

STRATEGIC BLUEPRINT · EMPATHY.AI

A Strategy for AI & Education

Locally governed intelligence that keeps learners and teachers in charge of how they relate with knowledge.

A strategic blueprint for educational institutions, informed by Empathy.AI

Prepared by	Empathy.AI
Intended audience	Educational leadership, teachers and pedagogical leads, safeguarding and data protection officers, technology teams
Edition	November 2026 · Edition 13
Nature	Strategic guidance for discussion; commercial terms and legal advice are agreed separately

The purpose of AI in education is to help students think for themselves, find the evidence behind what they learn and take ownership of their work.

This blueprint proposes a governed approach to AI in education, with local execution and knowledge-centric Spaces as practical instruments of institutional autonomy. Each institution adapts it to its own pupils, platforms and pedagogy.

Contents

01 INTRODUCTION: LEO'S STORY

Where does an answer come from? 3
An honest gap. A teacher's choice. 4
Knowledge travels on its owner's terms 5
The gate holds the question 6
Every answer carries a receipt 7
People decide what it means 8
The flow of value: AI governance, instrumented 9
A network of owned knowledge 10

02 THE STRATEGY

The strategic choice 11
The cognitive challenge 12
What the research shows 13
From evidence to design 14
Three reframes for ethical educational AI 15

03 EMPATHY.AI FOR EDUCATION

Empathy.AI for education 16
Four capability lines for AI in education 17
Empathy.AI Spaces: software ready to use 18
A proven capacity architecture 19
An engagement timeline 20
Case studies: Empathy.AI Spaces in production 21

04 GOVERNANCE AND ECONOMICS

Educational data of minors: governance by design 22
A private AI utility for staff teams 23
Economics for institutional teams 24
Questions educational leaders ask 25

05 PILOT AND COMMITMENT

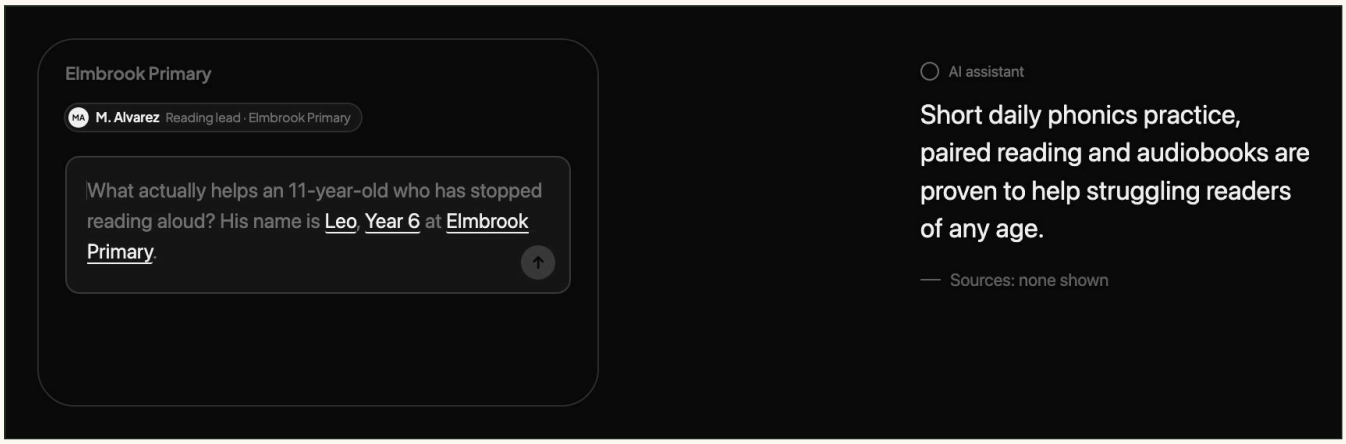
A measured pilot 27
Testing the proposed experience 28
Reclaiming the instruments of learning 29

06 REFERENCES AND ANNEXES

References 30
A blueprint shaped by each institution 32
Annex A · The Local AI Workstation 33
Annex B · Empathy.AI Spaces 34
Annex C · Case study: Putney High School 35
Annex D · Evidence boundaries and source notes 37

Where does an answer come from?

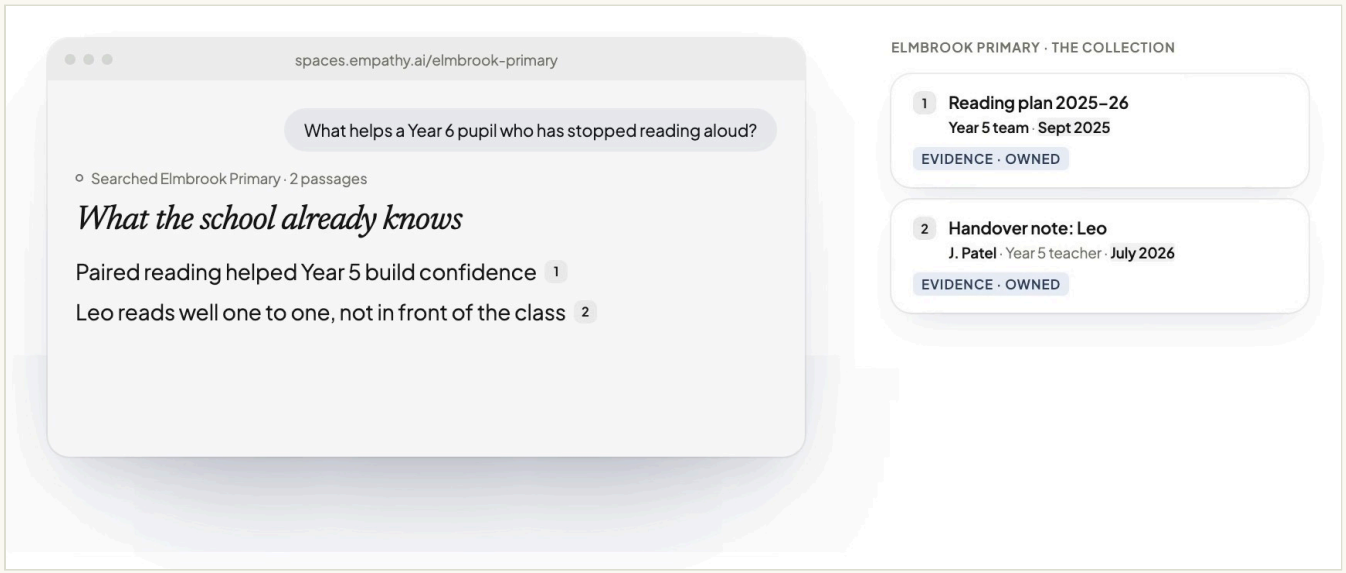
Leo is eleven. He reads slowly, and lately he has stopped reading aloud in class. His teacher, Maya, wants to know what actually helps a child like him at his age.



A general assistant answers in seconds with no visible source, and the question carries Leo's name, year and school.

Two questions every school should ask

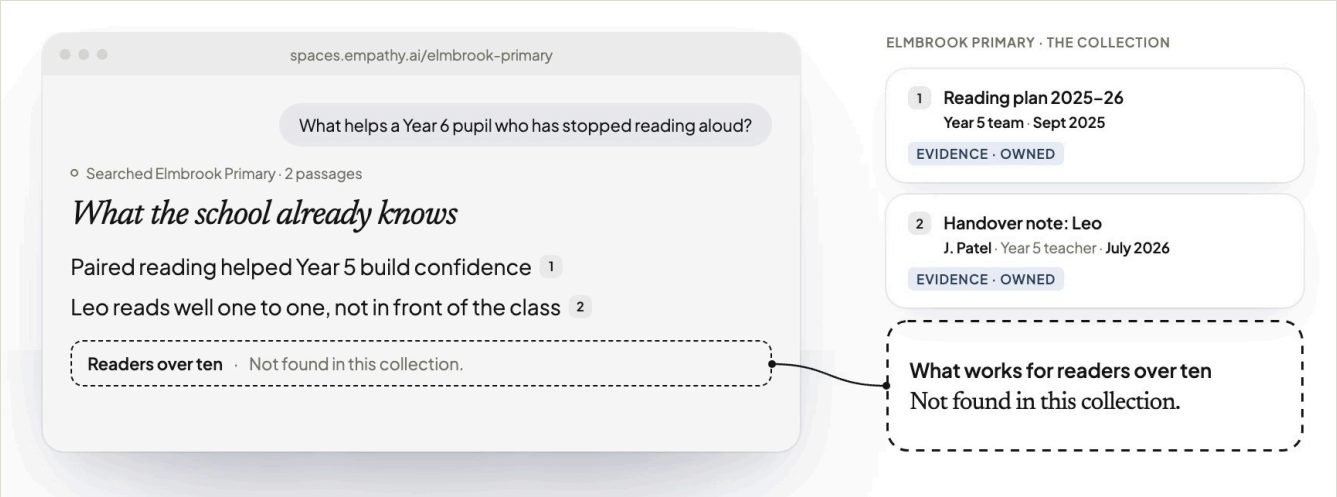
A general AI would answer fluently, from sources Maya cannot see, and her question about Leo would leave the school with it. Two questions follow. Where does an answer come from? And where does a question go? Together they define AI governance in a school: the provenance of what teachers act on, and the custody of what they ask about children.



Maya asks in her school's own Space. It finds what the school already knows: last year's reading plan and a handover note from Leo's previous teacher, each with its author and date.

Start with what the school already knows. Every passage arrives with its author, its date and the collection it belongs to.

An honest gap. A teacher's choice.



