

Data Flow and Data Handling Summary

v3.12 · Updated Aug 2026 · Reviewed Jun 2026 · Next review Jun 2027

Owner: Andrew James (CEO)

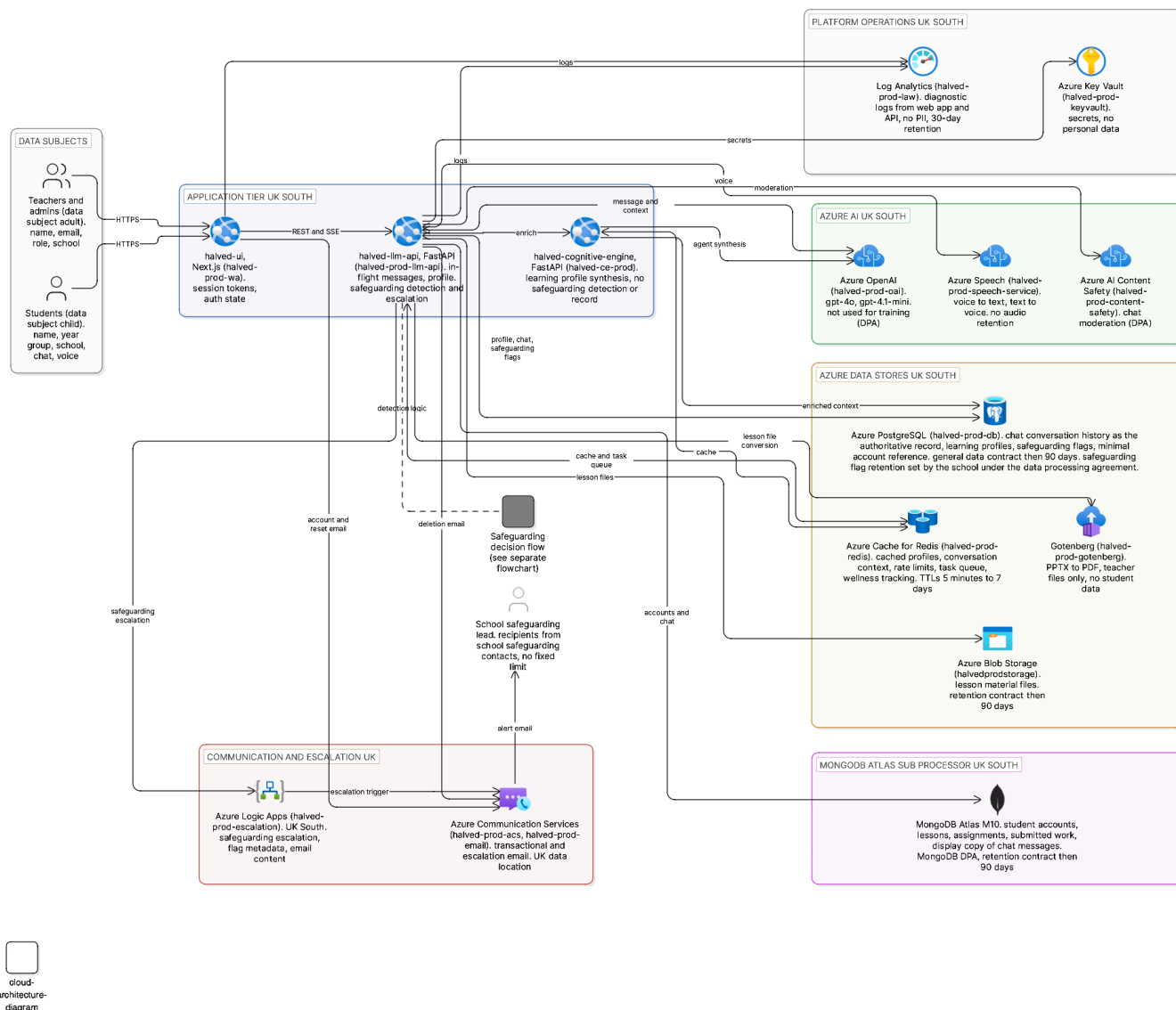
For school IT leads and Data Protection / Designated Safeguarding Leads

Overview

Halved is an always available learning support system that brings expert guidance to every student at the exact moment they need it. This document describes what data Halved collects, where it is stored, how it is processed, and how it is protected. It is intended to support your school's data protection assessment and due diligence process.

