

AURIVA: ADAPTIVE ENGLISH LANGUAGE LEARNING PLATFORM FOR AUTISTIC CHILDREN IN SRI LANKA

COMMON INTEGRATED SOLUTION REPORT

Group ID: R26-IT-006

Team member	Student ID	Individual component
H. H. Jayasundera	IT22084590	GNN-Enhanced Adaptive Concept Learning Module
R. S. Liluksha	IT22216700	Motor-Informed Adaptive Handwriting Learning
N. G. Samaranayake	IT22266378	Phoneme-Aware Adaptive Pronunciation Support Component
M. C. Gunawardena	IT22166838	Adaptive Dialogue Learning and Progression Tracker

Report submitted in partial fulfilment of the requirements for the

B.Sc. (Hons) Degree in Information Technology

Specialising in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

October 2026

DECLARATION AND SUPERVISOR CERTIFICATION

Candidate declaration. We declare that this common report is our own work and does not incorporate, without acknowledgement, material previously submitted for a degree or diploma at any university or institute of higher learning. To the best of our knowledge, it contains no material previously published or written by another person except where acknowledgement is made in the text. We grant the Sri Lanka Institute of Information Technology the non-exclusive right to reproduce and distribute this report, in whole or in part, in print, electronic, or other media, while retaining the right to use the content in future work.

Academic integrity and AI disclosure. We have disclosed all use of generative AI and other automated tools in accordance with the module policy. We remain responsible for the accuracy, originality, citations, code, data, analyses, and conclusions in this report.

Student name	Student ID	Signature	Date
H. H. Jayasundera	IT22084590		
R. S. Liluksha	IT22216700		
N. G. Samaranayake	IT22266378		
M. C. Gunawardena	IT22166838		

Supervisor certification. The above candidates carried out this research project under my supervision.

Role and name	Signature	Date
Supervisor: Ms. Manori Gamage		
Co-supervisor: Ms. Suriya Kumari		

ABSTRACT

Autistic children in Sri Lankan special education learn English as a second language with few digital tools that follow the national curriculum, suit their communication and motor needs, or give teachers interpretable evidence of progress. This report presents Auriva, a teacher-mediated tablet platform that integrates four components into one system: a graph-informed concept learning module for foundational vocabulary, a motor-informed adaptive handwriting module, a phoneme-aware pronunciation support module, and an adaptive dialogue learning and progression tracker. The components share authentication, student profiles, a PostgreSQL database and one Teacher Workspace. Learning decisions are made by transparent rules in the backend, while machine-learning services are advisory and fall back gracefully, so a child's session never depends on them. Following a design and development approach, the platform was deployed in special education settings and evaluated through unit, scenario and integration testing, controlled data collection, teacher review and classroom use. Over 27 active days the concept module recorded 441 scored attempts from 19 children, with Tier 1 mastery in 80.3 per cent of concept records; visual similarity predicted observed confusion, whereas a graph attention network did not beat heuristic baselines. The handwriting component reproduced its deployed features exactly and passed 7,609 automated tests. Pronunciation scores correlated with teacher scores at $r = 0.73$ on 32 reviewed attempts, against $r = 0.19$ for the originally proposed template matching. The dialogue tracker's pace model beat a literature-based formula on verbal attempts (balanced accuracy 0.695 against 0.498). Teacher usability evidence and the dialogue field study are still pending. The work shows that a rule-governed, teacher-controlled architecture with optional, audited machine learning is feasible in Sri Lankan classrooms, and it reports candid negative results to guide future work.

Keywords: autism spectrum disorder, adaptive learning, English as a second language, teacher-mediated learning, special education technology

ACKNOWLEDGEMENTS

We sincerely thank our supervisor, Ms. Manori Gamage, and our co-supervisor, Ms. Suriya Kumari, for their guidance, constructive feedback and encouragement throughout the research, and for insisting that every claim in this work be supported by evidence.

We are grateful to the principals, special education teachers and staff of the partner special education schools, who gave their time to requirement gathering, accommodated the platform in live classroom practice, reviewed children's attempts, and provided the professional judgement against which the teacher-facing design was repeatedly corrected. We also thank the Ministry of Education of Sri Lanka for the letter that authorised school visits in the Colombo district, and the children and their families and guardians whose participation made the evaluations possible.

We thank the Sri Lanka Institute of Information Technology and the Department of Information Technology for the facilities and academic support provided. Finally, we thank our families and friends for their patience and support.

TABLE OF CONTENTS

1 INTRODUCTION AND PROJECT CONTEXT	1
1.1 Background and context.....	1
1.2 Stakeholders and user needs.....	2
1.3 Alignment with the Sustainable Development Goals	3
1.4 Purpose and scope of the common report	3
2 OVERALL PROBLEM DEFINITION AND FOUR SUBPROBLEMS	5
2.1 Overall problem definition.....	5
2.2 Root causes and overall research gap.....	5
2.3 Subproblem and component mapping.....	6
2.4 Overall aim and group objectives	7
2.5 Scope assumptions and constraints	8
3 OVERALL SOLUTION STRUCTURE.....	10
3.1 Integrated solution overview.....	10
3.2 High level architecture	10
3.3 Component 1 summary	13
3.4 Component 2 summary	14
3.5 Component 3 summary	14
3.6 Component 4 summary	15
3.7 End to end user and data flow	16
4 COMPONENT INTEGRATION AND COHESIVE OPERATION.....	19
4.1 Integration strategy.....	19
4.2 Interfaces data contracts and shared models	20
4.3 Shared nonfunctional requirements.....	21
4.4 Integrated workflow and exception handling.....	21

4.5 Team contribution and dependency matrix.....	22
5 INTEGRATED IMPLEMENTATION AND PROJECT EXECUTION	24
5.1 Common technology stack and infrastructure	24
5.2 Development and integration process	25
5.3 Deployment and operational configuration.....	25
5.4 Evidence of team execution	26
6 INTEGRATED TESTING EVALUATION AND VALIDATION	28
6.1 Evaluation framework and success criteria.....	28
6.2 Integration and end to end testing	28
6.3 Overall results	29
6.4 Stakeholder and user validation	30
6.5 Achievement of common objectives.....	31
6.6 Integrated limitations and threats	32
7 COMMERCIALISATION SUSTAINABILITY AND RESPONSIBLE INNOVATION	33
7.1 Value proposition and beneficiaries.....	33
7.2 Market alternatives and differentiation	33
7.3 Business model and feasibility.....	34
7.4 Adoption intellectual property and compliance	35
7.5 Sustainable Development Goal contribution	37
8 CONCLUSION AND FUTURE DIRECTIONS	38
8.1 Overall conclusion	38
8.2 Principal integrated contribution.....	38
8.3 Future directions.....	39
REFERENCES.....	41
APPENDICES	43

Appendix A Integrated interface specification	43
Appendix B Team contribution declaration	44

LIST OF FIGURES

Figure 3.1: High-level architecture of Auriva showing users, shared services, four components, data stores and external services	12
Figure 3.2: End-to-end user and data flow through the integrated solution.....	17

LIST OF TABLES

Table 1.1: Stakeholders and confirmed needs.....	2
Table 2.1: Subproblems and component mapping.....	6
Table 2.2: Group objectives, components and evaluation evidence	7
Table 4.1: Shared integration surfaces	19
Table 4.2: Principal interfaces and contracts	20
Table 4.3: Shared nonfunctional requirements and their treatment	21
Table 4.4: Failure paths and recovery behaviour	22
Table 4.5: Team contribution matrix	22
Table 4.6: Dependency matrix	23
Table 5.1: Common technology stack and component additions	24
Table 5.2: Deployment and test configurations by component.....	26
Table 5.3: Dated evidence of team execution	27
Table 6.1: Evaluation designs by component	28
Table 6.2: Integration and system testing evidence	29
Table 6.3: Principal results by component.....	29
Table 6.4: Stakeholder and user validation	31
Table 6.5: Achievement of group objectives	31
Table 7.1: Buyers, users and beneficiaries.....	33
Table 7.2: Differentiation of the integrated product	34
Table 7.3: Main cost areas	35
Table A.1: Interface catalogue	43
Table A.2: Shared entities	44
Table B.1: Team contribution declaration	44