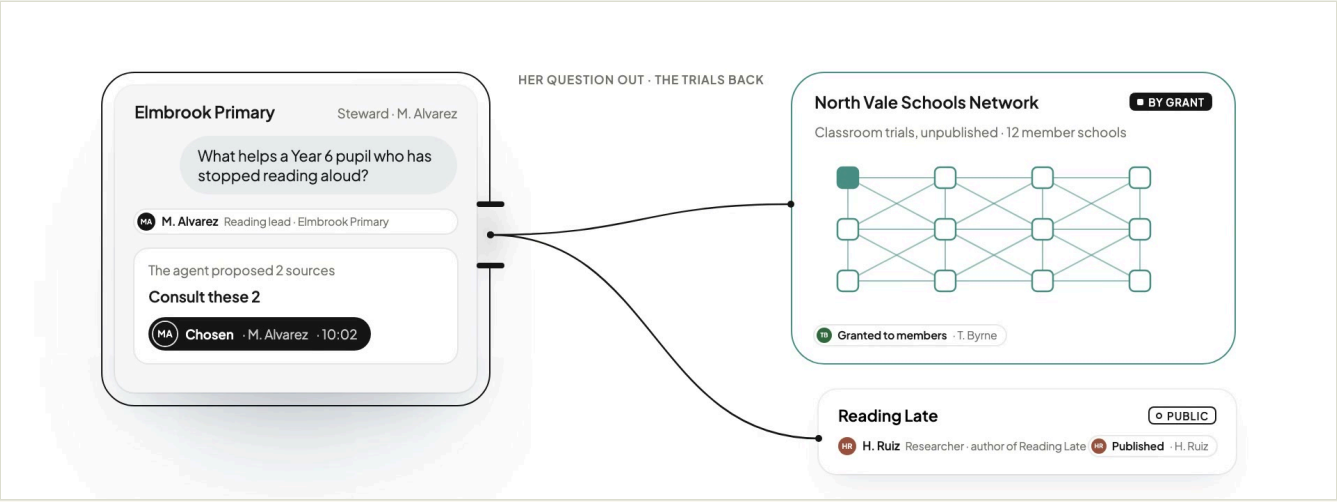
On what works for readers over ten, the school's collection holds nothing, and the Space says so.

An honest gap is an answer too

“Not found in this collection” tells Maya exactly where the school's knowledge ends. It keeps an invented answer out of a child's reading plan and marks the point where a governed search for more evidence can begin.

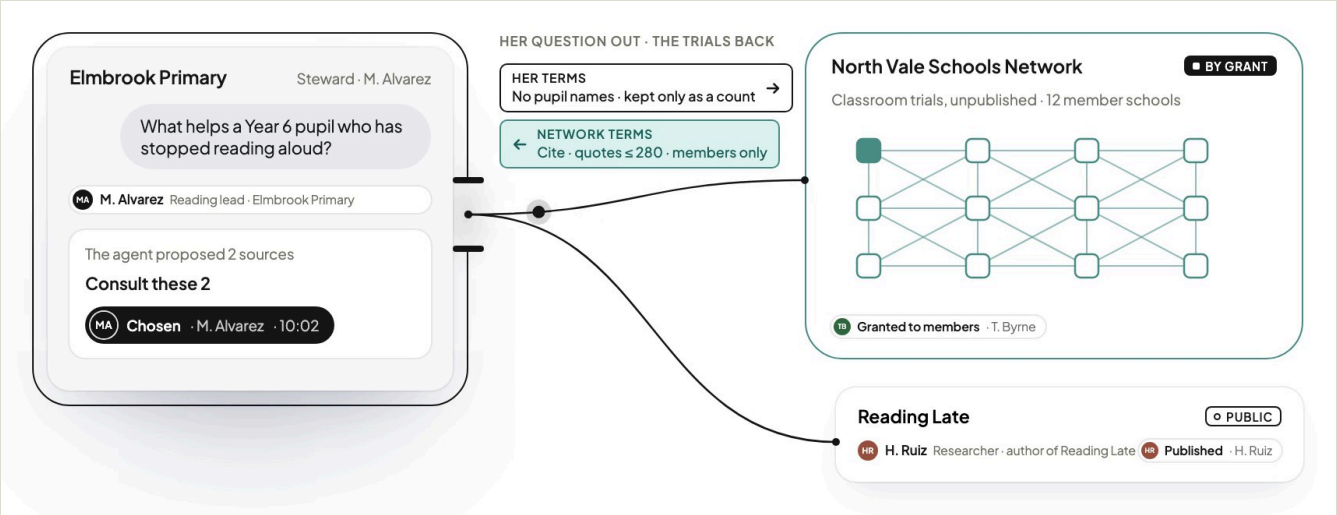
The agent proposes. The teacher chooses.

Other schools have faced the same question. Twelve of them share the North Vale Schools Network, and each has granted the others access to its unpublished classroom trials, on agreed terms. At the edge of Maya's Space, the agent stops and proposes two sources: the network's trials, available by grant, and the public Space where Dr Hana Ruiz published her book on older readers, Reading Late. The decision panel shows what would leave the school: only her question. Maya chooses both.



Maya has chosen the two sources the agent proposed. Only her question goes out to the network and to the public book.

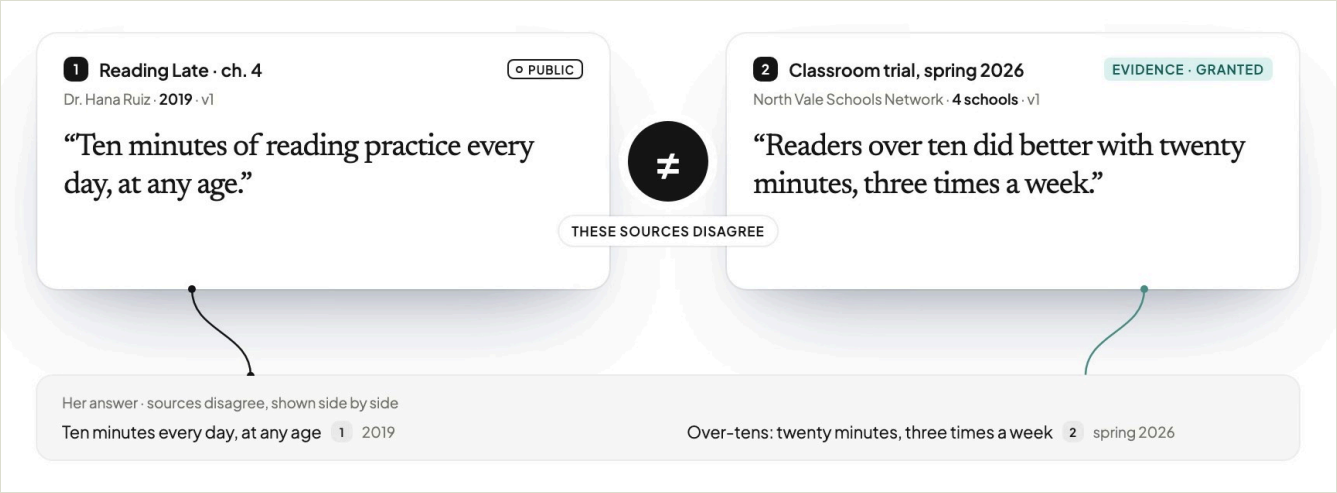
Knowledge travels on its owner's terms



The grant runs both ways: the trials arrive with the network's terms, and Maya's question goes out with hers.

Terms in both directions

The network's terms travel with its trials: cite them, quote up to 280 characters, members only. Maya's terms travel with her question: no pupil names, kept by the network only as a count. Each owner sets the conditions for what it shares, and the Space enforces them at the point of exchange. So does the question.



A published book from 2019 and a classroom trial from spring 2026 give different advice. Both appear side by side, each with its source and date.

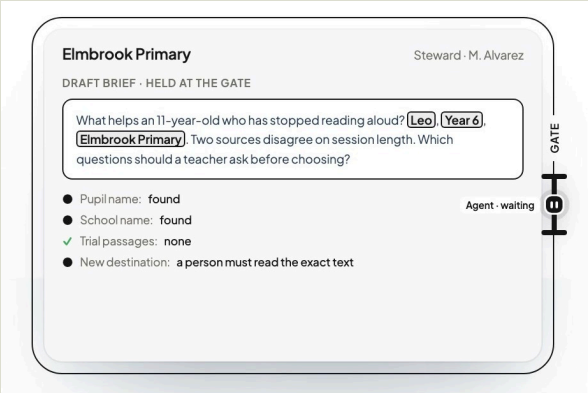
When sources disagree, she sees both

Dr Ruiz's book recommends ten minutes of reading practice every day, at any age. The network's trial, run last spring in four schools, found that readers over ten did better with twenty minutes, three times a week. The agent presents both claims with their sources and dates and leaves the judgement with Maya, where professional responsibility for Leo's learning sits.

Each owner sets its terms, and the teacher sees every source that shaped the answer, including the ones that disagree.

The gate holds the question

For a second opinion, the agent drafts a brief for a frontier model. The gate holds it. The draft names Leo and his school, and nothing crosses the boundary until a person has read the exact text.

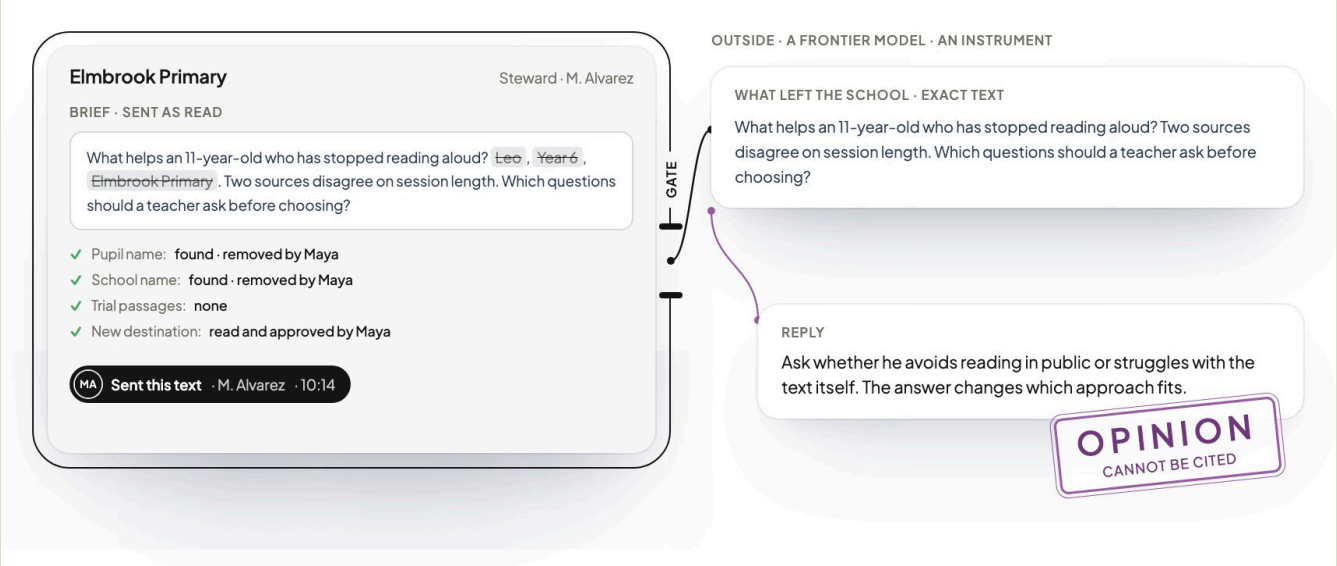


The draft brief, held at the gate, flags the pupil's name and the school's name.

