crisis topics) — those trigger a scripted, gentle redirect to "let's go find a grown-up" rather than an in-character answer.
- Every character shares the Tone Adaptation module from the eldercare spec (§5 in the parent document): simpler language when the child seems confused, calm pacing when they seem upset.

Storytelling Mode (shared with eldercare)

This is another already-built feature carried over directly, not a kids-specific addition — Storytelling Mode is Milestone 7 of the shared software core, and it works the same way for both audiences.

How it works. A story is imported and split into paragraph-aware, overlapping chunks, embedded the same way conversation turns are, and stored tagged with its title. A character's system prompt instructs it to treat a story's content as "something it has read," not something that happened to it, which is what lets a character discuss a story accurately without that story bleeding into the character's own claimed memories. Asked a

content-specific question about an imported story, the character correctly describes what happens and why — validated behavior, not a target.

For kids specifically, this is the mechanism behind "tell me the story about the little duckling again" or "what happened to the fox in that story" — a character discussing and revisiting a story it knows, in its own voice, drawing on the same recall path used for a child's own remembered facts and preferences.

Verbatim narration — designed for both, not yet built for either. The next planned extension is a true "read it to me" mode: the story's original unchunked text handed straight to the voice layer, bypassing the language model entirely for the actual reading, with a bookmark tracked outside the model so a mid-story interruption ("wait, who's that?") can be answered and reading resumed from the same spot rather than restarted. For kids, this is the bedtime-story use case in its most literal form — pause for a question, then "should I keep reading from where we left off?" This is scoped work, shared with the eldercare roadmap, not a kids-only feature to build separately.

Shared Codebase and Roadmap

Companion for Kids is **not a fork**. It runs on the same staged roadmap as the eldercare product, with the same gates:

Stage	Eldercare goal	Kids equivalent
A — Software core	Prove conversation + memory architecture	Reused as-is; already complete
B — Handheld field prototype	Validate the emotional hypothesis on a phone/tablet before hardware spend	Same — test character personas and parental controls on a tablet app first
C — Embedded port	Prove the experience survives \$60 hardware	Same target hardware (Pi 5 / CM5, RK3588-class)
D — Form factor & pilot	Plush device, real users	Plush/toy form factor; pilot with families, not individuals
E — Refinement & launch	Iterate, certify, ship	Adds toy-specific certification (see below)

The practical implication: **most of the engineering risk was already retired by the eldercare program.** Stage A is done. Stage B for kids is mostly a UI and content-library exercise on top of an already-working core, plus building out the parental control surface described next — which is genuinely new work.

What's genuinely new for this audience

- The character library's **content** (original kids' characters) and **per-character memory partitioning** — the mechanism itself (persona library, Storytelling Mode) is shared, already-working eldercare architecture; see the sections above.
 - The parental control system (below) — this is the largest net-new engineering effort in this document.
 - Toy-safety and children's-privacy compliance, which is a materially different regulatory track than the eldercare product's (see Regulatory section).
-

Parental Controls

This is the feature area where kids' Companion diverges most from the eldercare product, and it needs to be genuinely robust — not a settings toggle bolted on afterward.

Monitoring

- **Conversation review.** Parents can read a transcript or summary of what their child talked about with each character, on the same transparency principle as the eldercare caregiver memory review (Stage B3 in the parent document).
- **Interest tracking.** The dashboard surfaces what topics and characters a child gravitates toward over time — useful for a parent wanting to understand their kid's interests, not framed as surveillance of their emotional state.
- **Alerting on boundary triggers.** If a child raises a topic that triggers a scripted redirect (see Content boundaries above), the parent is notified — this is the one case where the device does flag something to an adult, and it's disclosed to the child as a standing rule, not a secret.

