

Kita DPIA – draft

This template is an example of how you can record your DPIA process and outcome. It follows the process set out in our DPIA guidance, and should be read alongside that guidance and the Criteria for an acceptable DPIA set out in European guidelines on DPIAs.

You should start to fill out the template at the start of any major project involving the use of personal data, or if you are making a significant change to an existing process. The final outcomes should be integrated back into your project plan.



Submitting controller details

Name of controller	Kita Education Ltd (company number 13264075), Studio 43 Great Western Studios, London W2 5EU
Subject/title of DPO	Chief Technology Officer
Name of controller contact / DPO (delete as appropriate)	Nathan Varughese, CTO, Kita Education Ltd

Step 1: Identify the need for a DPIA

Explain broadly what project aims to achieve and what type of processing it involves. You may find it helpful to refer or link to other documents, such as a project proposal. Summarise why you identified the need for a DPIA.

Kita is an AI-powered online coding education platform designed for secondary school students. This DPIA covers the introduction of an AI tutoring layer into the Kita platform and its deployment in a funded pilot with up to 10 UK secondary schools (up to approximately 300 students and up to approximately 12 teachers), under a grant from the Department for Education (DfE) for the purposes of Computer Science education. The pilot aims to evaluate whether AI-generated formative feedback can improve student learning outcomes and reduce teacher marking workload, and to assess the impact of AI-assisted tools on feedback quality and teacher confidence.

The AI tutoring layer uses a multi-agent architecture with a persistent per-student memory layer. During each interaction, students' code, error messages and conversation history are transmitted to an external AI provider (Azure OpenAI running at a UK data centre) for processing.

A DPIA is required under Article 35 UK GDPR and ICO guidance due to the large-scale systematic processing of students' data using novel AI and automated technology in an educational setting. Kita adheres to the ICO Children's Code (Age Appropriate Design Code) in the design and operation of its platform. An independent external legal review of this DPIA has been conducted by Tariq Sayfoo, Cripps LLP.

Step 2: Describe the processing

Describe the nature of the processing: how will you collect, use, store and delete data? What is the source of the data? Will you be sharing data with anyone? You might find it useful to refer to a flow diagram or other way of describing data flows. What types of processing identified as likely high risk are involved?

Kita processes data through two main streams: (1) technical platform processing to deliver AI-powered tutoring; and (2) research processing to evaluate the effectiveness of AI-assisted coding education as part of the DfE pilot.

Collection: Student data is collected at registration (first name, last name, auto-generated username, school association) and continuously during use (code in the editor, runtime errors, AI conversation history, I/O test inputs/outputs, quiz answers). Teachers provide name and email at registration. Student accounts are provisioned through the school, so teachers or administrators provide students' identity data on the school's behalf, while code, errors, chat and quiz answers are generated by the student during use.

Use: During each AI tutoring interaction, the student's current code, error messages, session conversation history, and lesson context are transmitted in real time to Azure OpenAI (UK South region) to generate the tutoring response. A persistent memory system maintains per-student context across sessions: AI help request history, recurring errors, completed tasks, and lesson goals.

Storage: Data is stored in a cloud-hosted PostgreSQL database (Neon). File and asset storage uses AWS S3. Authentication and user profiles are managed via Supabase. AI agent memory is stored via Letta. All four stores are hosted in UK/EU regions, supporting UK data residency.

Retention and deletion: Session working memory is cleared at session end. Core and long-term AI memory is deleted at course completion or within 12 months of last access, whichever is sooner. Code submissions, test results and quiz answers are retained until 31 August 2026 (end of the research evaluation window), after which they will be anonymised. Anonymised research data is retained for 2 years from the end of the pilot (until July 2028), then deleted. Consent forms are retained for the duration of the pilot plus 2 years, held by the University of Brighton in accordance with their research governance procedures.

Sharing: Pseudonymised student data (student work, assessments, submission times and time on task, with names and account IDs removed) is shared with the University of Brighton for independent evaluation under a data sharing agreement to be finalised prior to any data being shared. No data will be shared with the University until this agreement is in place. The DfE will receive an anonymised summary report of findings only. No other third-party sharing is intended.

University of Brighton data role: The University of Brighton acts as an independent data controller for the research data it receives and processes, in accordance with its own Research Privacy Notice and ethics approval from the University of Brighton Cross-School Research Ethics Committee (confirmed 1 June 2026). A data sharing agreement will be finalised prior to any data being shared. If on further legal review the University is determined to be acting as a joint controller, a joint controller agreement under Article 26 UK GDPR will also be required.