LIST OF ABBREVIATIONS

Abbreviation	Meaning
AAC	Advanced Audio Coding
ABA	Applied Behavior Analysis
AI	Artificial intelligence
API	Application Programming Interface
ARI	Adjusted Rand Index
ASD	Autism spectrum disorder
CI	Confidence interval
DTW	Dynamic Time Warping
EQ	Evaluation question
GAT	Graph Attention Network
GKB	Graph Knowledge Base
GO	Group objective
GOP	Goodness of Pronunciation
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
JSON	JavaScript Object Notation
JWT	JSON Web Token
MFCC	Mel-frequency cepstral coefficient
ML	Machine learning
ORM	Object Relational Mapper
OWASP	Open Worldwide Application Security Project
RC	Research contribution
REST	Representational State Transfer
SDG	Sustainable Development Goal

Abbreviation	Meaning
SHAP	SHapley Additive exPlanations
SUS	System Usability Scale
TC	Test case
TLS	Transport Layer Security

1 INTRODUCTION AND PROJECT CONTEXT

1.1 Background and context

Autism spectrum disorder (ASD) is a neurodevelopmental condition marked by differences in social communication and interaction and by restricted, repetitive patterns of behaviour [1]. The Global Burden of Disease Study 2021 estimated that about 61.8 million people worldwide are autistic, roughly 1 in 127 [2]. In Sri Lanka, an early community screening study estimated a prevalence of 1.07 per cent among children aged 18 to 24 months [3], and by 2025 more than 9,000 children had been formally identified as autistic, with many more thought to be unidentified [4]. Special education provision is correspondingly stretched: the Ministry of Education's 2024 school census records 7,574 students in special education units within government schools, and 27 government-assisted special schools with 1,956 students and 468 teaching staff [5].

English is a second language for most Sri Lankan children and is the language of higher education and much of employment. For an autistic child, the communication differences of autism combine with the demands of learning a new language, so functional English is among the hardest skills to acquire in an ordinary classroom. Special education teachers are expected to deliver the national English curriculum [6] to several children with very different abilities at once. They often rely on improvised materials or on mainstream applications that were not designed for autistic learners, follow no Sri Lankan syllabus, and give the teacher little evidence of what a child has actually learnt.

Auriva is a tablet-first adaptive English learning platform built for this setting. It rests on one organising decision: the teacher, not the parent and not the algorithm, runs the learning session. The teacher signs in, selects a child from an assigned list and conducts the session on a shared tablet while the child works directly with the on-screen activities [7]. The platform was developed by four students as four research components, each covering one foundational English skill: vocabulary concepts, handwriting, pronunciation, and functional spoken dialogue. The

components were integrated into a single application with one login, shared student profiles and a shared database.

1.2 Stakeholders and user needs

Stakeholders were identified through requirement-gathering sessions with the principal, teachers and staff of a partner special education school, through teacher review of children's attempts, and through supervised classroom and field sessions at several schools in the Gampaha and Colombo districts. The Ministry of Education issued a letter authorising the group's school visits in the Colombo district, and each participating school gave permission after receiving it [8]. Table 1.1 summarises the stakeholders and the needs confirmed through this engagement.

Table 1.1: Stakeholders and confirmed needs

Stakeholder	Role	Confirmed needs
Autistic children (approx. 5 to 15 years)	Primary beneficiary; interacts under teacher supervision	Structured, visual, repetitive content; clear and gentle feedback; consistent pacing; a non-verbal way to take part; a calm, predictable interface with no failure state
Special education teachers (approx. 21 to 60 years)	Primary user; runs every session	Curriculum-aligned, ready-to-use content; per-child scores and trends at a glance; explanations of why a child is held back or moved on; authority to stop a session or override the system
Principals and special education coordinators	Buyer, approver and system administrator	Predictable licensing, evidence of progress for reporting, one system for all children and teachers
Parents and guardians	Secondary beneficiary; give consent	Assurance of safe use of children's data; worksheets for guided practice at home (handwriting)
Ministry of Education	Authorising body; potential procurement route	Alignment to the national syllabus; evidence of outcomes at pilot scale

Teacher engagement changed the design in concrete ways. Teachers valued the concept module's handover flag because it bounded the number of consecutive failures a child could experience, asked that generated summaries describe rather than interpret, and requested vocabulary beyond the original nine categories [7]. Test rounds with children showed that conversation practice was too hard and that freshly generated conversations unsettled them, so the dialogue component's planned third level was removed [8]. The pronunciation component routes uncertain scores to a teacher review queue so that the teacher remains the final judge [9].

1.3 Alignment with the Sustainable Development Goals

The project is aligned primarily with Sustainable Development Goal (SDG) 4, Quality Education [10]. It targets target 4.5, equal access to education for persons with disabilities, and target 4.a, education facilities and learning environments that are disability-sensitive and inclusive. It also contributes to SDG 10, Reduced Inequalities, in particular target 10.2 on the social inclusion of all irrespective of disability, by helping autistic children acquire the everyday English on which later education and employment depend. Two of the individual theses additionally relate design choices that protect children's sensory and emotional well-being to SDG 3 and the use of information technology for inclusive education to SDG 17 [7]; the group treats these as supporting rather than primary alignments.

These are statements of intended contribution, not measured impact. The evidence in this report concerns system behaviour, short-term learning signals and teacher judgement in small samples; it does not show long-term educational outcomes. Section 7.5 evaluates the contribution and its possible negative effects.

1.4 Purpose and scope of the common report

This common report presents the integrated problem and the integrated solution. It explains how four components were designed as one platform, how they interoperate, how the complete solution was implemented and evaluated, and what shared value it offers. It does not repeat the component-level research. The detailed methods, experiments, datasets, statistical analyses and discussion are in the four individual theses: the Graph Neural Network-Enhanced Adaptive Concept Learning Module by H. H. Jayasundera [7], the Motor-Informed Adaptive Handwriting Learning component by R. S. Liluksha [11], the Phoneme-Aware Adaptive Pronunciation Support Component by N. G. Samaranayake [9], and the Adaptive Dialogue Learning and Progression Tracker by M. C. Gunawardena [8].

Chapter 2 defines the overall problem and the four subproblems. Chapter 3 presents the integrated solution and its architecture. Chapter 4 describes how the components were integrated and operate together. Chapter 5 covers the shared implementation and project execution. Chapter 6 reports integrated testing, evaluation and validation.

Chapter 7 considers commercialisation, sustainability and responsible innovation, and Chapter 8 concludes with future directions. Results that were still pending when the individual theses were completed are identified as pending throughout and are not presented as findings.

2 OVERALL PROBLEM DEFINITION AND FOUR SUBPROBLEMS

2.1 Overall problem definition

The common problem is stated as follows. Autistic children in Sri Lankan special education settings who are learning English as a second language have no curriculum-aligned, adaptive, teacher-controlled learning tool that teaches the foundational skills of vocabulary, handwriting, pronunciation and functional spoken use, and that gives their teachers interpretable evidence of each child's progress.