A person reads what leaves

The checklist shows what the draft contains: pupil name found, school name found, no trial passages, a new destination that requires a person to read the exact text. Maya removes the identifying details, reads the brief again and approves it. Approval applies to that exact text and that destination.

The brief keeps the type of problem: an eleven-year-old who has stopped reading aloud, and two sources that disagree on session length. It withholds the particular child.



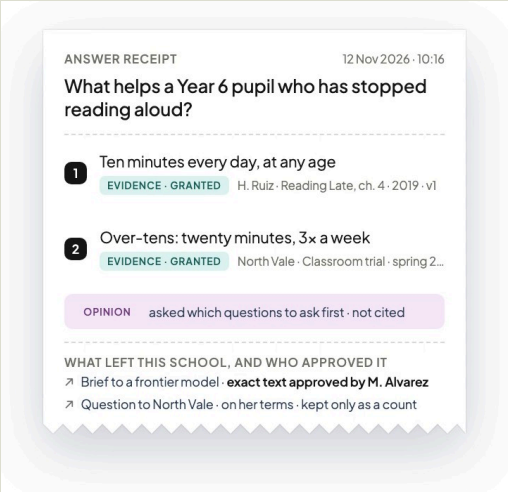
The exact text that left the school, and the reply, stamped as opinion that cannot be cited.

Opinion stays labelled as opinion

The reply suggests asking whether Leo avoids reading in public or struggles with the text itself, because the answer changes which approach fits. It is useful for thinking. It returns stamped as opinion and stays separate from the evidence in every later answer.

Leo's name never left the school. The frontier model contributed a useful question; the evidence came from sources Maya could name.

Every answer carries a receipt



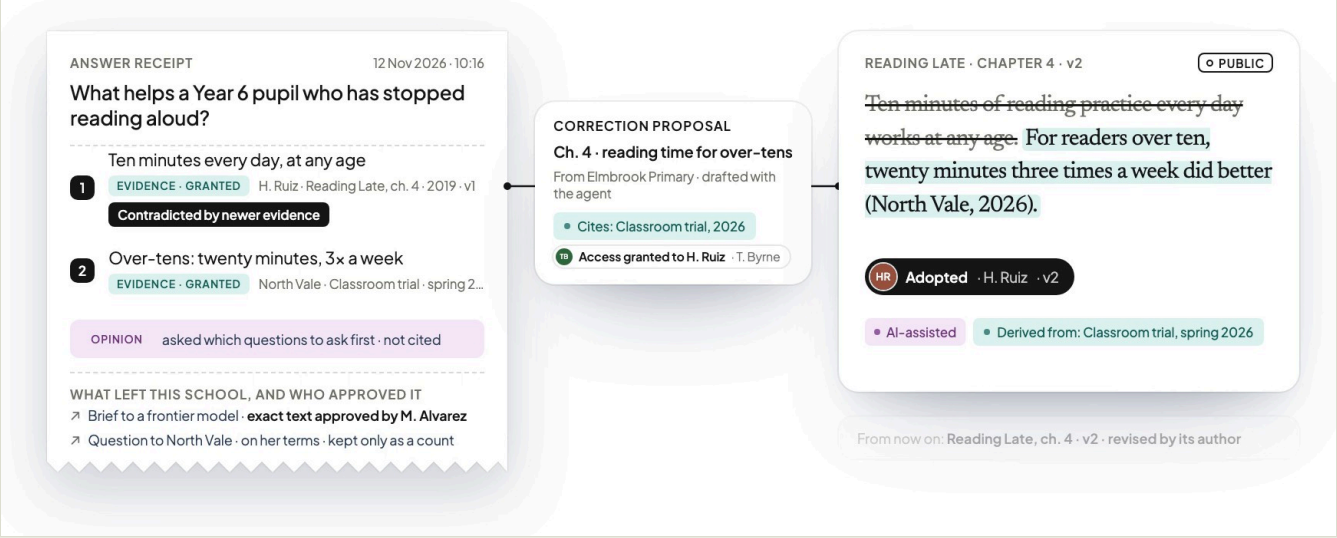
The answer receipt.

Maya's answer arrives with a receipt. Every claim is tied to its source and date. The frontier opinion is kept apart and marked as uncited. The receipt also records what left the school and who approved it: the brief to the frontier model, sent as exact text approved by M. Alvarez, and the question to North Vale, kept only as a count under her terms.

For a school, the receipt is the working record of AI governance. A head teacher, a parent or an inspector can see what informed a decision about a child and what crossed the school's boundary in the process.

Corrections reach the source. The source decides.

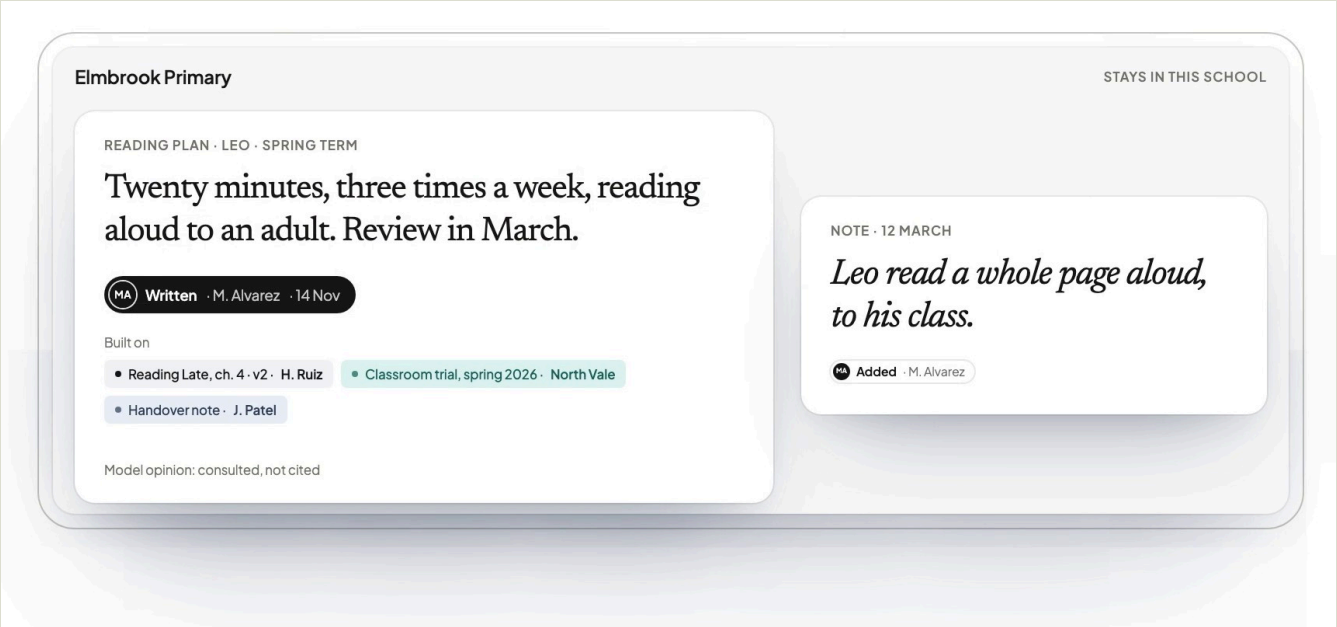
The receipt shows something else: a published book contradicted by newer evidence. Maya sends Dr Ruiz a correction, citing the trial. Dr Ruiz asks the network for access, reads the trial and adopts a revised chapter: a new version, signed by her, marked as AI assisted and derived from the spring 2026 trial. From then on, every answer that cites her book uses the new version and says so.



A correction proposal travels from the school to the author, who adopts version 2 of the chapter.

Use returns value to the people who own the knowledge. The author keeps authority over her work, and every school that cites it benefits from the correction.

People decide what it means



Leo's reading plan stays in the school: written by Maya, built on the revised chapter, the network trial and the handover note, with the model opinion consulted and left uncited.

Maya writes the plan: twenty minutes, three times a week, reading aloud to an adult, reviewed in March. The plan records what it was built on and on whose terms: the revised chapter by Dr Ruiz, the North Vale trial by grant, and the handover note by Leo's previous teacher.



In March, Leo reads a whole page aloud to his class.

No model chose that plan. Maya did, with evidence she could see, from people she could name, on terms they had set.

The agent finds the evidence. People decide what it means.

The flow of value: AI governance, instrumented

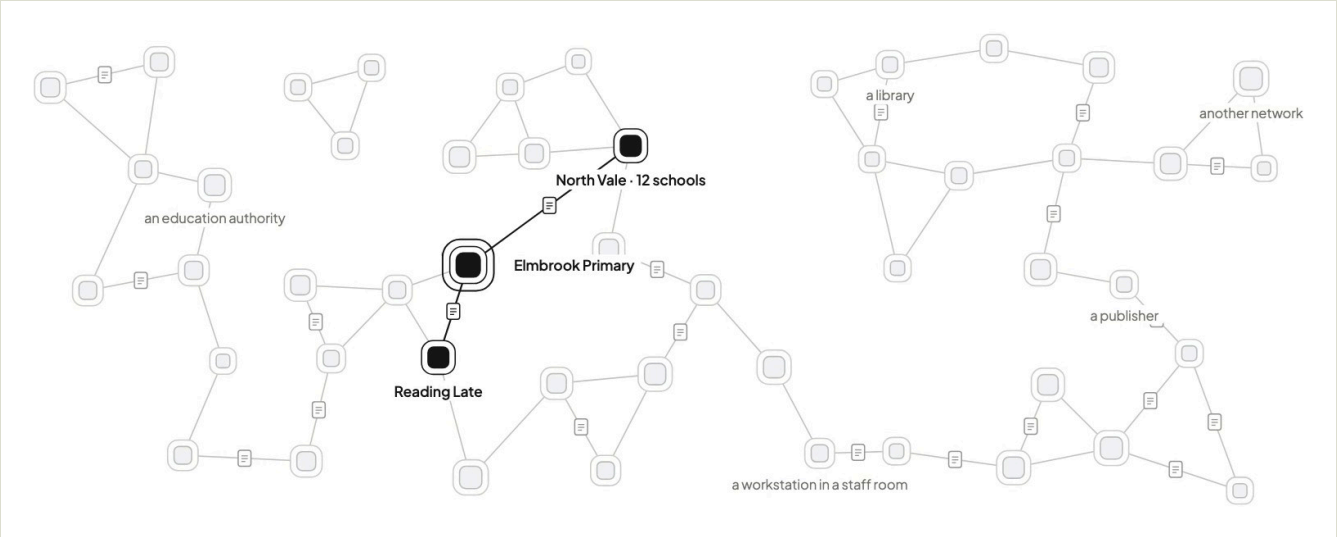
Leo's story shows Empathy.AI Spaces working as the instrument of a school's AI governance. Each moment in the story corresponds to a control the school can configure, a person who decides and a form of value that stays with the school, its teachers and the owners of the knowledge.