Content and access controls

Available from **two surfaces**, matching how parents actually manage family tech today:

1. **Parent dashboard (app or web).** Full configuration: which characters are enabled, topic allow/deny lists, daily time limits, quiet hours, and conversation history.
2. **On-device admin/config mode.** A PIN- or pattern-gated mode accessible directly on the device itself, for parents who want to adjust settings without pulling out a phone — useful for grandparents or a second caregiver who may not have the app installed.

Both surfaces write to the same underlying config; there is one source of truth, not two configs that can drift out of sync.

Specific controls

Control	Description
Character enable/disable	Turn individual characters on or off per child
Time limits	Daily usage caps, with a gentle in-character wind-down rather than an abrupt cutoff
Quiet hours	No wake-word response outside configured hours (e.g., after bedtime)
Topic filtering	Parent-adjustable sensitivity for what triggers a redirect-to-adult response
Multi-child profiles	Separate memory and settings per child on a shared device
Data export / delete	Full transcript export or permanent deletion, parent-initiated, same encryption-at-rest guarantee as the eldercare product
Two-factor admin access	On-device config mode requires the same PIN as the app, to prevent a child from disabling their own limits

Regulatory Position

This is a materially different track from the eldercare product and needs its own review, not an inherited assumption:

- **This is a toy**, without the ambiguity the eldercare document leaves open. ASTM F963 and CPSIA apply in full: small-parts testing, flammability, battery safety, and toy-specific electrical standards.
 - **COPPA (Children's Online Privacy Protection Act)** applies in the US, even though the product is offline-first — COPPA's data collection provisions are triggered by any information collected *about* a child, not only by internet transmission, and Companion for Kids does record conversation data locally. Verified parental consent and the parent-transparency guarantees above are the compliance backbone; this needs dedicated legal review before any pilot involving real children's data, exactly as flagged as a long-lead item in the eldercare document's Stage D.
 - **No camera**, inherited directly from the eldercare decision, and if anything an even easier call for a children's product than an eldercare one.
 - **International equivalents** (GDPR-K in the EU, age-appropriate design codes in the UK) should be scoped in parallel if international sale is planned, since they're stricter than COPPA on several points.
-

Success Metrics (Kids-specific additions)

The eldercare product's technical metrics (time-to-first-audio, sustained tok/s, battery life, thermal limits) carry over unchanged. Added for this audience:

Metric	Target
Parent dashboard engagement	Parent checks in via dashboard or admin mode at least weekly during pilot
Boundary-trigger false positive rate	Low enough that legitimate conversation isn't constantly redirected — needs pilot tuning
Character memory consistency	Child-perceived "it remembers me" across characters, measured per-character
Parent trust rating	Parents report confidence in the transparency and control tools, not just satisfaction with the toy

Risks Specific to This Audience

Risk	Prob.	Impact	Mitigation
A child attempts to role-play unsafe or inappropriate content	Medium	High	Content boundaries baked into every character file; scripted redirect; parent alerting
A child bypasses parental controls	Low–Medium	High	Two-factor admin access; controls enforced on-device, not just in the app
COPPA or international child-privacy non-compliance	Medium	Critical	Dedicated legal review before any pilot with real children's data
Parent perceives the device as a monitoring tool rather than a companion for their child	Medium	Medium	Transparency framing disclosed to both parent and child; no hidden data collection
Character IP dispute (if original-character boundary is not maintained downstream)	Low, if principle holds	Critical	Original characters only, designed and owned in-house — no franchise impersonation, in this version or future ones

Version History

- v1.0 — Initial concept extension of Companion v2.0 (eldercare) to a children's audience. Character system redesigned around original, in-house personas rather than licensed franchise characters. Parental control system specified as net-new work. Toy-specific regulatory track (ASTM F963, CPSIA, COPPA) added.
- v1.1 — Corrected framing: the persona library and Storytelling Mode are not kids-specific inventions — both are already-built eldercare features (validated via a historical-figure persona and the Milestone 7 story-recall test), reused wholesale here. Narrowed "genuinely new for kids" down to character content, per-character memory partitioning, and the parental control system.