Four observations from the individual problem statements support this definition. Digital tools deliver isolated activities without mastery-gated progression or graduated remediation [7]. Handwriting applications judge the finished trace and say little about how a letter was produced [11]. Pronunciation and speech assessment are built and tuned for adult or typically developing speakers, and detailed scores can discourage children who are sensitive to failure [9]. Vocabulary tools teach words rather than when to use them, and fuzzy matching of speech rewards the first-language transfer errors a Sinhala-speaking child most needs to correct [8]. In every case the teacher receives little explanation of the system's decisions.

2.2 Root causes and overall research gap

The component analyses point to six shared root causes.

- **Generic design.** Existing applications are global in content, not aligned to the Sri Lankan syllabus, and often assume continuous connectivity and a parent-child rather than a teacher-mediated model. A feature audit of two widely used autism-focused applications, Otsimo and Speech Blubs, found neither provided a tiered mastery architecture, a graduated relearn pathway, teacher-mediated control or a relational model of concepts [7].
- **Product-only assessment.** Handwriting tools evaluate whether a trace was completed rather than the movement that produced it, although process measures such as smoothness, pauses and direction are known to carry information [11], [12].

- **Assessment tuned to the wrong population.** General speech recognition is less accurate for non-native child speakers, and word-level template matching against an adult reference proved unreliable for autistic children's voices [9], [8].
- **Failure-oriented feedback.** Tools expose scores and failures to the child and give no non-verbal route, so children who do not speak cannot progress and echoed speech can be mistaken for learning [9], [8].
- **Opaque adaptation.** Where adaptation exists it is rarely explained to the teacher, who is excluded from the instructional loop [7], [11].
- **Thin local evidence.** Reported accuracy for artificial intelligence (AI)-driven tools for autistic children is mostly obtained on retrospective data or with few neurotypical evaluators, and local research platforms target general skills rather than curriculum-aligned English [7], [13].

The overarching gap is therefore the absence of a single platform that combines four things: curriculum-aligned teaching across the four English skills; mastery-based adaptation whose decisions are transparent and controlled by the teacher; speech and motor assessment suited to autistic learners in Sri Lanka; and prospective evidence from real classrooms. No single component can close this gap, because a child's English development depends on understanding words, forming them in writing, saying them, and using them in context.

2.3 Subproblem and component mapping

The overall problem was divided into four subproblems that do not overlap in method or data but depend on common platform services. Table 2.1 maps each subproblem to its component and owner.

Table 2.1: Subproblems and component mapping

Component and owner	Subproblem addressed	Principal research focus
1. Concept Learning Module H. H. Jayasundera (IT22084590)	Vocabulary is taught as isolated activities with no mastery gate, no remediation that varies the stimulus, and no model of how concepts relate or are confused	Tiered, mastery-gated vocabulary teaching; Graph Knowledge Base; concept similarity features; Graph Attention Network evaluation
2. Motor-Informed Handwriting R. S. Liluksha (IT22216700)	Handwriting practice judges the final trace and gives teachers no evidence of how the letter was produced	Trajectory capture, direction-invariant Dynamic Time Warping (DTW), interpretable motor measures, transparent three-attempt workflow, exploratory clustering

Component and owner	Subproblem addressed	Principal research focus
3. Phoneme-Aware Pronunciation N. G. Samaranayake (IT22266378)	Pronunciation is judged subjectively by ear, and automatic assessment is not designed for autistic children's speech	Three-layer scoring cascade (Mel-frequency cepstral coefficient (MFCC) with DTW, phoneme-level Goodness of Pronunciation, word verification), teacher-review recalibration, child-speech adaptation, adaptive next word
4. Adaptive Dialogue Tracker M. C. Gunawardena (IT22166838)	Children learn words but not when to use them; speech assessment penalises or rewards the wrong errors; pace is not adapted or explained	Three-phase contextual word learning, Sinhala transfer-aware speech assessment, latency-based echolalia detection, explainable learning-pace prediction

The subproblems are interdependent in three ways. They share the child (one student record, one set of sensory preferences), the teacher (one login, one workspace) and the data platform. They cover successive aspects of the same skill: a child can recognise a word through the concept module, write it with the handwriting module, say it with the pronunciation module and use it with the dialogue module. The dialogue design can also draw on concept, speaking and writing mastery when generating Level 2 activities, although that cross-component use was not evaluated [11]. All four are necessary because removing any one leaves a skill with no adaptive support and no teacher evidence, which would reproduce the fragmentation the project set out to remove.

2.4 Overall aim and group objectives

Common aim. To design, build, integrate and evaluate one teacher-mediated, curriculum-aligned adaptive English learning platform for autistic children in Sri Lankan special education, in which four components teach foundational vocabulary, handwriting, pronunciation and functional dialogue, and in which adaptive decisions are transparent and the teacher keeps professional authority.

The aim is broken into six measurable group objectives (GO1 to GO6), mapped in Table 2.2 to the components that serve them and to the evidence used to judge them in Chapter 6.

Table 2.2: Group objectives, components and evaluation evidence

ID	Group objective	Components	Evidence used
----	-----------------	------------	---------------

ID	Group objective	Components	Evidence used
GO1	Deliver four curriculum-aligned learning components inside one deployed application with shared authentication, student profiles and database	All	Integrated build; shared-surface checks (Section 6.2)
GO2	Provide mastery-based adaptive progression in each component, with decisions made by explicit rules	All	Rule scenarios, mastery tests, adaptive-behaviour verification (Sections 6.2, 6.3)
GO3	Give teachers interpretable evidence of each child's progress through one consistent workspace	All	Teacher Workspace tabs; reports; teacher feedback (Section 6.4)
GO4	Ensure intelligence layers can never interrupt or harm a child's session	1, 2, 3, 4	Fallback, timeout and kill-switch verification (Section 6.2)
GO5	Evaluate with autistic learners and teachers in real settings under consent and ethical oversight	All	Classroom deployment, controlled collection, teacher review, field study (Sections 6.3, 6.4)
GO6	Report evidence candidly, including negative and pending results, and establish a credible adoption pathway	All	Limitations; commercialisation analysis (Section 6.6, Chapter 7)

2.5 Scope assumptions and constraints

Scope. The solution covers foundational English up to Grade 3 level, aligned to the National Institute of Education syllabus [6], for autistic children in Sri Lankan special education. It is designed for teacher-mediated use on shared tablets. English is the target language, with Sinhala support in some activities.

Exclusions. The platform is an educational support tool. It does not diagnose autism, estimate severity or replace therapy, and it is not a medical device. It is not designed for independent child use or unsupervised home use. It has no parent application, covers English letters only (Sinhala and Tamil scripts are future work), and does not assess long-term retention.

Assumptions. The supervising teacher is the right authority on whether a child is engaged and whether a session should continue. The national curriculum is the correct instructional target. Behavioural logs are a valid signal of conceptual difficulty when verbal self-report is unavailable. Partner schools have suitable tablets and a local network.

Constraints and dependencies. Evaluation samples are small and drawn from a limited number of schools within one academic year, with data collection ending in October 2026. Some components depend on external speech and language services

whose quotas and costs bound scaling. Several research builds ran with prototype-level security. All four components depend on the shared backend, database and authentication layer, so a change to a shared schema or contract affects every module.

3 OVERALL SOLUTION STRUCTURE

3.1 Integrated solution overview

Auriva is delivered as one system: a React Native tablet application, a Node.js backend exposing a Representational State Transfer (REST) interface, a set of supporting Python services, a PostgreSQL system of record and a Neo4j graph database. Its central value proposition is that a teacher with no technical background can run a structured English session for an autistic child on a shared tablet, and afterwards see, for each of four skills, what the child did, what the system decided and why.

Four design principles hold the four components together, and each component implements them in its own domain.