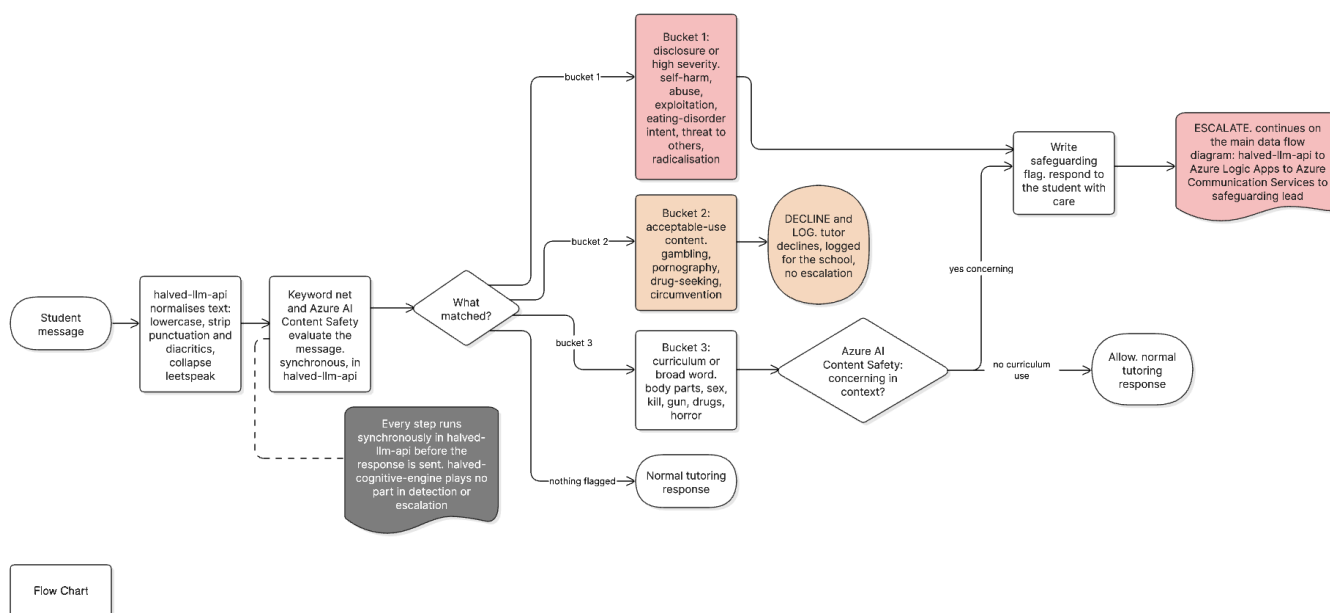
All data storage and all data processing, including artificial intelligence inference and voice processing, takes place within the United Kingdom. No student personal data is stored or processed outside the United Kingdom.

Data Flow Diagram



The Halved platform architecture and data flow, showing how data moves between the student's browser, the Halved application layer, and Microsoft Azure UK South services.

Safeguarding Decision Flowchart



How student messages are checked in real time: detection via Azure AI Content Safety, severity assessment, and escalation to the school's Designated Safeguarding Lead.

1. What data Halved holds about students

When a student account is created, Halved stores:

- A unique student identifier as the primary student record, and a school email address retained for authentication and account communications. Full student names are held by the school, not by Halved. By default a student's first name is held for use in chat; schools may choose a no-names configuration in which even the first name is school-held.

The student identifier is a reference, not a credential. Knowing an identifier grants no access. All access is authenticated and scoped server-side to the authorised user and their school.

Roster ingestion. For the September 2026 pilot, student and staff records are provisioned by Halved-managed spreadsheet ingest only; the school self-entry mode described below is not enabled for the pilot. Schools provide their roster as a spreadsheet, and what Halved receives and retains differs by whether names are included. Spreadsheet without names: Halved receives identifiers and minimal account fields only; this is the default and the cleanest. Spreadsheet with names: any full names received at ingestion are stripped on receipt and not persisted on the Halved side, leaving only the identifier, email and, by

default, first name. School self-entry, in which school staff enter accounts directly and Halved holds only the model fields, is a planned future option gated on the school super-admin role and is not in use for the pilot. Account creation is always performed by schools or their staff; students do not self-register.