High-risk processing types present: Kita falls under the ICO's 'Innovative Technology' category given its use of novel AI agents with persistent memory in an educational setting. The processing also falls under the ICO's 'vulnerable individuals' category because the data subjects are children (aged 13–18). Incidental special-category data could be disclosed through open-ended AI chat; this is addressed at R5 and in Step 6.

Describe the scope of the processing: what is the nature of the data, and does it include special category or criminal offence data? How much data will you be collecting and using? How often? How long will you keep it? How many individuals are affected? What geographical area does it cover?

Nature of data: Identity data (first name, last name, auto-generated username, Kita-assigned school ID); behavioural/educational data (code written in the editor, runtime error messages, Python error types and line numbers, AI chat history, I/O test inputs/expected/actual outputs, quiz answers, course progress). No special category data or criminal offence data is intentionally collected. Students are asked not to share personal information in AI chat.

Volume and frequency: Up to 300 students and 12 teachers across up to 10 UK secondary schools. All students are aged 13–18; no students under 13 are participating in the pilot. Students are expected to use the platform regularly during lessons, with multiple AI interactions per session.

Retention: As set out above under 'Retention and deletion'.

Geography: Students and schools are located in the United Kingdom. AI processing occurs in the UK via Azure OpenAI UK South region.

Describe the context of the processing: what is the nature of your relationship with the individuals? How much control will they have? Would they expect you to use their data in this way? Do they include children or other vulnerable groups? Are there prior concerns over this type of processing or security flaws? Is it novel in any way? What is the current state of technology in this area?

Relationship and control: Students and teachers access Kita through their school. Kita is the data controller for the processing described and appoints its providers (Microsoft Azure, Neon, AWS, Supabase and the AI-memory provider) as processors under Article 28 agreements. Students have no direct contractual relationship with Kita, so transparency and rights are delivered with and through the school. Expectation of use: Students and their parents/guardians would reasonably expect their code and test results to be processed for educational purposes. However, they may not anticipate that their code and AI conversation history are transmitted to an external third-party AI provider in real time. The Participant Information Sheet provides explicit disclosure of this. The Participant Information Sheet is provided to students and their parents/guardians (and to schools). It is written in age-appropriate language for 13-18 year-olds and links to Kita's privacy notice. Teachers are also data subjects: Kita processes their name, email and platform-usage data to administer accounts, deliver the service and support the pilot evaluation, which they would reasonably expect in their professional role.

Novel technology: AI-powered tutoring in schools is a novel form of educational technology in the UK. The DfE published Generative AI Product Safety Standards in January 2026, requiring that AI products used with learners do not use learner-generated data for model training, operate transparently, and comply with data protection law. Kita closely follows this standard.

Intentional design decisions with privacy implications:

- The AI is clearly labelled as an automated system and does not simulate a human teacher.
- AI agents receive a structured payload (current code, error message, conversation history, lesson context) rather than open-ended access to the student's broader account data.
- AI agents cannot access the internet, external APIs, or any system outside the Kita platform. This is enforced at the architecture level.
- No profiling, automated decision-making, or scoring with significant effects on students beyond the current lesson context is carried out.
- No engagement-maximising features (push notifications, streaks, rewards) are used.
- Students can view their own learning activity, progress metrics and performance statistics within the platform.

Prior concerns and state of technology: LLM-based tutoring systems are increasingly deployed in educational settings globally. Known concerns include: potential for students to share sensitive personal information in chat interfaces; risk of AI providing incorrect or harmful advice; algorithmic bias; data leaks. Each concern is addressed in Step 5 and Step 6.

Describe the purposes of the processing: what do you want to achieve? What is the intended effect on individuals? What are the benefits of the processing – for you, and more broadly?

Primary purposes of processing (these are consistent with the overarching purpose set out in Step 1):

- Delivery of AI-powered tutoring support: to provide students with timely, personalised learning assistance while they complete coding lessons.
- Assessment support: to enable teachers to review student code submissions, I/O test results and assignment progress, reducing teacher marking workload.
- Research and evaluation: to evaluate the effectiveness and safety of AI-assisted coding education as part of the DfE-funded pilot, generating evidence to inform national policy on responsible AI deployment in schools.

Broader benefits: developing students' AI literacy and computing skills; building teacher capability in AI-assisted teaching; widening access to high-quality Computer Science support across state schools; contributing to the DfE's policy framework for AI in education. The pilot is designed to test these benefits rather than assume them.

Step 3: Consultation process

Consider how to consult with relevant stakeholders: describe when and how you will seek individuals' views – or justify why it's not appropriate to do so. Who else do you need to involve within your organisation? Do you need to ask your processors to assist? Do you plan to consult information security experts, or any other experts?