- **Teacher mediation.** The teacher chooses the child, the module and the activity, can stop a session at any time, and sees explanations rather than bare scores. The child never signs in and never reaches an adult-facing control unaided.
- **Rules decide, models advise.** Mastery, progression, remediation and unlocking are governed by explicit, auditable rules in the backend. Machine-learning outputs (graph models, phoneme scores, motor clusters, pace predictions) inform the teacher or adjust difficulty within safe bounds but never gate a child's progress on their own.
- **Graceful degradation.** Every intelligence service has a timeout, a fallback or a switch, so that a slow or absent service is invisible to the child.
- **Child-safe feedback.** Children see no numerical scores and no failure messages. Errorless and non-verbal pathways let every child take part.

The end-to-end outcome is a single record per child, spanning concept mastery, handwriting evidence, pronunciation scores and dialogue progress, presented through one Teacher Workspace with a tab per module and one consistent report layout.

3.2 High level architecture

Figure 3.1 shows the architecture. Users reach the platform only through the tablet application, whose shell provides login, the student profile, sensory preferences and

the Teacher Workspace. Each module contributes its own screens inside this shell. The application calls the Node.js and Express backend over REST with a JavaScript Object Notation (JSON) Web Token (JWT) on every request. The backend holds the business logic and the authoritative decision rules, and writes synchronously to PostgreSQL. Intelligence services run as separate Python processes that only the backend calls, never the tablet: the Graph Knowledge Base (GKB) service over Neo4j, the handwriting motor-pattern service, the pronunciation Goodness of Pronunciation (GOP) worker with word verification, and the dialogue Flask microservice.

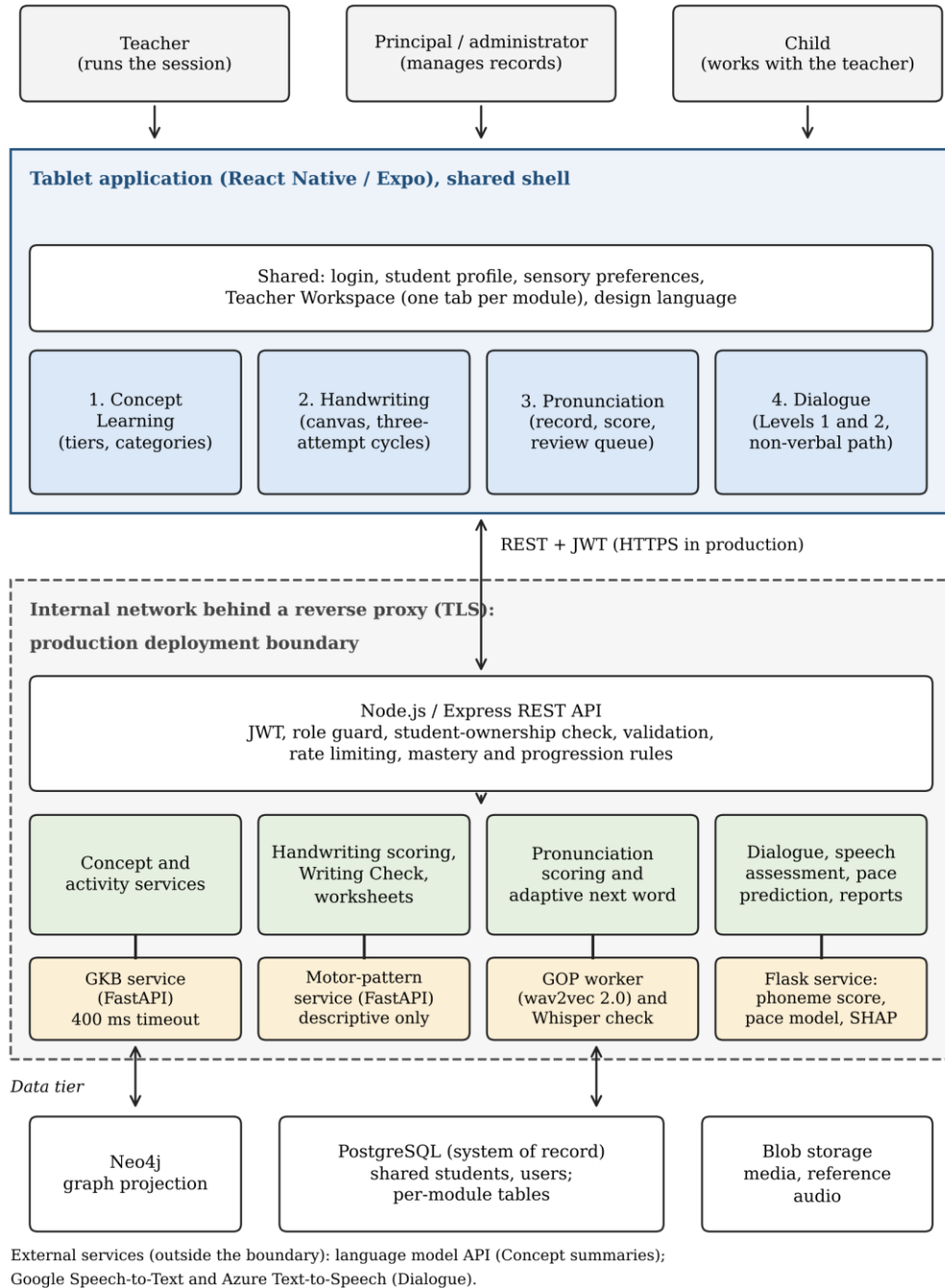


Figure 3.1: High-level architecture of Auriva showing users, shared services, four components, data stores and external services

In the production deployment the backend and the Python services are reachable only on an internal container network behind a reverse proxy that terminates Transport Layer Security (TLS) and is the only publicly exposed service [7]. Three external managed services lie outside this boundary: a language model Application Programming Interface (API) used for teacher-facing summaries in the concept

module, and Google Speech-to-Text and Azure Text-to-Speech used by the dialogue component. The pronunciation component's speech models run on the project's own server, so its recordings are not sent to third-party speech services [9].

3.3 Component 1 summary

Component and owner. Graph Neural Network-Enhanced Adaptive Concept Learning Module; H. H. Jayasundera (IT22084590).

Subproblem. Foundational vocabulary is taught without mastery gates, varied remediation or any model of how concepts relate and are confused.

Input. A teacher-selected category and concept from 96 curriculum-aligned concepts in nine categories, each with a photograph, cartoon, short animation, narration audio and attribute metadata; the child's tap responses; the stored sensory preferences.

Core method. Each concept is taught through three tiers: Find it (visual familiarisation and identification), Name it (word recognition) and Watch and colour (attribute introduction, unscored). A rule-based engine enforces an Applied Behavior Analysis (ABA) mastery criterion of two correct out of three in a block, adjusts the number of distractors with consecutive successes or errors, and runs a three-level relearn pathway that varies the stimulus (photograph, cartoon, animation) before raising a teacher handover flag. A Neo4j Graph Knowledge Base records students, concepts, categories and weighted confusions, supplies confusion-informed distractors under a 400 millisecond timeout, and underpins a similarity feature layer and a shadow-mode Graph Attention Network (GAT) [14].

Output. Tier mastery records, relearn and handover state, a concept progress tab in the Teacher Workspace, and plain-language teacher summaries.

Contribution to the solution. Provides the vocabulary foundation and the relational model of concepts; demonstrates the pattern of an advisory intelligence layer behind hard fallbacks that the other components follow. Detailed methods and results are in [7].

3.4 Component 2 summary

Component and owner. Motor-Informed Adaptive Handwriting Learning; R. S. Liluksha (IT22216700).

Subproblem. Handwriting practice judges the finished trace and gives teachers no evidence of how a letter was produced.

Input. Time-ordered stylus coordinates with stroke boundaries captured on a tablet, compared with shared letter templates; teacher requests for Writing Checks and worksheets.