- Year group and subject assignments, set by the teacher.
- Information the school supplies about where a student is starting from: a reading age, and whether the student is working at age expectation in English and in maths. Each is supplied as a fixed choice. These are attainment and reading measures the school already holds, recorded as the school states them and neither assessed nor inferred by Halved. They are not SEND status, not a diagnosis, not a condition and not a disability label, and there is no free-text field anywhere in the set.
- Assignment work, the written work the student produces within Halved.
- Student-uploaded content, files and documents uploaded by the student. Screened through the same Azure AI Content Safety pipeline as chat messages, with keyword detection running before content-safety scanning. A malware-at-rest gate (Defender for Storage) applies to all uploaded files; the Standard tier, which performs active scanning and tagging, will be active on the production environment from 1 September 2026.
- Student-authored free-text notes, work-completion notes and jotter text authored by the student. Stored and visible to authorised staff; not routed through the live safeguarding pipeline and not live-escalating. This is the single unscreened text surface on the platform, mitigated by staff visibility and point-of-use framing. Access defaults: submitted work is visible to the task teacher and class staff; free-text notes are visible to the safeguarding lead and form tutor only; role terms are school-configurable.
- Chat conversation history, the text of conversations between the student and Halved's learning support.
- A learning profile, a background summary built by Halved from how the student interacts during tutoring, together with the fixed-choice answers the student gives about how they prefer to learn and the information the school supplies about where the student is starting from. It records learning behaviour, stated preferences and those school-supplied measures, such as topics covered, apparent strengths, areas of difficulty, preferred ways of learning, and recent engagement. It does not record any medical diagnosis, neurodevelopmental condition, or disability label. This summary is used to personalise future sessions and is held within Microsoft Azure UK South.

- A communication register, a coarse value on a scale of 1 to 7 that sets the language Halved uses with the student, meaning vocabulary level and sentence complexity. It starts from the student's year group and from the reading age the school supplies, and adjusts deterministically based on how the student responds during use. It is a setting for how Halved talks to the student rather than a judgement of the student: it is not itself a record of ability, attainment or reading age, and it records no diagnosis, disability or special-category data. It is held per student in the learning profile and is not shown to the student or to other students.
- A task-confidence self-assessment, where the student has chosen to give one: a number on a simple four-step scale recording how confident the student felt about a task. It is student-provided, separate from the learning behaviour Halved observes, carries no free text and no mood or wellbeing content, and is used to personalise support and to show teachers a trend over time.
- Answers to fixed-choice questions about how the student prefers to learn and how they approach their work. These are student-provided, chosen from options Halved sets, and carry no free text. They record a preference or a working approach only, and do not ask about health, disability, SEND status, a diagnosis or any other special category of data. Answering is optional. They are used to personalise the support the student receives and are not shown to teachers.
- Activity records, a record of when the student was working in Halved and on which surface. Each record holds a start time and an end time, each held to the minute, the exact length of the period in seconds, the surface the student was on, and, where one is known, the session the activity belonged to. Because these are clock times, they cover whenever a student chooses to work, including evenings and weekends. This is set out in full under "Activity and time spent" below, including how the clock times and the length are held.

Halved does not collect:

- Biometric data.
- Location data.
- Device identifiers beyond a standard browser session.
- Payment information.

2. Where data is stored

All data is stored within the Microsoft Azure UK South region (London datacentres), or on MongoDB Atlas hosted in Azure UK South. No student personal data is stored outside the United Kingdom.

What	Storage technology	Location
Student accounts, lessons, assignments, submitted work, and the display copy of conversation messages shown in the interface	MongoDB Atlas	Azure UK South
Full chat conversation history as the authoritative turn-by-turn record, learning profiles derived from conversations, attainment records, activity records, and safeguarding flags	Azure Database for PostgreSQL	Azure UK South
Uploaded lesson files (PPTX, PDF)	Azure Blob Storage	Azure UK South
Student-uploaded files	Azure Blob Storage	Azure UK South
Student-authored free-text notes	MongoDB Atlas / Azure PostgreSQL	Azure UK South
Temporary cache, conversation context, rate limiting, and background task queue	Azure Cache for Redis	Azure UK South
Secrets and credentials	Azure Key Vault	Azure UK South
Application and audit logs	Azure Log Analytics workspaces	Azure UK South

All data is encrypted at rest. Secrets such as API keys, passwords, and encryption keys are stored in Azure Key Vault and never stored in plain text.