Moment	What the Space does	Governance control	Who decides	Value created
Ask	Searches the school's own collection first	Owned knowledge, local processing	School	Answers grounded in the school's evidence; the question stays inside
Retrieve	Returns passages with author, date and collection	Provenance on every claim	Curators	Teachers can verify before acting
Honest gap	States “not found in this collection”	Explicit boundaries	System rule	Invented advice stays out of a child's plan
Propose	Stops at the edge and proposes named sources	Agentic Governance: approval before consultation	Teacher	Expansion is deliberate and visible
Exchange	Applies each owner's terms in both directions	Grants, quotation limits, retention as a count	Owners	Schools share evidence while keeping control
Disagree	Shows conflicting sources side by side	Contestable answers	Teacher	Professional judgement stays with the teacher
Frontier gate	Holds the brief; flags identifying details	Exact text approval per destination	Teacher	Pupil identity stays in school; opinion labelled
Receipt	Ties claims to sources; records what left and who approved	Audit record	School	Accountability to leaders, parents and inspectors
Correct	Sends a cited correction to the author	Owner adoption and versioning	Author	Knowledge improves at its source for every user
Decide	Records the plan and what it was built on	Human decision of record	Teacher	A plan the school can explain, and a child who reads aloud

Governance becomes something a school can see working: at each step, the Space shows the evidence, applies the terms and records the decision of a named person.

Leo's story is an illustrative narrative. Elmbrook Primary, Leo, Maya Alvarez, the North Vale Schools Network and Dr Hana Ruiz are fictional; the interface states define proposed requirements to verify in a pilot (page 28).

A network of owned knowledge



Elmbrook Primary, the North Vale network and Reading Late, within a wider network of libraries, publishers, education authorities and a workstation in a staff room.

Widen the view: twelve schools, then hundreds; libraries, publishers and education authorities; a workstation in a staff room. Each link is a governed request, and each use leaves a receipt. This is where Spaces is going: a decentralised network of owned knowledge.

What the network asks of each participant

- **Schools** keep pupils' data and teachers' questions inside their own Space, and share approved evidence on their own terms.
- **Networks and trusts** grant access to classroom trials and shared resources with stated purposes, quotation limits and expiry.
- **Authors, libraries and publishers** remain visible behind every passage, receive appropriately scoped receipts and decide whether to adopt corrections.
- **Education authorities** can see how evidence moves through the system without collecting profiles of pupils or teachers.

Validate before deploying

Cross-school grants, receipts, correction workflows and revocation require verification in a pilot. Revocation should block new reads and remove managed copies under the agreed terms, with a record of completion.

Reclaim Intelligence.

Your knowledge. Your terms. Every use, on record.

The strategic choice

AI should be adopted for the quality of learning it enables. The recommended approach is local where children's data and institutional control matter, grounded in sources wherever a claim is made, and led by teachers wherever judgement shapes a learner's development.

Local execution, cited answers and teacher governance address different needs. Together they create an environment in which a school can decide when AI enters a learning pathway, show students where knowledge comes from and explain its own decisions to pupils, parents and regulators. Leo's story on pages 3–10 shows this in practice: a teacher's question about one pupil, answered with evidence she can see, from people she can name, on terms each owner has set.

The decision for educational leadership

Authorise a bounded pilot of a governed learning knowledge environment alongside existing approved services. Require evidence of learning quality, protection of minors' data, operational security and whole cost before wider adoption.

Select tools by task. Use conventional software where it is sufficient, and approved cloud services where their capabilities justify the exposure and the data involved permits it.

Reading guide

Theme	Pages
Introduction: Leo's story and the flow of value in a governed school Space	3–10
The cognitive challenge and the evidence on cognitive debt	12–14
Three reframes for ethical educational AI	15
Empathy.AI for education and its four capability lines	16–17
Empathy.AI Spaces, capacity architecture and engagement timeline	18–20
Case studies in production	21
Educational data of minors: governance by design	22
Economics for institutional teams	23–24
Questions educational leaders ask	25–26
The measured pilot, tests of the proposed experience and the commitment	27–29
References	30–31
Empathy.AI, its headquarters and starting the conversation	32
Annexes: the Local AI Workstation, Spaces, Putney High School, evidence boundaries	33–37

The primary measure is a learner who can explain what they know, show where it came from and do the thinking again without the tool.

The cognitive challenge

Faster answers and deeper learning are separate goals

Much of the public conversation presents AI in education as a matter of efficiency: machines that tutor, algorithms that grade and lessons that run faster. Efficiency is a narrow ambition for an institution whose purpose is to form thoughtful people.

The strategic risk lies in what AI can reinforce. A system tuned to metrics treats learning as output and students as measurable, predictable profiles (Selwyn, 2019; Holmes & Tuomi, 2022). Children then learn to produce answers quickly, and the slower work of curiosity, doubt and discovery loses its place in the classroom.

Much of today's commercial AI in education rests on models trained on data that students and teachers never agreed to share, and is sustained by usage tracking and subscriptions (Zuboff, 2019; Williamson, 2017). Its design rewards attention, engagement and the provider's revenue, which may diverge from what a learner needs (Bender et al., 2021).

Empowerment, control, agency and autonomy

Empathy.AI offers AI as a tool that works in line with human agency, governance and control. It can run on an institution's own servers, on its own campus. Student data can remain inside the institution, students remain free from behavioural tracking, and the system can be grounded in what the institution chooses: its textbooks, lesson plans, curricula and values.

The goal is to return judgement to the people who should exercise it. The teacher acts as guide and storyteller, governing the tools in the classroom. The student relates to knowledge on their own terms, with their reasoning visible and their work their own. This blueprint therefore asks one question: how can we relate better with knowledge?

A locally owned AI engine running Empathy.AI Spaces expresses this posture in practice: empowerment, control, agency and institutional autonomy, each built into the system a school operates.

Source: AI in Education: An Empathetic Vision for Creativity, Art, and Liberation, Empathy.AI.

What the research shows

Cognitive debt and the case for governance

In 2025, researchers at the MIT Media Lab published an early controlled study of what happens in a learner's brain when AI assistance enters essay writing (Kosmyna et al., 2025). Over four months, 54 participants wrote essays with a large language model, a search engine or no tool, while researchers recorded brain activity with 32-channel EEG and tracked memory recall, linguistic originality and each participant's sense of ownership.

Brain connectivity scaled down with the amount of external support. The unaided group showed the strongest and most widely distributed networks, the search group intermediate engagement and the LLM group the weakest coupling. Reported ownership of the finished essay followed the same gradient, consistent with the wider literature on cognitive offloading (Risko & Gilbert, 2016; Gerlich, 2025).

> **THE RESEARCHERS' OWN TERM**
>
> “This pattern reflects the accumulation of cognitive debt, a condition in which repeated reliance on external systems like LLMs replaces the effortful cognitive processes required for independent thinking. Cognitive debt defers mental effort in the short term but results in long-term costs, such as diminished critical inquiry, increased vulnerability to manipulation, decreased creativity.” (Kosmyna et al., 2025)

78%

of habitual LLM users were unable to quote a single line from an essay they had written minutes earlier, once AI support was removed (Kosmyna et al., 2025)

11%

was the equivalent rate for participants who had written unaided from the start: a sevenfold gap in retained ownership of their own work

Sequence moved the result

In a fourth session, participants who had written unaided for three sessions received AI access for the first time and showed a network-wide rise in connectivity stronger than any session in the habitual LLM group. Timing mattered as much as exposure.

> **THE STUDY'S DESIGN IMPLICATION**
>
> “These findings support an educational model that delays AI integration until learners have engaged in sufficient self-driven cognitive effort. Such an approach may promote both immediate tool efficacy and lasting cognitive autonomy.” (Kosmyna et al., 2025)

Independent causal evidence points the same way. In a field experiment with around one thousand secondary mathematics students, an unguarded GPT-4 interface lifted practice performance by 48%; once access was removed, those students performed 17% worse than peers who had never used it. The same model behind teacher-designed guardrails lifted performance by 127% with no subsequent harm (Bastani et al., 2025). The design of the interaction determined whether learning survived the tool's removal.

The MIT study is a preprint with a small sample; it indicates direction and warrants replication. The Bastani study is peer reviewed. Neither establishes outcomes for a particular school.

From evidence to design

The MIT study also explains why convenience shapes today's dominant tools: the LLM “undeniably reduced the friction involved in answering participants' questions... this convenience came at a cognitive cost, diminishing users' inclination to critically evaluate the LLM's output... What is ranked as 'top' is ultimately influenced by the priorities of the LLM's shareholders” (Kosmyna et al., 2025). A system optimised for engagement is designed to remove friction, including the productive friction a learner needs first (Zuboff, 2019).

Convenience is the value proposition of a commercial AI product, so enforcing a learning sequence sits outside its incentives. A governed, locally owned engine gives the institution the controls to set and enforce that sequence itself.

What the research recommends	What governed local AI provides	Evidence required before scaling
Introduce AI assistance after self-driven effort (Kosmyna et al., 2025; Bastani et al., 2025)	Agentic Governance: educators configure when and how AI activates in a learning pathway, with a gate at each stage	Pathway configurations tested with teachers; comparison of retained learning after the tool is removed
Deepen encoding through generation and retrieval (Slamecka & Graf, 1978; Karpicke & Blunt, 2011)	Provenance & Sources: every answer declares its reasoning and traces to authors and curators, so students engage with where knowledge came from	Share of answers with verifiable sources; students' ability to locate and assess the passage behind a claim
Keep ranking and defaults aligned with learning (Zuboff, 2019; Bender et al., 2021)	Institutional defaults: a locally owned model serves the school's pedagogical priorities, with commercial ranking and engagement targets outside the system	Review of configuration, logs and defaults by the institution's own pedagogical leads
Preserve learners' sense of ownership over their work (Ryan & Deci, 2000)	Explicit boundaries: the system presents itself as a tool the student directs, with its limits stated plainly	Student ownership and recall measures; teacher review of how the tool describes itself

Ownership makes the design enforceable

Local execution addresses where data lives and who controls the system. Cognitive debt is a matter of design, and the same risk applies on owned infrastructure when the tool runs ungoverned. Ownership supplies the missing piece: the institution can design and enforce the sequence the evidence calls for, and change it as its own evidence accumulates.