Core method. Each attempt is compared with a reference using direction-invariant Dynamic Time Warping (DTW) and summarised by interpretable measures of smoothness, pause control, direction control and speed variation. A weighted, null-aware letter Motor Score drives three-attempt cycles with decreasing support; mastery is decided from the unguided third attempt at a threshold of 70, with at most three cycles per letter and case per day, exact-letter remediation, Additional Home Practice and rule-based word learning. Exploratory K-Means clustering is served separately and never influences decisions.

Output. Motor Score and exact mastery evidence, a Teacher Report, teacher-initiated Writing Check results, printable A4 worksheets with photo upload for review, and descriptive motor-pattern summaries.

Contribution to the solution. Adds process-level evidence for the motor side of English, and a worked example of keeping machine learning descriptive and outside decisions. Detailed methods and results are in [11].

3.5 Component 3 summary

Component and owner. Phoneme-Aware Adaptive Pronunciation Support; N. G. Samaranayake (IT22266378).

Subproblem. Pronunciation is judged by ear, and automatic assessment is built for adult or typically developing speakers.

Input. A teacher-chosen mode (alphabet or word), category and target word from 78 items (26 letters and 52 words in four categories) with reference audio; the child's

recording (mono Advanced Audio Coding (AAC), 16 kHz); the student's population group and reduced-stimulation setting.

Core method. After picture, reference audio and phoneme-level support, the child completes listen-and-choose, an errorless sound-segmentation game and spoken-word practice. Recordings pass a three-layer cascade: Layer 1, an accept-only Mel-frequency cepstral coefficient (MFCC) and DTW comparison; Layer 2, phoneme-level Goodness of Pronunciation scoring [15] with a wav2vec 2.0 model [16] and Whisper word verification [17]; Layer 3, recalibration from teacher reviews that is applied only when held-out validation shows an improvement. The next word is chosen from the weakest phoneme, word difficulty and recurring difficulty. A pipeline that separates child and teacher speech from classroom recordings supports a child-speech adaptation model.

Output. Encouraging feedback only for the child; word and phoneme scores, replay, history and an uncertainty-ranked review queue for the teacher.

Contribution to the solution. Supplies objective, phoneme-level pronunciation evidence and the platform's model for teacher-in-the-loop review. Detailed methods and results are in [9].

3.6 Component 4 summary

Component and owner. Adaptive Dialogue Learning and Progression Tracker; M. C. Gunawardena (IT22166838).

Subproblem. Children learn words but not when and how to use them, and existing speech assessment mis-scores Sinhala transfer errors and echoing.

Input. A teacher-selected category or topic; the child's spoken or non-verbal responses; the child's own profile details for personalised sentences; 30 curriculum words and phrases in Greetings, Magic Words and Abilities.

Core method. Level 1 teaches each word in three phases (Familiarisation, Production, Contextual Awareness) and Level 2 teaches personalised sentences for three Grade 3 speaking topics, each traced to the national activity book, with a non-verbal pathway throughout. Three research contributions (RC) address problems the

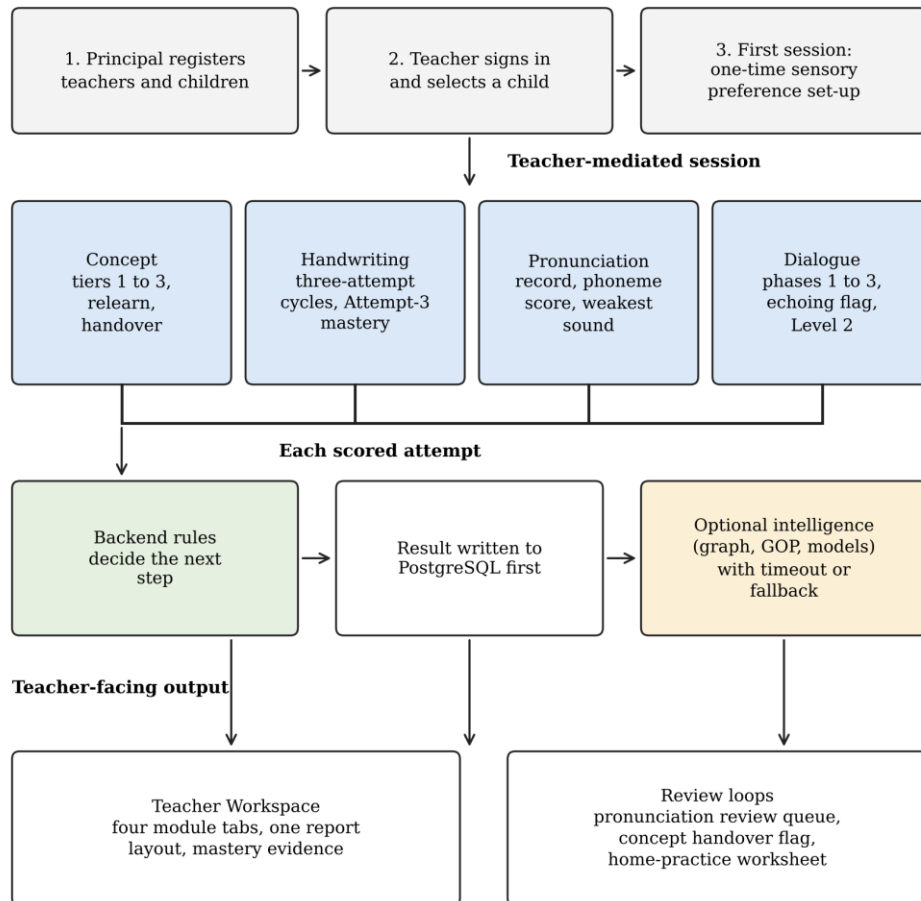
original design could not solve: phoneme-aware assessment that gives partial credit to Sinhala transfer errors (RC1); a latency-based rule that flags probable immediate echolalia without penalising the child (RC3); and a two-tier model that predicts learning pace from a child's first session with a word and explains the prediction to the teacher (RC2).

Output. Scores, next-word and difficulty decisions, echolalia flags, explained pace predictions and Level 1 and Level 2 teacher reports.

Contribution to the solution. Moves the platform from recognising and producing words to using them, and shows how a small-data learning model can be audited and gated. Detailed methods and results are in [8].

3.7 End to end user and data flow

Figure 3.2 traces a user and a data item through the complete solution. The numbered steps below correspond to it.



The child sees only encouraging feedback; scores and diagnostics go to the teacher.

Figure 3.2: End-to-end user and data flow through the integrated solution

1. The principal creates teacher and student records once; they are immediately available to all four modules, so no per-module registration is needed.
2. A teacher signs in and selects a child from the assigned list. Role and student-ownership checks apply to every later request.
3. On the child's first session the one-time preference set-up stores avatar, theme and font choices against the student record. These preferences are then applied by all four modules.
4. The teacher opens a module tab and starts an activity. The child interacts on the tablet; the teacher can stop at any time.

5. Each scored attempt is sent to the backend, where explicit rules decide the next step. The attempt is written to PostgreSQL first; optional intelligence services are then consulted or updated, in parallel or under a timeout.
6. Results appear in the Teacher Workspace. Low-confidence pronunciation scores enter a review queue, repeated concept failures raise a handover flag, and repeated handwriting difficulty can generate a home-practice worksheet. The child sees only encouraging feedback.

4 COMPONENT INTEGRATION AND COHESIVE OPERATION

4.1 Integration strategy

The four components were developed in parallel, each member working on a personal branch, and were then merged into shared application and backend repositories by the group leader, who also deployed the integrated system. Each member's Python service stayed in its own repository as a separate microservice. Integration followed an interface-first approach built on four shared surfaces that were fixed early and that every module uses unchanged.