3. AI and voice processing

All AI inference and voice processing is performed within Microsoft Azure UK South. No student data is transmitted to servers outside the United Kingdom for AI or voice processing.

AI chat processing, Azure OpenAI (UK South)

When a student sends a message to Halved's learning support, that message, along with structured context about the lesson topic, assignment criteria, and the student's learning profile, is transmitted to the Microsoft Azure OpenAI Service, hosted in Azure UK South (London datacentres).

What is sent to Azure OpenAI:

- The student's message text, exactly as typed.
- Lesson notes and assignment criteria for the current task.
- A summary of the student's learning profile, expressed as observed learning behaviour (for example, that the student benefits from shorter explanations).
- A role-based identifier, a unique ID, not the student's name.

What is not sent to Azure OpenAI:

- The student's name or email address as structured fields.
- Diagnoses or condition labels. Halved is configured so that diagnosis and condition labels are never written into the learning profile or transmitted to the AI service.
- Submitted work from previous assignments.
- Any data not directly relevant to the current tutoring session.

Azure OpenAI Service is governed by the Microsoft Data Processing Addendum. Data sent via the Azure OpenAI API is not used by Microsoft to train foundation models, and Halved does not use student data to train third-party models.

Halved operates as a closed-loop tutoring system: it does not browse the internet and does not return public-search results.

Note: if a student types their name or personal details directly into the chat box, that text will be included in what is sent to Azure OpenAI, because Halved processes the message as written. Students should be advised not to include personal information in chat messages.

Voice features, Azure Speech Service (UK South)

Halved supports optional voice features, both processed entirely within Azure UK South:

- Speech to text, a student speaking their answer, processed via the Microsoft Azure Speech Service, UK South region.
- Text to speech, Halved reading content aloud, using the Microsoft Azure Speech Service (voice: en-GB-AdaMultilingualNeural), UK South region.

Voice features are optional. If your school prefers to disable them, this can be requested.

Handwriting, annotation and images

Handwriting-to-text (input method). Student writes on a canvas in the browser -> the image is sent directly to Azure AI Vision (Read OCR, UK South) -> the recognised text is returned -> the image is discarded (held in memory for the duration of the call only; no write to Blob, database or logs) -> the recognised text passes Halved's content-safety and safeguarding screening before it is stored or reaches the tutor. The recognised text, not the ink, is what Halved reads. If handwriting cannot be read, Halved tells the student it cannot read it.

Submitted annotations (student work). Student's strokes are drawn over the lesson slide, flattened into a single image, and uploaded to Halved's Blob storage (container student-uploads, path scoped by school and user). A record is created in the platform database and referenced from the submitted work. The image is screened through the same pipeline as any other student upload. Retained with the assignment record; no automatic expiry; removable by staff or by the student's school through the deletion process.

Teacher annotations. While marking, a teacher's marks are held as coordinates in the platform database. On confirmation, the marked-up slide is flattened into an image and stored in Blob storage in the same way. Retained with the assignment record as staff-authored feedback.

Photographed work. The photograph is uploaded to Halved's Blob storage first. Content moderation and then text recognition (Azure AI Vision Read, UK South) run against the stored file. The recognised text is screened. The photograph is retained as the image of that page of work; the screened text is the text of that page.

What the sub-processor holds. Azure AI Vision receives an image, returns text, and retains nothing. Halved's own retention of images (submitted annotations, teacher annotations, photographs) is described above and is governed by the Data Retention Policy.

Teacher-to-student messaging

A teacher composes a message and selects an audience (an individual student, a class, or a group). The message is stored in the platform database with the sender, the recipients and a timestamp, and is shown to the recipient student in the platform. Teachers can edit or recall a message they have sent.

Messages are staff-authored content. They are not passed to the AI tutor as context, and they do not contribute to the AI learning profile.

Messages are automatically deleted 7 days after they are sent.

Attainment and progress

Three sources feed the attainment record:

Teacher marks. A teacher records a score against a piece of work (a score and a total). This is entered by the teacher, not generated by Halved.

Code-judged answer correctness. Where a student answers a question that can be checked deterministically (the answer is either right or it is not), Halved records the outcome of that check. This is a record of a code-evaluated result. **An AI model's opinion never enters the attainment record.** Where correctness cannot be determined by a straightforward check, no attainment record is created.