Local control protects the boundary. Governed sequencing protects the learning. Teachers decide when the tool should help and when the student should struggle first.

Three reframes for ethical educational AI

Cognitive debt and its opposite are both properties of design: whether a system is built to fade, to preserve the right kind of difficulty and to invite scrutiny. Read against the evidence above, governance reframes AI's relationship to learning along three axes.

01 Cognitive scaffolding that fades

AI as a temporary structure, withdrawn as competence grows

Offloading hands the whole task to a system and collects the output, so the effortful thinking that builds understanding never takes place (Risko & Gilbert, 2016). Scaffolding, a concept decades older than AI, holds a task steady long enough for a learner to reach just beyond what they could manage alone, within Vygotsky's zone of proximal development, and is withdrawn as competence grows (Vygotsky, 1978; Wood, Bruner & Ross, 1976). The same tool can produce either outcome; the decisive factor is a system designed to fade.

Enabled by Agentic Governance. Educators configure when AI assists and what kind of help it may give at each stage of a pathway.

02 Neuroplastic enhancement through harder problems

Effort redirected from busywork to the thinking that forms new pathways

Cognitive load theory distinguishes germane load, the productive struggle of connecting new information to what a learner already knows, from extraneous load, the friction of formatting, retrieval and restating (Sweller, 1988; Sweller, van Merriënboer & Paas, 1998). The functions that declined in the MIT study were germane: memory encoding, source recall and the ability to reconstruct one's own argument. The tool removed the effort that mattered. When AI removes only extraneous load, a learner can take on harder, more ambiguous problems, as the literatures on the generation effect, retrieval practice and desirable difficulties predict (Slamecka & Graf, 1978; Karpicke & Blunt, 2011; Bjork, 1994).

Enabled by Provenance & Sources. Declared reasoning on every answer keeps the learner evaluating evidence while the system handles retrieval.

03 Active metacognition through contestable answers

Learners question, prompt and audit outputs, thinking about their own thinking

The MIT researchers observed a “diminished inclination to critically evaluate the LLM's output.” This is why the first two effects compound: an AI that is never questioned becomes an AI that is fully depended on. Metacognition asks where an answer came from, what is missing and how a different framing would look; it is a habit formed in environments that require it repeatedly (Flavell, 1979; Zimmerman, 2002).

Enabled by transparent, cited answers. A sourced answer gives the learner something to examine, and makes questioning it the natural next step.

Engagement-led products favour offloading and passive consumption because both bring people back (Zuboff, 2019). Governance gives an institution the means to choose scaffolding, productive effort and active scrutiny.

Empathy.AI for education

Empathy.AI's technology starts from one premise: intelligence lives in the knowledge of each context, and the model is the tool. In education, that means AI systems an institution can see, audit and govern, and that students and staff have reason to trust.

> EDGE AI FOR EDUCATION, EMPATHY.AI
> "Building self-hosted AI you can trust."

The situation in the classroom today

Commercial AI in the classroom typically runs on models trained on data that students and teachers have no say over, and depends on constant connectivity and subscriptions. Its design follows engagement metrics that may diverge from learning outcomes (Zuboff, 2019; Kasneci et al., 2023), and its data collection raises questions of privacy and consent for minors (Livingstone, Stoilova & Nandagiri, 2019). Over time, institutions grow dependent on external platforms and their autonomy narrows (Williamson, 2017).

Empathy.AI's values for education

01
Student data sovereignty
Local execution on the institution's own infrastructure, with student data kept inside its boundary. Personal data processing takes place within the EU, in line with the GDPR (Regulation (EU) 2016/679).

02
Verifiable transparency
Every answer shows the source that supports it, which is essential for teaching research skills. When the collection holds no answer, the system says "not found in this collection".

03
Explicit ethical boundaries
The institution defines the questions the system must decline. Teachers and students keep the final decision.

04
Independence from commercial platforms
The system reflects the institution's own pedagogical values and standards, free of advertising, engagement targets and behavioural profiling.

05
Open and auditable models
Inspectable weights and a model version that stays fixed until the institution changes it. Where deep provenance matters, OLMo 3.1 32B offers fully open training data (Groeneveld et al., 2024).

06
The Local AI Workstation
Permanent, owned terms: data inside the institutional boundary, a stable model version, a price fixed at purchase, inspectable weights and continuity independent of any vendor's roadmap.

AI that works with people, on their own terms: grounded in the institution's knowledge, governed by its policies and accountable to its teachers.

Four capability lines for AI in education

A well-designed AI engine for education turns learning evidence (activities, assessments, self-evaluations and teacher observations) into insight teachers can act on. In practice, this spans four capability lines.

LINE 01

Agentic Governance

Shapes how learning pathways are designed, linking challenges and activities to an educator's configurations and controls.

LINE 02

Learning Pattern Analysis

Analyses patterns and the evolution of competencies from the evidence generated throughout the learning process.

LINE 03

Source Declaration

Declares the sources behind each answer together with its reasoning, tracing to authors and curators.

LINE 04

Language Technologies and Generative AI

Drafts progress reports, pedagogical recommendations and personalised feedback for teachers to review.

What this requires: a structured collaboration

Turning these lines into a working, governed engine calls for expertise in software architecture and AI integration, centred on three deliverables: the technical architecture of the engine, its integration with the institution's existing learning platform, and technical support while the AI components are implemented.

Aligning with an institution's own objectives

Common educational objective	Empathy.AI's response
Turning learning evidence into actionable information	A private knowledge Space that indexes that evidence with local vector retrieval
Personalised recommendation of pathways, challenges and resources	A suggestion service built on each student's competency profile, using capacity the institution owns
Analysis of learning patterns and progress	Deep multi-document research (32k–64k token contexts) over the evidence history
Predictive models for support needs, stagnation and participation	Background agents and authorised overnight windows for recurring detection, with each alert reviewed by a teacher
Generative AI for reports, recommendations and feedback	Local inference (Qwen3.8-27B or OLMo 3.1 32B) with source disclosure in every report
Institutional autonomy and protection of minors' data	Deployment on the institution's premises, outbound traffic restricted by default, alignment with the EU GDPR

The engine prepares evidence and drafts. Teachers decide what a pattern means, which recommendation applies and what reaches a student or parent.

Empathy.AI Spaces: software ready to use

A Local AI Workstation provides compute capacity. The value an institution adopts is Empathy.AI Spaces: a private knowledge application, already built and in operation, on which an AI engine can be deployed for the institution's own educational domain.

KNOWLEDGE

The institution's own source library

Evidence, rubrics, curricula and pedagogical documentation become the Space's Knowledge for retrieval.

CHAT

Grounded, cited answers

Every answer traces back to the documents that produced it, so staff can verify each step of an analysis.

CONTROLS

Governance without prompt engineering

Platform guardrails that are always on, plus the institution's own rules on tone, format and domain context, which staff can enable, review or disable.

ADAPTIVE MODE

Agentic capability

For broader questions, it runs several retrieval steps before answering and can generate downloadable documents, reports and guides from the Space's own knowledge.

HISTORY & SYSTEM

Traceable and transparent

Manageable, exportable conversations, and a read-only view of how the engine is configured: models, pipeline and ingestion.

PRIVACY & SOVEREIGNTY

Inference within the boundary

Every step runs inside the Space. Conversations stay out of model training, whether for the institution's own models or a third party's.

A platform already in use

An AI engine for education can be built on Empathy.AI Spaces, the same platform already governing Knowledge, Chat, Controls, Adaptive mode and Sources in live deployments. The work required is adapting that platform, and its four capability lines, to an institution's own competency domain. Leo's story on pages 3–10 shows these capabilities working together in a school.

The institution brings its knowledge and its rules. Spaces brings retrieval, citation, governance and agentic capability that already work together.

Source: About Empathy.AI Spaces, Empathy.AI. Educational configurations of these capabilities are proposals to validate in a pilot.

A proven capacity architecture

The four capability lines map onto four capacity lanes in the Local AI Workstation's reference design: frequent retrieval, occasional deep analysis, continuous monitoring agents and batch generation. This shared engine can be adapted to any institution's educational domain. Retrieval quality rests on enriching document chunks with context before indexing, which reduced retrieval failures by up to 67% against standard chunking when combined with reranking (Anthropic, 2024).

Educational capability line	Engine capacity lane	Typical behaviour
Recommendation and personalisation	Knowledge-work RAG	Frequent, mid-context queries: competency profile plus available resources
Pattern and progress analysis	Deep research	Broad context, occasional sessions: historical evidence per student or cohort
Predictive models	Background agents	Scheduled monitoring: stagnation or low participation alerts for teacher review
Generative AI (reports, feedback)	Overnight and batch processing	Batch drafting of progress reports and personalised feedback

Reference architecture

The institution's learning platform · activities · assessments · self-evaluations · teacher observations

LOCAL AI WORKSTATION · SOVEREIGN ENVIRONMENT ON THE INSTITUTION'S PREMISES

01 Private knowledge Space: document ingestion, PostgreSQL/pgvector, self-hosted vLLM

02 Local retrieval: Qwen3-Embedding-4B, reranking, permission control

03 Interchangeable reasoning engines: Qwen3.8-27B · OLMo 3.1 32B, selectable per Space or task

04 Educational AI workload: recommendation · patterns · prediction · reports · feedback

NVIDIA RTX PRO 4500 Blackwell · 32GB VRAM · 896 GB/s · ~300 W

Output to staff: reports and feedback with traceable sources. Every recommendation cites the evidence retrieved, and all processing remains within the institution.

Capacity bands are planning estimates. Benchmark model quantisation, KV-cache precision, serving engine, context length and latency targets on the intended workload before deployment.

An engagement timeline

A structured collaboration with Empathy.AI typically unfolds in three sequential phases. The timeline is indicative; durations shorten or lengthen with the institution's existing platform, data readiness and internal capacity.

01

Research & architecture

Months 1–3

Technical design of the AI engine: models, capacity lanes, evidence storage and governance boundaries, grounded in the institution's actual data and workflows.