Table 4.1 lists the four shared surfaces and the member who owns each.

Table 4.1: Shared integration surfaces

Surface	What is shared	How each module uses it
Authentication and authorisation	Token verification, role guard, student-ownership check	All routes of every module sit behind the same guards; a teacher can reach only children assigned to them
Student and teacher data model	Student and teacher records, population group, avatar and theme, reduced-stimulation setting	Modules add their own tables but reference the common records
Teacher Workspace	Student profile with one tab per module, common card and report layout	Each module renders its progress in the shared layout so a teacher learns one reporting interface
Sensory preference layer	Avatar, theme and font captured once in the first-session set-up	Applied by all modules from the single stored record
Deployment	One build, one backend, one database	Services are containerised and run as one composed stack in the production deployment

Integration risk was managed in four ways. First, intelligence services were isolated so that no module's decision logic depends on a model. Second, each component kept its learning logic stable during integration: the dialogue component's merge required only frontend changes (restyled screens and extended Level 2 reports) and no changes to its assessment or models [8]. Third, every schema change was applied through a reviewed migration where migrations were used, and database access was limited to the four members. Fourth, pre-session smoke tests (Section 5.3) confirmed that the build, the backend and the speech services were reachable before each testing day.

4.2 Interfaces data contracts and shared models

Table 4.2 summarises the principal interfaces. The complete catalogue is in Appendix A. All client-to-backend traffic uses JSON over REST with a JWT bearer token [18]. The tablet never calls a Python service directly.

Table 4.2: Principal interfaces and contracts

Interface	Direction	Contract and error handling
Tablet to backend	App to Node.js API	REST/JSON; JWT on every request; request validation and rate limiting; role and ownership checks
Backend to PostgreSQL	Node.js to database	Sequelize Object Relational Mapper (ORM); authoritative interaction, progress and report records; per-module tables referencing shared student and teacher records
Concept backend to GKB service	Node.js to FastAPI over Neo4j	Writes are fire-and-forget; reads have a hard 400 ms timeout with a fallback chain (graph, interaction log, curriculum sequence); an environment flag disables inference entirely
Handwriting backend to motor-pattern service	Node.js to FastAPI	Three letter features in, cluster identifier and descriptive summary out; not reachable from the app; no operational decision depends on it
Pronunciation backend to scoring workers	Node.js to Python GOP worker and Whisper	Scoring endpoint rate-limited to 30 requests per 5 minutes per teacher; low-confidence results go to the review queue
Dialogue backend to Flask microservice	Node.js to Flask, port 5001	Endpoints /health, /score-phoneme, /predict-trajectory, /explain-trajectory; if the phoneme service does not answer within 2 seconds, fuzzy matching scores the attempt; if the service is down, the backend still decides every outcome
External services	Backend to third parties	Language model API (concept summaries, anonymised statistics); Google Speech-to-Text and Azure Text-to-Speech (dialogue); each failure degrades to a defined fallback

Shared identifiers and models. The student identifier is the join key across all modules. The population group (ASD level or learning disability, as entered by the school) is read by the pronunciation component to drive review coverage and per-population calibration [9]. The avatar theme and reduced-stimulation setting are read by several modules. **Validation, versions and ownership.** The backend validates all inputs. Each model artefact carries its own versioning: the dialogue model file stores its column list and label encoder so a retrained model can replace it without code changes [8], and the concept analyses run on versioned dataset snapshots with content hashes [7]. Each member owns the contract of their own service and the

tables it writes; the shared student, teacher and authentication surfaces are owned jointly.

4.3 Shared nonfunctional requirements

Table 4.3 states how the shared nonfunctional requirements are met and, equally important, where the current builds fall short.

Table 4.3: Shared nonfunctional requirements and their treatment

Requirement	Treatment across the platform	Status and limits
Performance	Intelligence calls are bounded: 400 ms graph read timeout; 2 s phoneme-service timeout; pronunciation scoring about 3.0 s per attempt once models are loaded	Pronunciation cold start about 12.8 s; handwriting scoring time not measured; no load testing reported
Usability	Landscape child screens, stored sensory preferences, one consistent teacher workspace, no scores or failure messages for the child	No System Usability Scale (SUS) result is yet reported for any component
Accessibility	Sensory themes, font and avatar choice, reduced-stimulation setting, non-verbal pathways, errorless activities	Stylus input only was evaluated for handwriting; finger input untested
Security	JWT authentication, role and ownership checks, hashed passwords, security headers, rate limiting; reviewed against the Open Worldwide Application Security Project (OWASP) Top 10 [19]	Some research builds used local Hypertext Transfer Protocol (HTTP) rather than HTTPS (HTTP Secure), development secrets and unauthenticated service calls; production hardening is required (Section 7.4)
Privacy	Pseudonymised research data, no video, minimal personal data, access limited to the research team, alignment with the Personal Data Protection Act [20]	Practices differ by component (Section 7.4); consent for secondary use of data is not uniformly confirmed
Reliability	Fire-and-forget graph writes, fallbacks for every read, idempotent attempt finalisation, session resume after interruption	Automated tests are mostly unit-level with mocked dependencies; database-backed integration was largely checked manually
Maintainability	Decision logic as pure functions with named constants; reviewed migrations; research code separated from serving code	Two defects in the dialogue component remain open (Section 6.6)
Scalability	Stateless backend instances, independently scalable services, managed databases	Tested at pilot scale only; external speech quotas bound growth
Professional standards	ABA mastery criteria, curriculum traceability, non-diagnostic positioning, teacher final authority	Formal accessibility audit not performed

4.4 Integrated workflow and exception handling

In normal operation the workflow is the one in Section 3.7: a teacher-led session, a rule-based decision per attempt, a synchronous write to PostgreSQL and asynchronous or bounded calls to intelligence services. The guiding correctness property, stated explicitly in the concept thesis, is that the intelligence layer's absence must be invisible to the child. Table 4.4 lists the main failure paths and the behaviour verified or designed for each.

Table 4.4: Failure paths and recovery behaviour

Failure	Behaviour	Component
Graph service unavailable or slower than 400 ms	Distractors fall back to the interaction log, then to the curriculum sequence; no visible delay (verified in scenario tests)	Concept
Language model summary fails	Numeric dashboard renders and states that the summary is unavailable; no value is fabricated (verified)	Concept
Child fails relearn Level 3	Handover flag to the teacher; no further looping	Concept
Motor-pattern service unavailable	Letter practice and word learning continue because no decision uses its output; Teacher Report behaviour in this case was not separately evaluated	Handwriting
Attempt finalised twice or Attempt 3 technically incomplete	Only one result recorded; Attempt 3 can be retried without using up the cycle (test cases (TC) TC17 and TC10 defined)	Handwriting
Phoneme service silent or slower than 2 s	Fuzzy-matching fallback produces the score; pre-session checklist and smoke test prevent silent fallback	Dialogue
Trajectory model unsure or its flag off	Literature-based formula or "typical" label is used; the model can never change recorded mastery	Dialogue
Low-confidence pronunciation score	Attempt is flagged for teacher review (26.1 per cent of attempts in the study) rather than shown as certain	Pronunciation
Layer 1 template match unreliable	Layer 1 may accept but never fail a child; uncertain attempts escalate to phoneme scoring, then to the teacher	Pronunciation

4.5 Team contribution and dependency matrix

Table 4.5 maps each student to their component and recorded work, and Table 4.6 records which component depends on which. The shared platform, integration and deployment were provided by the group, with integration and deployment carried out by the group leader. Appendix B records the agreed contribution declaration for confirmation by all four members.