Time in productive difficulty. Halved records how long a student spends working on a task that is appropriately stretching for them (as opposed to breezing through it, or being stuck). This is a task-fit signal, used to tell a teacher whether the work is pitched right.

Where a student reworks a piece of work after support from Halved, the change between the first attempt and the reworked attempt is recorded. This change is the only element Halved attributes to its own support.

The attainment record is surfaced to the student's teachers and to the student as a trend. It is not used for ranking, comparison between students, or any automated decision about a student. It does not feed any inference about ability, SEND, or developmental stage.

Activity and time spent

What the record is. Halved holds a table of student activity intervals. One row records that a named student was active from a start time to an end time on one surface. Each row holds the student identifier, the school identifier, an optional session identifier, the surface, the start time and the end time, the length of the interval in seconds, and the time the row was

created. The start time, the end time and the creation time are full timestamps with a time zone, each truncated to the minute. The length is computed before that truncation and stored exactly, in seconds, so the duration of an interval survives at full precision while the clock times do not. The truncation is enforced by the database as a constraint on each of the three timestamps, rather than by the code that writes them, so a later endpoint or a direct insert cannot store a finer time.

The surface value is one of six: chat, work-completion, lesson-content, lesson-notes, homepage, or unknown. The two lesson surfaces are deliberately separate: lesson-content is the lesson page a student reads or works through, and lesson-notes is the list page they browse to reach it, so that time spent choosing what to open is recorded as navigation rather than counted as time spent studying. “Unknown” is a real stored value, used where the surface could not be determined. The session identifier can also be absent. A row with no session and a surface of unknown attributes a period of time to a named student with no further context about what they were doing.

What it does not hold. No message content, no work content, no location, no device or browser identifier, and no inference about the student. It is a record of clock time and surface only.

Out-of-hours activity. Because these are clock times, the record covers whenever a student chooses to work, including evenings, weekends and school holidays. Halved does not restrict when a student may use the platform, so the record follows the student’s own choice of working time.

What it is for. Three outputs are produced: the total time a student has spent working, the proportion of that time spent with Halved’s support, and a distribution showing when in the day a student works. These are for the student’s teachers, and for the student’s own view of their own record.

What is built. Recording, the teacher measures and the student’s own view of the record are built and were released as a single change. Recording of a student’s activity begins in a given environment of the platform when that release is deployed there, and where it has not been deployed no activity record exists.

How the measurement is bounded. Only student accounts are recorded; the timer does not run for staff. An interval ends at the last interaction actually observed, not at the later point the platform notices the student has stopped, and the same rule applies when the browser tab is hidden, since a student may have walked away before that happened. The measure is

therefore conservative by construction: it can under-report a student who sits still, and cannot over-report anyone. For a figure shown to a teacher about a child, and to that child about themselves, that is the direction in which to err.

Who sees it. The student dashboard shows a student their own activity record, including a version of the when-in-the-day view, so that a student can see what is held about their own working time. Halved committed to the order: the timer that creates activity records ships with the student view, not before it. That commitment was met, the two having been released together, so there is no build in which a student's working time is recorded while the student has no way to see it.

Retention. A maximum of 12 months from the start time of the interval, purged by a scheduled job, and in any event deleted on contract end plus 90 days, as set out in the Data Retention Policy.

4. How access is controlled

Students and teachers log in with an email address and password.

- Passwords are stored using bcrypt hashing and are never stored or transmitted in plain text.
- Sessions use encrypted JSON Web Tokens with a seven-day expiry.
- Communication between Halved's web app and AI backend uses an internal API key, not exposed to the browser.
- Teacher accounts and student accounts are role-separated. Teachers can view only their own students' work.
- System administrator access is separately controlled and limited to Halved staff.

5. Safeguarding

Halved includes a live safeguarding pipeline operating across all environments. Student messages are checked in real time. Where a welfare concern is detected, the concern is logged and an alert is sent to the school's nominated safeguarding lead or leads. Lower-severity concerns are flagged and escalated without interrupting the student's session. Higher-severity concerns return appropriate support information to the student and are escalated immediately.

The safeguarding pipeline uses Azure AI Content Safety and Azure Logic Apps, both hosted in Azure UK South and covered by the Microsoft Data Processing Addendum. Safeguarding alert content, including a short excerpt of the flagged message, is sent to the school’s nominated safeguarding contacts by email through Azure Communication Services (United Kingdom).