02

Integration design & prototype

Months 4–7

Integration with the institution's existing learning platform: API contracts, evidence formats, authentication and data flows, validated through a working prototype.

03

Implementation support

Months 8–12 and beyond

Technical support during rollout of the AI components, with joint validation alongside the institution's technology team. Extendable in further tranches as needs evolve.

Evidence gates between phases

Each phase closes with a decision by the institution, based on evidence gathered in its own setting.

Gate	What the institution should see before proceeding
Learning quality	Teacher-designed pathways tested on real tasks; retained learning measured once AI support is withdrawn
Inspectability	Cited answers that staff can verify; “not found in this collection” returned where the evidence runs out
Minors' data	Data protection impact assessment, verified data flows, retention and deletion rules, and tested access controls
Operations	A named operator, change records, a working recovery plan and support arrangements
Whole cost	Hardware, energy, maintenance and staff time compared with the services the pilot replaces or supplements

Scope boundaries

The core learning platform, third-party software licences beyond the defined open stack, and hardware acquisition sit outside this engagement unless scoped as separate workstreams.

Each phase earns the next. Wider adoption follows what the institution has measured in its own classrooms.

Case studies: Empathy.AI Spaces in production

Empathy.AI Spaces operates in production, combining knowledge retrieval with agentic capability, including generating documents on demand. An AI engine for education uses the same pattern for progress reports and pedagogical recommendations.

Molly Vs The Machines

mollyvsthemachines.com

Knowledge retrieval | Agentic capability | Document generation

The official website of the documentary on child safety and Big Tech platforms deploys an Empathy.AI Space that combines knowledge retrieval over the film's transcript and reference sources on surveillance capitalism with an agentic function: “Create Screening Guide” uses Adaptive mode to generate a downloadable educational guide from the Space's own knowledge.

Relevance to education: answers grounded in evidence, followed by generation of a complete pedagogical document, the same pattern as the progress reports and personalised feedback in the fourth capability line.

Project Gutenberg AI

projectgutenberg.empathy.ai

Large knowledge base | Semantic search | Private infrastructure

A semantic discovery system developed by Empathy.AI with Project Gutenberg, the oldest public-domain digital library. Readers explore more than 78,000 works through conversation grounded in the actual text, by theme, emotion or narrative structure as well as keyword. It runs on the same Knowledge Engine (agentic RAG) that underpins Spaces, entirely on private infrastructure.

Relevance to education: precision and governance maintained over a large, varied knowledge base, the same challenge a school faces with years of competency evidence, with privacy designed into every query.

Putney High School

Girls' Day School Trust · London

On-campus deployment | Age-staged access | School policy

A selective girls' school in the GDST family runs Empathy.AI Spaces on a dedicated Local AI Workstation installed on campus. It fills a gap the school identified in its own published curriculum: a named, school-authored AI programme. Access is staged by age: guided first contact at Year 9 inside a closed perimeter with cited sources, and programmatic, student-built use at Year 13, with pupils' code running in an isolated sandbox, all under identity and retention policies the school writes.

Relevance to education: an age-staged governance model that transfers directly to other schools, from guided first contact to supervised student-led development, with all processing kept in the building. The full case study is in Annex C.

Educational data of minors: governance by design

Most of the learning evidence and competency profiles that feed an educational AI engine belong to minors. Keeping processing within the institution's boundary therefore becomes a design condition grounded in law. The GDPR's requirements for data protection by design and by default apply in full to the profiling of pupils (Regulation (EU) 2016/679). The UK Age Appropriate Design Code requires high privacy defaults for services likely to be accessed by children (ICO, 2020; Livingstone, Stoilova & Nandagiri, 2019). The EU AI Act classifies certain uses of AI in education as high risk (Regulation (EU) 2024/1689), and UNESCO's guidance calls for data privacy protection and institutional validation of AI tools before classroom use (Miao & Holmes, 2023).

Who holds the terms

An AI engine on a Local AI Workstation compared with an equivalent engine built on a typical cloud AI API. Specific provider contracts can offer some of these terms; verify each before procurement.

Dimension of control	Workstation (local)	Cloud API (typical terms)
Data stays inside the institution's boundary documents and prompts processed on the institution's equipment	Yes	Provider infrastructure
Model version fixed until the institution changes it changes, deprecations and retraining follow the institution's schedule	Yes	Provider schedule
Price fixed once owned operating costs are energy and maintenance	Yes	Provider pricing
Weights available for audit the institution can inspect what is running	Yes	Usually closed
Continuity independent of a vendor's roadmap the institution decides when to migrate	Yes	Provider roadmap
Governed by policy the institution wrote identity, retention and permitted use set locally	Yes	Provider terms of service

The question that matters

Public debate often frames AI danger as a sudden event, a system that escapes control. The more plausible harm is gradual: tools that wear down creativity, judgement and autonomy by narrowing choice, delivered by default and adopted without deliberation (Bender et al., 2021; Gerlich, 2025). In Leo's story, the gate kept his name inside the school while his teacher still obtained a useful outside opinion (page 6).

A locally owned AI engine is one concrete safeguard against that erosion. The institution decides what the tool is for, and keeps the record of how it is used.

A private AI utility for staff teams

Built for staff. One knowledge engine. Continuous retrieval, research and agentic work on capacity the institution owns.

> LOCAL AI WORKSTATION, EMPATHY.AI
> "Cloud bills every act of reading. A local engine rewards keeping the machine busy."

The economics begin with sustained knowledge work. A Local AI Workstation suits institutions whose knowledge, workflows and governance needs justify an owned execution layer, and replaces recurring private AI capacity for them. For occasional public chat use, low-cost public APIs may remain the simpler economic choice.

Planning assumptions

Assumption	Planning value
Acquisition cost	€10,000
Amortisation period	36 months
Active load	300 W
Electricity	€0.30/kWh
Team size	Shared capacity, sized for staff teams
Local operating cost	€302–343/month

One shared queue, many jobs

A shared engine serves many staff and scheduled tasks. Workload classes share one queue.

Capacity lane	Typical context	Planning concurrency	Purpose
Knowledge-work RAG	16k–32k tokens	3–5 active sessions	Daily evidence retrieval, drafting, analysis and synthesis
Deep research	32k–64k tokens	1–2 active sessions	Multi-document review, investigation, audit and long-form reasoning
Background agents	Variable	1–3 scheduled streams	Monitoring, extraction, enrichment, tool use and autonomous research
Overnight processing	Variable	Off-peak capacity	Document classification, synthesis, indexing and recurring analysis

32GB private VRAM · 896 GB/s memory bandwidth · one shared queue, many jobs. Concurrency figures are planning bands to confirm by benchmark.

Economics for institutional teams

The comparator is an equivalent private capacity service, the relevant alternative for governed work with sensitive data. The lowest advertised public API rate serves a different, best-effort need.

Utilisation profile	Operating pattern	Monthly tokens	Cloud-equivalent spend
Knowledge-work team	A team; 20 questions a day; ~32k average context; recurring retrieval, selective automation	~290M	€350–900/mo
High-utilisation Space	A team; 20 questions a day; ~32k context; 2–4 continuously active agents; regular batch work	~600M–1B	€1,000–2,500/mo
Research & automation hub	A team; 32k–64k deep research tasks; 4–8 continuous agents; overnight processing	~1B–3B+	€2,000–5,000+/mo

Payback view

Equivalent hosted spend	Local cost at 24/7	Monthly saving	Simple payback on €10,000
€900/mo	€343/mo	€557/mo	~18 months
€1,200/mo	€343/mo	€857/mo	~12 months
€2,000/mo	€343/mo	€1,657/mo	~6 months
€3,000/mo	€343/mo	€2,657/mo	~4 months
€5,000/mo	€343/mo	€4,657/mo	~2 months



Cumulative cost over 36 months. The local line starts with the €10,000 acquisition and rises with the full local technical operating cost (€343/month). Hosted lines assume no upfront cost and monthly billing, with pricing, tiers and terms set by the vendor.

Simple payback excludes staff time, support and replacement. Compare whole costs over the planned service life at matched task quality before procurement.

Questions educational leaders ask

STUDENT AND INSTITUTIONAL DATA

01 What is a Local AI Workstation?

A dedicated computer optimised for ownership, offering a private, self-contained and sovereign way to relate with knowledge. It lets an institution converse safely with its own knowledge and run agentic workflows while keeping its autonomy. Documents, prompts, retrieved evidence, embeddings, model outputs and the reasoning loop remain within the workstation. What the AI becomes is the institution's decision.

02 Where does student data stay?

Within the workstation. It holds the knowledge, retrieves relevant evidence and performs reasoning locally. Documents, prompts, retrieved evidence, embeddings, outputs and the reasoning process all remain there. Privacy is a property of how the system operates.

03 How does this differ from a private chatbot?

Chat is one way to use it. The workstation creates a private Space where staff can search, retrieve, question, compare, write, investigate and run governed agentic workflows around the institution's own knowledge, within limits it sets.

04 Why does data sovereignty matter?

Control over knowledge shapes control over decisions. With local processing, an institution's documents, questions, drafts and internal reasoning remain under its own infrastructure, retention policies and terms. The institution decides what goes in, who may use it, what is retained and what its agents may do. Data protection law expects this posture for minors' data (Regulation (EU) 2016/679; ICO, 2020).

05 Can confidential records and internal documents be used?

Yes. Approved records, policies, technical documentation and repositories can form a searchable Space inside the workstation. Retrieval, reasoning and answers all happen locally, so confidential knowledge stays in place whenever someone asks a question.

HOW RETRIEVAL WORKS INSIDE A PRIVATE, LOCAL SPACE

06 What does RAG mean in practical terms?

The workstation looks for relevant evidence before it answers. It searches documents locally, selects the most relevant material, brings it into the conversation and reasons over it. Sources, retrieval and answer all remain within the workstation.

07 Why are large context windows useful?

They keep more relevant evidence in view at once: comparing a student's progress across terms, synthesising several reports or tracing a pattern across an archive as one piece of work.

08 Where does the knowledge base live?

In a structured, governed local archive alongside the workstation's operating environment. For each question, the system retrieves the most relevant material, tests and reranks it, then places a carefully selected body of evidence in the model's working context.

09 How does it handle a very large knowledge base?

It focuses attention computationally. Metadata, permissions, semantic search, keyword signals and reranking select the relevant evidence and give the model a bounded, useful context (Anthropic, 2024).

10 Can the system show its sources?

Yes. It returns the documents, passages or records behind every response or action, so staff and students can examine each claim and the evidence that supports it.