Table 4.5: Team contribution matrix

Member	Component	Research tasks	Development and test evidence
H. H. Jayasundera (IT22084590)	1. Concept Learning	Similarity-feature study; Graph Attention Network evaluation with five baselines; exposure-policy confound analysis	Three-tier engine, relearn pathway, Graph Knowledge Base service, teacher analytics; 14 adaptive scenarios verified; 21-child classroom deployment
R. S. Liluksha (IT22216700)	2. Handwriting	Direction-invariant DTW features; Motor Score; clustering and stability analyses on two datasets	Capture, scoring, cycles, Writing Check, worksheets, ML service; 7,609 automated tests passed on 30 Aug 2026; controlled collection from 7 learners
N. G. Samaranayake (IT22266378)	3. Pronunciation	Three-layer cascade; agreement with teacher scores; child-speech adaptation pipeline	Recording, scoring, review queue, adaptive next word; 87 backend and 16 frontend tests passed; 249 in-app attempts

Member	Component	Research tasks	Development and test evidence
M. C. Gunawardena (IT22166838)	4. Dialogue	Sinhala-transfer assessment; echolalia rule; explainable trajectory model; three audits	Level 1 and 2 flows, microservice, reports, Level 2 audio; 114 unit tests passed; supervised field study begun 7 Oct 2026

Table 4.6: Dependency matrix

Component	Depends on the shared platform for	Depends on other components	Others depend on it for
1. Concept	Authentication, student model, workspace	None for operation	Sensory preference layer shape; reference pattern for fallbacks
2. Handwriting	Authentication, student model, workspace	None for operation	Letter practice evidence; dialogue design may read writing mastery (not evaluated)
3. Pronunciation	Authentication, student records and population field, backend and database	None for operation	Speech-quality and review-queue pattern
4. Dialogue	Authentication, student profile, database, deployment	Design may read concept, speaking and writing mastery for Level 2 (not evaluated)	Level 1 and 2 reports in the common workspace

5 INTEGRATED IMPLEMENTATION AND PROJECT EXECUTION

5.1 Common technology stack and infrastructure

The components deliberately share one technology base so that integration required few translation layers. Table 5.1 lists the common technologies and each component's additions, with the reason for the shared choice.

Table 5.1: Common technology stack and component additions

Layer	Common technology	Component-specific additions
Mobile client	React Native 0.81 with Expo SDK 54 and React 19; React Navigation 7; Axios; react-native-svg	Concept: Reanimated, Lottie, expo-audio. Handwriting: drawing canvas, PDF rendering. Pronunciation: expo-av recording, Zustand store. Dialogue: expo-print, Sinhala spelling helper
Backend	Node.js with Express 5; Sequelize 6; JWT authentication; security headers; rate limiting	Dialogue: ffmpeg audio conversion, Jest 30 tests. Pronunciation: Winston logging, scoring orchestration
Data	PostgreSQL system of record (managed); Neo4j graph (managed); blob storage for media	Concept: Graph Knowledge Base on Neo4j. Pronunciation: stored recordings with retention script
Python services	FastAPI or Flask microservices; scikit-learn	Concept: GKB service and offline graph research code. Handwriting: motor-pattern model. Pronunciation: GOP worker (wav2vec 2.0), Whisper. Dialogue: phoneme scoring, Random Forest, SHapley Additive exPlanations (SHAP)
External services	Language model API; Google Speech-to-Text; Azure Text-to-Speech	Concept: summaries. Dialogue: transcription and personalised audio
Tooling	Git and GitHub; Visual Studio Code	Dialogue: Google Colab for model training. Handwriting: Python 3.13 research environment

A shared stack was chosen for four reasons: a single client codebase reaches Android and iOS tablets; one Node.js backend and one relational database give every module the same authentication, ownership checks and reporting joins; keeping intelligence in separate Python services lets each research strand use the libraries it needs without enlarging the production image; and managed data services reduce operational load for a small team. The shared code lives in the auriva-frontend and auriva-backend repositories, with the component microservices in separate repositories under the project's GitHub organisation (R26-IT-006).

5.2 Development and integration process

Work proceeded in short iterations within each component, with a common interface-first discipline for everything the components share. Each member developed on a personal branch and merged into the shared repositories, where the frontend integration branch is dev/merge and the backend integration branch is dev/integration. The shared contracts (authentication, student model, workspace layout, sensory preferences) were agreed before the modules were merged, and the components were then integrated without changing their learning logic.

Quality practices were applied consistently. Decision logic was written as pure functions over explicit state with named constants, so mastery, difficulty and relearn rules could be unit tested without a database. The dialogue component used about 45 specified tasks, each with written acceptance criteria, and a numbered decision log recording the reasons for every significant choice [8]. Defects found by testing, audits and pilot use were tracked to resolution, and regression tests were added when defects were fixed. Research code was kept apart from serving code so that experiments could not affect a live classroom. Each thesis declares its use of generative AI and other tools in its own Appendix B, in line with the module policy.

5.3 Deployment and operational configuration

The integrated system is deployed as a composed container stack on a single virtual machine. A reverse proxy terminates TLS and is the only public service; the backend API and the three Python services are reachable only on the internal network. Container images are built and published automatically from the main branch, and a single service can be rolled back independently by pinning its image tag. External managed services comprise the PostgreSQL database, blob storage, the Neo4j database and the language model API, and the application time zone is fixed to Asia/Colombo so that daily practice limits and session dates use local time [7].

Operational controls follow the same philosophy as the design. The graph inference path is governed by an environment flag that disables it and returns the system to heuristic behaviour. The dialogue machine-learning path is switched on by a flag and was enabled only for supervised sessions. Pronunciation configuration (database

connection, secrets, rate limits, worker and model paths) is held in environment variables, structured logs and read-only quality reports monitor scoring, and a retention script deletes stored audio after 60 days while keeping score records [9]. Before each testing day the dialogue team confirms the device build, confirms the backend and phoneme service are reachable, completes one Phase 2 attempt and one Phase 3 scenario as a smoke test, and checks that scores and latencies were stored [8].

Table 5.2 records the deployment and test configuration used for each component.

Table 5.2: Deployment and test configurations by component

Component	Devices used	Backend and services	Notes
1. Concept	Android tablets in the partner school	Composed container stack, managed PostgreSQL and Neo4j, language model API	Production data from ordinary teacher-run use
2. Handwriting	Samsung Galaxy Tab S7+ with stylus	Backend over local-network HTTP; PostgreSQL; FastAPI service	Research build: HTTPS, managed secrets and service authentication are production requirements
3. Pronunciation	Physical Android tablet and emulator; Expo Go for pilot	Node.js service on the project server; Azure PostgreSQL; speech models run locally (GOP worker about 1.3 GB)	No recordings leave the server
4. Dialogue	Amazon Fire HD 10 (11th generation) and a Samsung tablet	Azure-hosted backend and microservices; frozen testing build auriva-db5c387	Changed during testing only to deploy single-voice Level 2 audio

Backup and restoration. PostgreSQL is the system of record and is a managed service. The Graph Knowledge Base is a projection that can be rebuilt from the interaction event log, which is how the research dataset builder works. A formal backup and restore procedure is listed as a production requirement by the handwriting thesis and has not been documented as complete for the platform; it is carried into the future work in Section 8.3.

5.4 Evidence of team execution

Evidence of execution comes from version-control history on the member branches and shared repositories, the numbered decision log and task acceptance records, dated audits and test runs, frozen and identified deployed builds, and the data-collection windows summarised in Table 5.3. Raw activity logs are not reproduced.