Internal consultation: This DPIA has been reviewed by the Kita leadership team: Heather Lyons (CEO), Nathan Varughese (CTO), Sebastian Vasquez-Lopez (Head of AI Engineering), and Fiona MacNeill (Head of Product / Research Lead). Independent external legal review has been conducted by Tariq Sayfoo, Cripps LLP.

Pilot schools (all): All pilot schools are being consulted as part of this DPIA distribution, prior to the pilot commencing. School DPOs are invited to raise questions or concerns before signing the data processing agreement with Kita. School-level consent is the basis for platform data processing for all participating schools.

Case study schools — parental consent: For the qualitative research strand (focus groups and classroom observations), parental consent is required in addition to school-level consent. Parents and guardians of students aged 13–18 at case study schools receive the University of Brighton's Participant Information Sheet and Consent Form prior to any focus group or observation activity.

Students (13–18): Students at case study schools provide their own assent via the student consent form. Direct consultation with all students across all pilot schools is not proportionate at pilot scale; student interests are represented through the schools and the University of Brighton ethics review. Consistent with the Children's Code.

University of Brighton ethics review: The independent evaluation is subject to University of Brighton Cross-School Research Ethics Committee review. A favourable ethical opinion was confirmed on 1 June 2026.

Step 4: Assess necessity and proportionality

Describe compliance and proportionality measures, in particular: what is your lawful basis for processing? Does the processing actually achieve your purpose? Is there another way to achieve the same outcome? How will you prevent function creep? How will you ensure data quality and data minimisation? What information will you give individuals? How will you help to support their rights? What measures do you take to ensure processors comply? How do you safeguard any international transfers?

Lawful basis: Kita relies on Article 6(1)(f) UK GDPR (legitimate interests) as the lawful basis for processing — delivering AI-supported computing education and generating independent evaluation evidence. This is supported by a Legitimate Interests Assessment (LIA) which weighs these interests against the rights of children and is available on request. No special category data is intentionally processed; should any be disclosed incidentally through the AI chat interface, it is handled through the safeguarding procedure described in Step 2.

Transparency and rights: Individuals are informed through an age-appropriate privacy notice and the Participant Information Sheet (Articles 13–14 UK GDPR and the Children’s Code). Data subjects and their parents/guardians can exercise access, rectification, erasure and objection rights, delivered with and through the school, with Kita’s support.

Processor compliance: All processors (Microsoft Azure, Neon, AWS, Supabase and Letta) are appointed under Article 28-compliant data processing agreements covering security, confidentiality, sub-processor controls, breach notification and deletion.

International transfers: None. All processing takes place in the UK via Azure OpenAI UK South region, confirmed by Microsoft’s UK data-residency commitments and the Azure data processing agreement.

Necessity and effectiveness: The AI tutoring function requires real-time transmission of student code and conversation context to the AI provider to generate meaningful responses. A less privacy-intrusive alternative (e.g. an entirely on-device model) is not currently feasible at the required quality level.

Data minimisation: Students’ full names are never exposed to the AI provider. AI prompts include only code, error messages, and conversation context. No additional personal data (name, age, demographics) is sent to the AI provider. Quiz answers and test results are stored per school to enforce data isolation.

Automated decision-making: The AI system provides educational hints and feedback only. No automated decisions with legal or similarly significant effects on students are made by the AI. Final grades, progression decisions and assessments remain entirely within teacher and school control.

Children’s Code (Age Appropriate Design Code) compliance: describe the measures adopted to comply with the Code and provide an assessment of the risks to children listed in the Code.

Kita has applied the ICO’s Children’s Code in the design and operation of its platform. The following measures have been adopted:

- Best interests of the child: the platform is designed to support learning and independent problem-solving, not to maximise engagement or collect data beyond what is necessary.
- Data minimisation: only data necessary for the tutoring and evaluation purposes is collected. Full names and identifiers are not transmitted to the AI provider.
- No profiling: no profiling or automated decision-making beyond the current lesson context is carried out.
- No nudge techniques: no engagement-maximising features (push notifications, streaks, rewards, social comparison) are used.
- Transparency: the AI is clearly labelled as automated. An age-appropriate privacy notice and Participant Information Sheet are provided to students and parents/guardians.
- Safeguarding: AI agents are instructed via system prompt not to invite personal information and to redirect any safeguarding disclosure to a trusted adult. The AI uses no human persona.
- Parental consent: for case study school activities (focus groups and classroom observations), parental consent is obtained in addition to school-level consent.
- High privacy settings by default: data sharing with the AI provider is limited to the minimum context required for each interaction. Session working memory is cleared at session end.