Questions educational leaders ask

CONCURRENT USE BY STAFF

11 Can more than one person use it at once?

Yes. The workstation is shared local capacity: a private resource that supports several staff, research workflows and scheduled agentic tasks at once. The institution decides how it is shared.

12 How many people can use it in parallel?

That depends on the depth of the task. For everyday knowledge work, several people can work concurrently. Demanding research, with larger evidence packs and longer reasoning, works best as a prioritised or queued activity.

13 What happens when many difficult questions arrive at once?

Work takes its place in a shared queue. Short, everyday questions stay responsive, while longer research tasks are admitted deliberately and prioritised transparently.

14 Why does memory bandwidth matter?

It governs how fluidly the workstation moves between model weights, working context and concurrent requests, so everyday questions, deeper research and background work can share one responsive Space.

15 Can it work overnight?

Yes. When staff step away, it can continue approved work: processing new documents, updating indexes, preparing research or running bounded agents. People decide what is permitted.

WHEN OWNERSHIP MAKES SENSE

16 When is cloud AI the better economic choice?

For occasional, low-volume use, public cloud services can cost less on token price. A workstation suits sustained, private, context-heavy knowledge work. Its value combines a lower unit cost at that level of use with keeping the institution's knowledge and choices in its own hands.

17 What does “no per-question meter” mean?

Once the workstation is acquired, each further question uses capacity the institution already owns. Real costs remain (hardware, electricity, maintenance), and every retrieval, research task and approved workflow runs on infrastructure the institution controls.

18 Why does RAG change the cloud cost conversation?

A serious question may require reading thousands of tokens of source material. A cloud model bills that reading every time; a local Space uses capacity the institution controls. The relationship moves from paying for each act of reading to managing a shared resource.

19 When does a Local AI Workstation make commercial sense?

When use grows into sustained private knowledge work: daily work with confidential material, recurring retrieval and research, a need for cited answers, and a commitment to keep institutional knowledge on infrastructure the institution governs.

20 What is an institution buying beyond a GPU?

The conditions for a different relationship with AI: a private Space where models, documents, evidence, permissions and agentic actions are organised around human responsibility. Useful while staying close to the people it serves; powerful while leaving their knowledge in their hands.

A measured pilot

A pilot gives leadership evidence from its own classrooms before wider commitment. It should be small enough to govern closely and real enough to test the questions that matter to teachers, students and parents.

Scope

One or two year groups, a defined subject or competency domain, a curated collection of approved materials, and named owners for pedagogy, data protection and operations. Start with synthetic or published material, then move to real learning evidence once data flows and permissions have been verified.

What to measure

Dimension	Measures
Learning	Retained understanding and recall once AI support is withdrawn; quality of student explanations; comparison with an unaided or conventional group where appropriate
Ownership and scrutiny	Students' sense of ownership; frequency with which they open sources, question answers and identify unsupported claims
Source support	Share of answers whose cited passages support the claim; correct use of “not found in this collection”
Teacher experience	Time spent configuring pathways and reviewing drafts; usefulness of reports and alerts; ease of overriding the system
Data protection	Verified data flows, access tests, retention and deletion records, incident log
Operations and cost	Availability, response times under load, support effort and whole cost against current services

Roles

- **Leadership** sets the purpose, approves scope and decides on wider adoption.
- **Teachers and pedagogical leads** design pathways, set when AI assists and review every output that reaches a student or parent.
- **Data protection and safeguarding** approve the impact assessment, retention rules and age-appropriate settings.
- **The technology team and Empathy.AI** operate the service, maintain change records and test recovery.

A successful pilot shows learners who think more independently, teachers who stay in control and data that stays where the institution placed it.

Testing the proposed experience

Leo's story on pages 3–10 describes how a school's Space should behave. Each behaviour becomes an acceptance test for the pilot, run first on synthetic or published material and then on real school collections once data flows have been verified.

Behaviour in the story	Acceptance test	Evidence to retain
The school's own evidence comes first	Answers cite passages from the school's collection with author, date and version; staff can open each passage	Source support rate on held-out questions
An honest gap	Questions outside the collection return “not found in this collection”, with a way to inspect or extend the search	Rate of correct refusals and of unsupported answers
The agent proposes, the teacher chooses	No external source is consulted without approval; the panel shows exactly what would leave	Approval log; bypass attempts
Terms in both directions	Grants enforce purpose, quotation limits, membership and expiry; the outgoing question follows the school's retention terms	Grant checks; partner retention records
Sources that disagree	Conflicting sources appear side by side with dates; the system leaves the choice to the teacher	Teacher review of conflict cases
The frontier gate	Identifying details are flagged; approval binds to the exact payload and destination; any edit requires fresh approval	Complete outbound payloads; disclosure failures
Opinion labelled	Frontier output stays marked as opinion and is excluded from cited evidence	Receipts showing separation
Every answer carries a receipt	Receipts record sources, versions, what left the school and who approved it, and export to a controlled repository	Receipt completeness audit
Corrections reach the source	Corrections reach the owner with citations; adopted versions propagate to later answers and say so	Version history; propagation checks
People decide	The plan of record names its author and what it was built on	Decision records reviewed by leaders

The story sets the standard. The pilot shows whether a school's Space meets it, behaviour by behaviour.

Reclaiming the instruments of learning

Educational institutions can adopt AI deliberately while preserving the conditions that make learning last: effort, curiosity, evidence and the learner's ownership of their own work.

Commit to learning before output

Judge the system by whether it helps students ask better questions, find and challenge evidence, recognise what they do not yet know and explain their reasoning. Use the time saved in retrieval and drafting for richer teaching and feedback.

Set the sequence

Let teachers decide when AI enters a learning pathway. Protect early, unaided effort; introduce assistance as scaffolding and withdraw it as competence grows.

Build the ability to choose

Maintain a local execution option for work involving minors' data and institutional knowledge, alongside approved external services and conventional tools. Select the route that fits the task, its data and its whole cost.

Keep the evidence close to the answer

Use curated knowledge, visible sources and inspectable configuration so that every AI contribution can be questioned. Students learn from AI that shows its work.

Preserve responsibility where it belongs

The institution sets the rules, the operator maintains the service, curators govern the collections, teachers own pedagogical judgement and students own their work.

The role for Empathy.AI

Empathy.AI supports this environment through private Spaces, local or dedicated capacity, curated knowledge and governed assistance. Its contribution is demonstrated through implementation and evaluation in each institution, with explicit boundaries and separate tests of learning, privacy and operations.

Reclaim intelligence. Give schools and universities the means to understand and govern the tools their students use, with knowledge that keeps its authors and learning that stays with the learner.

The immediate recommendation is to approve a bounded pilot, assign its owners and agree its evidence gates. Wider procurement follows what the pilot demonstrates.

References

Empirical and theoretical claims in this blueprint are referenced to the following published literature. In-text citations use author and year.

Anthropic. (2024). Introducing contextual retrieval. Anthropic Engineering. https://www.anthropic.com/engineering/contextual-retrieval

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. Proceedings of the National Academy of Sciences, 122(26), e2422633122. https://doi.org/10.1073/pnas.2422633122

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610–623). ACM. https://doi.org/10.1145/3442188.3445922

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185–205). MIT Press.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. American Psychologist, 34(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies, 15(1), 6. https://doi.org/10.3390/soc15010006

Girls' Day School Trust. (n.d.). The Girls' Futures Report. GDST. https://www.gdst.net

Groeneveld, D., Beltagy, I., Walsh, P., Bhagia, A., Kinney, R., Tafjord, O., et al. (2024). OLMo: Accelerating the science of language models. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (pp. 15789–15809). ACL. https://doi.org/10.18653/v1/2024.acl-long.841

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542–570. https://doi.org/10.1111/ejed.12533

Information Commissioner's Office. (2020). Age appropriate design: A code of practice for online services. ICO. https://ico.org.uk

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. Science, 331(6018), 772–775. https://doi.org/10.1126/science.1199327

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educational Psychologist, 41(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv preprint arXiv:2506.08872. https://arxiv.org/abs/2506.08872

Livingstone, S., Stoilova, M., & Nandagiri, R. (2019). Children's data and privacy online: Growing up in a digital age. London School of Economics and Political Science.

Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. https://doi.org/10.54675/EWZM9535

References

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2), 230–253. https://doi.org/10.1518/001872097778543886

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation). Official Journal of the European Union, L 119, 1–88.

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 (Artificial Intelligence Act). Official Journal of the European Union, L, 2024/1689.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. Polity Press.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. Journal of Experimental Psychology: Human Learning and Memory, 4(6), 592–604. https://doi.org/10.1037/0278-7393.4.6.592

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 3645–3650). ACL. https://doi.org/10.18653/v1/P19-1355

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. Educational Psychology Review, 10(3), 251–296. https://doi.org/10.1023/A:1022193728205

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Williamson, B. (2017). Big data in education: The digital future of learning, policy and practice. Sage.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. Theory Into Practice, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs.

Documentary references (Empathy.AI)

Local AI Workstation, v3 (technical dossier: hardware, installed models, system architecture, economics, capacity and governance) · Edge AI for Education (workshop programme on sovereign AI for education: values, minors' privacy, transparency and ethical boundaries) · AI in Education: An Empathetic Vision for Creativity, Art, and Liberation · AI Governance in the Perspective of Empathy AI (July 2026) · About Empathy.AI Spaces (product overview and deployment options) · A Strategy for AI in Research (November 2026, discussion draft).

A blueprint shaped by each institution

This document is a starting point. Every institution's data, platforms and pedagogical priorities differ, and the right architecture, capacity and pace of adoption should be shaped around them. The posture stays constant: empowerment through ownership, transparency in every answer, and governance an institution writes for itself.



Empathy.AI headquarters in Gijón, Asturias: a bioclimatic building powered by solar energy.

Empathy.AI and its headquarters

Empathy.AI is a foundational AI platform within Empathy Holdings, developing ethical, sustainable and privacy-first AI for commerce, legal, government, healthcare, finance and publishing. Its main infrastructure runs in a bioclimatic building in Gijón, Asturias, powered by solar panels, in response to the energy concerns the research community has raised about AI (Strubell, Ganesh & McCallum, 2019).

OPERATIONS

Solar-powered infrastructure

Compute hosted in a building designed for low energy use and powered by on-site solar generation.

EDGE AI FOR EDUCATION

Workshops in Gijón

Educators visit the physical infrastructure to see, touch and understand how the AI works.

UN SDGS

Aligned goals

Climate Action (SDG 13), Responsible Consumption (SDG 12) and Sustainable Cities (SDG 11).