Student-uploaded content is screened through the same safeguarding pipeline as chat messages: keyword detection runs first, followed by Azure AI Content Safety scanning (UK South). A malware-at-rest gate (Defender for Storage) applies to all uploaded files; the Standard tier will be active on the production environment from 1 September 2026.

Student-authored free-text notes (jotter and work-completion text) are not routed through the live safeguarding pipeline and do not trigger live escalation. This is the single unscreened text surface on the platform; the safeguarding mitigation is staff visibility and point-of-use framing.

6. Data security summary

Control	Status
Encryption at rest	Yes, all Azure storage
Encryption in transit	Yes, HTTPS/TLS throughout
Secrets management	Azure Key Vault
Password storage	bcrypt hashed
Role-based access control	Yes, student, teacher, admin
UK data residency (all storage and processing)	Yes, Azure UK South and MongoDB Atlas Azure UK South

7. Data retention and deletion

Halved retains student personal data for the duration of the school’s contract. Following termination of services, general personal data is securely deleted within 90 days, unless Halved is required to retain it for longer to comply with legal, accounting, or regulatory requirements.

Safeguarding records are treated differently. Where a safeguarding concern has been recorded, the associated record is retained in line with statutory safeguarding guidance (Keeping Children Safe in Education) and the school's own retention schedule, which is typically up to seven years. This applies regardless of the general deletion timeline above.

Requests for deletion of an individual student's data during the contract period (the right to erasure under UK GDPR) should be directed to dataprivacy@halved.io. Halved will action these requests without undue delay and in any event within one month of receipt, extendable by up to two further months for complex or numerous requests. Where a record is subject to a statutory safeguarding obligation, Halved may retain it to the extent required by law. The school, as controller, determines the retention and erasure of safeguarding records, as set out in the Data Processing Agreement and aligned to the school's safeguarding retention schedule and Keeping Children Safe in Education.

In some circumstances, Halved may anonymise personal data so that it can no longer be associated with any individual. Anonymised data may be retained indefinitely and used to improve the platform.

8. Third-party services and sub-processors

All third-party services used by Halved process data within the United Kingdom and are covered by a Data Processing Agreement. A full sub-processor register is available on request.

Service	Purpose	Location	Student data involved	DPA
Microsoft Azure (App Service, PostgreSQL, Redis, Blob, Key Vault, Log Analytics)	Application hosting, primary data storage, secrets management	UK South	All student and teacher data	Microsoft DPA
Azure OpenAI Service	AI learning support responses	UK South	Chat messages, lesson context, learning profile summary	Microsoft DPA
Azure Speech Service	Speech to text and text to speech	UK South	Student voice audio, AI-generated text	Microsoft DPA

Service	Purpose	Location	Student data involved	DPA
Azure AI Content Safety	Safeguarding and content moderation of messages	UK South	Chat message content	Microsoft DPA
Azure Logic Apps	Safeguarding escalation workflow to the school's Designated Safeguarding Lead	UK South	Safeguarding alert content and flagged message excerpts	Microsoft DPA
Azure Communication Services	Transactional emails (account creation, password reset) and safeguarding escalation emails to school leads	United Kingdom	User email addresses, names, account setup links, safeguarding alert content	Microsoft DPA
Azure Container Instances (Gotenberg)	Lesson material document conversion, PPTX and PDF converted to page images for display	UK South	Teacher-uploaded lesson material content	Microsoft DPA
MongoDB Atlas	Student accounts, lessons, assignments, submitted work, and the display copy of conversation messages	Azure UK South	All structured student and teacher data	MongoDB DPA
Cloud202 (technical contractor)	Production infrastructure management, deployment, and Terraform state	London UK	Administrative access to production systems holding student and teacher data	Data Processing Agreement

9. Contact

For data protection questions related to your pilot, or to request a Data Processing Agreement, please contact:

Halved Limited

dataprivacy@halved.io

Registered in England and Wales, company number 15261677.

10. Review

This summary is reviewed at least once a year, and sooner when the platform architecture, data processing activities or sub-processors change. The next scheduled review is June 2027.

This document reflects the technical architecture of the Halved platform as at the review date shown at the top of this page.

Questions about data protection? dataprivacy@halved.io

Website halved.io · students log in at my.halved.io

Halved Limited · registered in England and Wales, company number 15261677