Assessment of risks to children: The primary risks to children are addressed in Step 5 (R4: memory retention; R5: sensitive disclosures via AI chat; R6: platform data breach). No additional risks specific to the Children’s Code have been identified beyond those already assessed. Residual risks are rated Low following the mitigations set out in Step 6.

Step 5: Identify and assess risks

Describe source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary.	Likelihood of harm	Severity of harm	Overall risk
R1 - Unintentional international data transfer: student code, error messages and AI conversation context are transmitted to Azure OpenAI UK South region. As processing remains within the UK at all times, no international data transfer occurs.	Remote	Minimal	Low
R2 - Data breach at third-party AI provider (Azure OpenAI): a security incident could expose pseudonymised student code, AI conversation history and associated session context.	Remote	Minimal	Low
R3 - Student data used for AI model training without consent: Azure DPA explicitly prohibits use of customer data for model training or product improvement.	Remote	Minimal	Low
R4 - Persistent AI memory retained beyond appropriate period: per-student AI interaction memory (hint requests, recurring errors, lesson goals) persists between sessions. Without a documented deletion schedule, data may be retained longer than necessary.	Possible	Minimal	Medium
R5 - Students sharing sensitive personal information via AI chat: the open-ended chat interface could lead to disclosure of sensitive information (health, family circumstances, safeguarding concerns), which would be transmitted to the AI provider and logged in session history.	Possible	Significant	Medium
R6 - Kita platform data breach (PostgreSQL/S3/Supabase): an attack or misconfiguration could expose student account data, code submissions, test results and quiz answers across all participating schools.	Remote	Severe	Medium

Step 6: Identify measures to reduce risk

Identify additional measures you could take to reduce or eliminate risks identified as medium or high risk in Step 5. (Mitigations below focus on medium risks R4–R6; low-risk items R1–R3 are included for completeness.)

See mitigation table below.

Risk	Options to reduce or eliminate risk	Effect on risk	Residual risk	Measure approved
R1	Use of Azure OpenAI Service deployed in UK South region eliminates the international transfer entirely as data remains within the UK at all times, removing the need for an IDTA or Transfer Risk Assessment.	Eliminated	Low	Yes
R2	Data Processing Agreement with Microsoft Azure covering security obligations, breach notification timelines, and sub-processor restrictions.	Reduced	Low	Yes
R3	Azure DPA explicitly prohibits use of customer data for model training or product improvement. Documented and verified.	Reduced	Low	Yes

R4	Documented retention schedule implemented: working memory cleared at session end; core/long-term memory deleted at course completion or within 12 months of last access; code submissions anonymised after 31 August 2026; anonymised data deleted July 2028. Automated deletion to be implemented and tested before pilot launch.	Reduced	Low	Yes
R5	System prompt instructs all AI agents not to encourage sharing of personal information and to redirect safeguarding concerns to a trusted adult. AI is clearly labelled as automated. Escalation procedure via school Designated Safeguarding Lead in place.	Reduced	Low	Yes
R6	Supabase row-level security (RLS) policies enforce per-user data isolation; all data keyed by school_id; HTTPS enforced for all connections; AWS IAM policies restrict S3 access per environment; encryption key applied per school.	Reduced	Low	Yes

Step 7: Sign off and record outcomes

Item	Name/position/date	Notes
Measures approved by:	Heather Lyons, CEO, Kita Education Ltd — 8 June 2026	Integrate actions back into project plan, with date and responsibility for completion
Residual risks approved by:	Heather Lyons, CEO, Kita Education Ltd — 8 June 2026. No residual high risks identified requiring ICO consultation.	If accepting any residual high risk, consult the ICO before going ahead
DPO advice provided:	Tariq Sayfoo, Cripps LLP (independent external legal review)	DPO should advise on compliance, Step 6 measures and whether processing can proceed
Summary of DPO advice:	External legal review conducted by Tariq Sayfoo, Cripps LLP. Advice: processing may proceed subject to: (i) completion of a data sharing agreement with the University of Brighton prior to any data being shared; and (ii) finalisation of data processing agreements with each pilot school.	
DPO advice accepted or overruled by:	Heather Lyons, CEO, Kita Education Ltd — 8 June 2026	If overruled, you must explain your reasons
Consultation responses reviewed by:	Heather Lyons, CEO, Kita Education Ltd. School DPO responses to be reviewed upon receipt and incorporated as required.	If your decision departs from individuals' views, you must explain your reasons
This DPIA will be kept under review by:	Nathan Varughese, CTO, Kita Education Ltd. Review to be conducted at the end of the pilot (August 2026) and upon any material change to the processing.	The DPO should also review ongoing compliance with DPIA