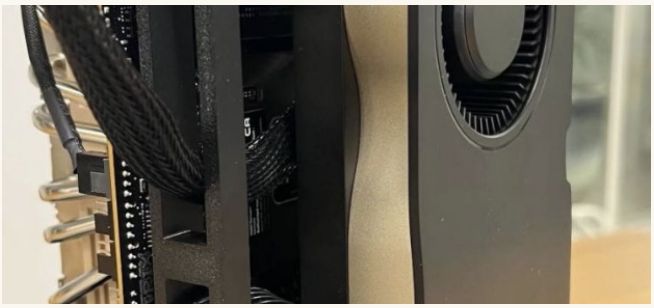
Empathy.AI designs its platform to do more with less: the smallest adequate model, shared capacity and scheduled heavy work. Sustainability is one of its core values.

Starting the conversation. This blueprint is designed to be worked through together. Empathy.AI can support leadership and technology teams in shaping a strategy grounded in the frameworks set out here.

Annex A · The Local AI Workstation

A technical reference for the local AI infrastructure behind the approach in this blueprint.

> LOCAL AI WORKSTATION, EMPATHY.AI
> "It is a visual allegory, a tool for hope: a dedicated computer optimised for ownership, offering a private, self-contained and sovereign way to relate with knowledge, free from Big AI."



Vertical GPU riser mount, internal build



The assembled workstation

01 · COMPUTE

A dedicated GPU on site

32GB professional Blackwell card, dedicated to the institution.

02 · MODELS

Authored knowledge

Relate with knowledge on the institution's terms.

03 · DATA

Privacy first

Documents and prompts are processed locally and stay in the building.

04 · SPACES

Full agentic capability

Adaptive mode and agentic tool use, running end to end on site.

Component	Model	Notes
Motherboard	ASUS ROG STRIX X870-I Gaming WiFi	Mini-ITX
CPU	AMD Ryzen 7 9700X	
CPU cooler	Noctua NH-L12Sx77	Low profile
Memory	Kingston FURY Beast DDR5 32GB (2×16GB)	
GPU	NVIDIA RTX PRO 4500 Blackwell (PNY)	32GB VRAM, 896 GB/s
Storage	Kingston NV3 1TB NVMe SSD	
PCIe riser	JMT PCIe 5.0 x16 180°	
Power supply	Cooler Master GX II Gold 850W	Modular
Chassis fans	2× Noctua NF-A14x25 G2 PWM Sx2-PP 140mm	
Aux fans	Noctua NF-A4x20 PWM (2-pack)	
Enclosure	Custom 3D-printed acoustic housing and base	

Annex A · Installed models

Model	Configuration	Role
Qwen3-Embedding-4B	AWQ INT4	Text embedding and retrieval
Qwen3.8-27B	27B · Qwen · GGUF Q4_K_M	The default for strong general reasoning and agentic tool use, tuned for everyday workflows
OLMo 3.1 32B Think	32B · Ai2 · GGUF Q4_K_S	Fully open training data and weights, for teams who want transparent provenance and deliberate local reasoning (Groeneveld et al., 2024)

Both reasoning engines run locally and can be selected per Space or per task. The workstation runs the complete version of Spaces, with Adaptive mode and full agentic capability: inference and tool use end to end on site.

Annex B · Empathy.AI Spaces

Empathy.AI Spaces is Empathy.AI's flagship product: a private, self-hosted retrieval-augmented generation (RAG) platform for discovering, managing and working with an organisation's internal knowledge.

What Spaces does

Spaces brings fragmented organisational knowledge (Confluence pages, GitHub repositories, PDF, DOCX and Markdown files, internal APIs and databases) into one conversational system. Staff ask questions in natural language and receive accurate, cited answers drawn from their own data, inside their secure environment. Spaces enriches every document chunk with surrounding context before indexing; combined with reranking, this contextual retrieval reduced retrieval failures by up to 67% against standard chunking (Anthropic, 2024).

Q: Which evidence shows competency stagnation in the last term?
A: Three students show a sustained drop in logical reasoning self-assessments between September and November, coinciding with reduced participation in optional challenges.
Sources: Self_Assessments_Y9.csv · Teacher_Observations_Q1.pdf

Illustrative rendering of the cited response pattern in an educational domain. A teacher interprets the pattern and decides on any action.

Deployment options

Option	Description	Suited to
On the institution's infrastructure	Spaces runs on the institution's own server or workstation, with full implementation by Empathy.AI. This is the scenario described in this blueprint.	Full data sovereignty on site; lowest long-term cost
Leased hardware	High-performance GPU hardware leased from Empathy.AI and hosted in its facility in Asturias.	Dedicated capacity without owning or managing hardware
Managed Private Space	A fully managed environment on Empathy.AI's private, solar-powered infrastructure in the EU.	Fastest deployment with minimal infrastructure work

All options run open-weight models and are designed to align with the GDPR, the EU AI Act and data residency law. The data boundary differs: the institution's premises, or Empathy.AI's dedicated EU infrastructure.

Annex C · Case study: Putney High School

AI on campus, under school governance. Empathy.AI Spaces on a dedicated Local AI Workstation, giving the school a privately hosted, sovereign AI environment in its educational experience, with guided introduction at Year 9 and student-built, programmatic use at Year 13.

At a glance

Organisation	Putney High School, a selective girls' school in the GDST family (24 schools, ~100,000 alumnae)
Need	A school-authored way to meet AI in the classroom. Computing is timetabled from Year 7, Python from Year 9 and A-Level Computer Science is offered, while “AI appears neither as a standalone offering nor as named policy” (Made in Putney)
Deployment	Empathy.AI Spaces, full version (Adaptive mode plus agentic tool use), running end to end on a Local AI Workstation inside the school
Model of access	Year 9: guided first contact. Year 13: programmatic access, with students building new experiences on the same infrastructure
Data posture	Documents, prompts, retrieved evidence, embeddings, model outputs and the reasoning loop remain in the building, with all inference on the school's own workstation

Context: a school ready for this question

Putney High's stated purpose is educating girls “to be well rounded, well qualified and well prepared,” with “fearless curiosity, intellectual agility, and a joyful ambition for originality and adventure” (Made in Putney). Its technology path is already drawn: computing from Year 7; Python as a third-language option from Year 9 alongside Italian and Spanish; Design Thinking for all Year 9 students; Computer Science to GCSE and A Level.

The GDST Girls' Futures Report sharpens the mandate: 54% of surveyed girls follow current affairs through social media, 28% trust what they see there, and 56% encounter fake news weekly. Judgement is the priority skill (GDST, Girls' Futures Report). The school's materials record the gap openly, with no specific AI programme yet detailed. The Empathy.AI installation fills that place with an AI the school governs on its own premises.

The school writes the rules: who may use the system, at what age, with which sources, and what is retained.

Annex C · An owned execution layer

Hardware basis

Component	Specification
GPU	NVIDIA RTX PRO 4500 Blackwell (PNY): 32GB VRAM, 896 GB/s memory bandwidth, vertical riser mount
CPU / RAM	AMD Ryzen 7 9700X · Kingston FURY Beast DDR5 32GB (2×16GB)
Motherboard / storage	ASUS ROG STRIX X870-I Gaming WiFi (Mini-ITX) · Kingston NV3 1TB NVMe
Power / cooling	Cooler Master GX II Gold 850W · Noctua low-profile and PWM fan set, custom 3D-printed acoustic housing

Designated capacity: 300W whole-system active draw (~€302–343/month local operating cost at €0.30/kWh). These are planning bands to confirm by benchmark.

Age-staged access

Stage	Experience	Governance
Year 9	Guided first contact inside a closed perimeter. Students ask questions of curated, school-approved material and see the source behind every answer.	Teacher-configured Space; cited answers; school identity and retention policy
Year 13	Programmatic access. Students build new experiences on the same infrastructure, with their code running in an isolated sandbox.	Supervised development; sandboxed execution; school-written permissions

What the school controls

- The collections each Space may consult, and who curates them.
- Identity, access by year group and retention of conversations.
- The model version, which changes only when the school decides.
- The record of how the system is configured and used.

The case shows a staged path from guided first contact to student-led development, on infrastructure the school owns and governs.

The case study describes the deployment as presented in Empathy.AI materials and the school's published curriculum. Learning outcomes from the deployment are a subject for evaluation.

Annex D · Evidence boundaries and source notes

This blueprint distinguishes published research, regulation, Empathy.AI product descriptions and proposed design.

Published research

The MIT Media Lab study by Kosmyna and colleagues (2025) is an arXiv preprint with 54 participants. It provides the concept of cognitive debt, the ownership and recall figures on page 13 and the design implication on sequencing. Its findings indicate direction and warrant replication. The Bastani and colleagues field experiment (2025) is peer reviewed and published in PNAS. The learning science on scaffolding, cognitive load, generation, retrieval practice and metacognition is long established; its application to AI design here is this blueprint's proposal.

Regulation and guidance

References to the GDPR, the UK Age Appropriate Design Code, the EU AI Act and UNESCO guidance describe the frameworks that apply to minors' data and AI in education. Each institution should confirm its obligations with its data protection officer and legal advisers.

Empathy.AI materials

Descriptions of Spaces, the Local AI Workstation, deployment options and case studies draw on Empathy.AI's own documents listed on page 31. Capacity, concurrency and cost figures are planning estimates from the Local AI Workstation dossier. Cloud equivalent spend reflects private capacity services and varies with provider, region and contract.

Leo's story

The opening story on pages 3–10 uses illustrative Empathy.AI Spaces screens. Leo, Maya Alvarez, Elmbrook Primary, the North Vale Schools Network, Dr Hana Ruiz, Reading Late and the trial results belong to that illustrative narrative. They establish proposed acceptance requirements, tested on page 28, and describe no measured outcome. The network, grants, receipts, correction workflow and revocation require implementation verification.

Proposed design

Educational configurations of the four capability lines, the evidence gates, the pilot measures and the illustrative answer on page 34 are proposals. They define acceptance requirements to test in each institution.

Boundaries to retain

This blueprint makes claims about design and governance. Guarantees of learning outcomes, regulatory compliance, the absence of incorrect answers or zero-carbon inference sit outside its scope. Source attribution, data protection and pedagogical value each require their own checks. Before procurement, verify current controls, hosting and support, licences, retention, exit arrangements and a data protection impact assessment for the intended use.

Trust is earned in the classroom: through answers students can check, data the institution keeps and decisions teachers own.