Table 5.3: Dated evidence of team execution

Date (2026)	Evidence	Component
2 May	Concept sessions begin in the partner school (27 active days to 9 Oct)	1
7 May	Pronunciation attempts begin to be recorded (249 scored attempts to 9 Oct)	3
25 Jul	Phoneme service confirmed running in pilot sessions after fallback was found to be silent	4
6 Aug	Counter and non-verbal threshold defects fixed after label audit	4
16 Aug	Structured code audit	4
19 Aug	Audit of trajectory model calibration; faulty calibration removed	4
30 Aug	Full automated test run: 7,609 tests passed	2
6 Oct	Audit of training labels	4
7 Oct	Supervised dialogue field study begins on frozen build auriva-db5c387	4
11 Oct	Concept graph snapshot taken: 570 nodes and 2,663 relationships	1
Third week of Oct	Planned end of dialogue field study	4

The audits in particular show that the team verified its own work rather than relying on headline accuracy. They exposed a silent fallback in early pilot sessions, a calibration fitted on nine rows that made the model predict one class for every input, corrupted early labels, and a confound in the concept ranking evaluation. Each was reported openly in the relevant thesis.

6 INTEGRATED TESTING EVALUATION AND VALIDATION

6.1 Evaluation framework and success criteria

The group asked four evaluation questions (EQ). **EQ1:** does the integrated platform operate end to end in real settings with shared authentication, data and reporting? **EQ2:** do the adaptive mechanisms behave as specified and show evidence of learning or of agreement with teacher judgement? **EQ3:** do the safeguards make the intelligence layers non-disruptive to the child? **EQ4:** do teachers find the platform usable and useful? Each component answered EQ2 with the design in Table 6.1, chosen to fit its data. The following acceptance criteria are used to judge the evidence: every module operational in the integrated build; every specified fallback verified; each learned or computed method outperforming the baseline it was proposed against; and, for usability, a mean System Usability Scale score above the commonly cited average of 68, the benchmark named in the dialogue evaluation plan [8].

Table 6.1: Evaluation designs by component

Component	Design and participants	Primary metrics and analysis
1. Concept	Observational deployment in routine teacher-run use; 21 registered children, 19 with tier-level data; 441 attempts across 42 of 96 concepts over 27 active days	Tier pass rates; within-concept gain (sign test, 95% confidence interval (CI)); similarity of confused versus non-confused pairs (Mann-Whitney U, Cohen's d); GAT versus five baselines under leave-one-student-out validation
2. Handwriting	Controlled collection from 7 learners (42 shape trajectories, 420 letter-attempt records); 140 third-attempt observations; teacher usability tasks	Feature reproducibility; silhouette; leave-one-participant-out adjusted Rand index (ARI); automated test status; SUS (pending)
3. Pronunciation	249 scored in-app attempts, 32 teacher-reviewed (chosen by uncertainty); 167 child clips from 9 classroom sessions	Pearson correlation, mean absolute error and pass/fail agreement with teachers; Layer 1 and Layer 3 analyses
4. Dialogue	87 real first-session records from 26 accounts; supervised field study with 10 children and about 6 teachers at 3 schools	Balanced accuracy under leave-one-child-out validation; observer agreement, echolalia precision and recall, SUS (pending)

6.2 Integration and end to end testing

Table 6.2 summarises the testing evidence across the platform. It reports what was tested at each level and, where relevant, what was not.

Table 6.2: Integration and system testing evidence

Test type	Evidence	Result
Unit and regression	Concept: unit tests of tier engine, validation and utilities. Handwriting: 4,150 mobile tests (142 suites), 3,209 backend tests (139 suites), 250 machine learning (ML) tests with 33 subtests. Pronunciation: 87 backend and 16 frontend tests. Dialogue: 114 Jest tests in four suites	All passed; Handwriting 7,609 of 7,609 (30 Aug 2026)
Interface and integration	Concept routes tested against a test database with guards active. Pronunciation: login, student selection, full word session, scoring, saving, history replay and review submission on a physical tablet and an emulator. Handwriting: manual end-to-end checks across app, backend, database and ML service. Dialogue: component ran in the integrated testing build	Passed; Handwriting manual functional cases defined (18) but results still being confirmed; Dialogue integration observations pending
Scenario and system	Concept: 14 adaptive-behaviour scenarios including fallback cases. Dialogue: sweep of 6,482 generated inputs to confirm that adding explanations changed no predicted label	14 of 14 passed; 0 labels changed
Security and authorisation	Role and student-ownership negative tests (Concept); authorisation, ownership and token tests (Handwriting); review against the OWASP Top 10 [19]	Passed in tests; research-build weaknesses recorded (Section 7.4)
Performance	Pronunciation timing on a development machine	About 3.0 s per attempt with models loaded; about 12.8 s cold start; not measured for other components
Defects found	Five in pronunciation, all fixed; ten in dialogue, eight resolved and two open; regression tests added for handwriting defects; concept option-set logging gap	Resolved except as noted

Two features of this evidence matter for interpretation. First, much of the automated testing is unit-level with mocked dependencies, so database-backed integration and device behaviour were verified largely by manual end-to-end use, and no automated real-device test framework exists. Second, the three behaviours that matter most for a platform whose intelligence is optional, namely graceful fallback, invisibility of failures to the child and teacher-only access to a child's data, were each tested explicitly in the concept component and designed for in the others.

6.3 Overall results

Table 6.3 synthesises the principal results. Each is the result of the component's own experiments, and the individual theses give the full methods, tables and caveats.

Table 6.3: Principal results by component

Component	Principal results	Reading
-----------	-------------------	---------

Component	Principal results	Reading
1. Concept	Tier 1 mastery in 80.3% and Tier 2 in 98.3% of concept records; 98.95% of concepts passed at Tier 1 and later attempted passed Tier 2. Within-concept gain of 0.358 score points (95% CI 0.235 to 0.480; 47 sequences improved, 8 declined; sign test $p < 0.001$). Visual similarity higher for confused pairs ($z = 3.83$, $p = 0.00013$, $d = 0.62$); phonetic similarity not ($p = 0.178$). Similarity covers 96.3% of untouched concepts. GAT did not beat the production heuristic or any baseline	Learning signal and a useful computed feature; negative result for learned graph attention at this data scale
2. Handwriting	Deployed feature values reproduced exactly by the research pipeline; 7,609 of 7,609 tests passed. Dataset A: the same two-cluster partition from K-Means, Ward and Gaussian mixture across 100 initialisations (silhouette 0.412; leave-one-participant-out mean ARI 0.643). Dataset B: weak-to-moderate structure (silhouette 0.2563; mean ARI 0.8413)	Transparent scoring is reproducible; clusters are descriptive only; score weights still need calibration
3. Pronunciation	System and teacher scores on 32 reviewed attempts: $r = 0.73$ (95% CI 0.51 to 0.86), mean absolute error 7.9 points, 71.9% within 10 points, 87.5% pass/fail agreement. Word-level MFCC-DTW alone: $r = 0.19$ and 0 of 151 attempts accepted. 26.1% of attempts flagged for review. Recalibration lowered internal adaptive-score error from 11.2 to 8.8 but did not improve the teacher-visible score. 167 child clips extracted	Phoneme-level scoring agrees substantially with teachers; adaptation model awaits labels
4. Dialogue	On 87 real first-session records (leave-one-child-out) both Random Forests beat the Tier 1 formula on verbal attempts (balanced accuracy 0.695 real-trained and 0.723 synthetic-trained against 0.498). 114 of 114 tests passed. Speech-assessment, echolalia and usability results pending the field study	Pace prediction viable for verbal attempts; real-trained model not yet better than synthetic-trained

Three findings cut across the components. First, explicit, mastery-gated rules deployed with real children worked as designed and were accepted by teachers, with learning signals visible in the concept data. Second, at classroom data scales, computed and interpretable features outperformed or matched learned models: a similarity feature that needs no training data carried real signal while a graph attention network did not, phoneme-level scoring built on pretrained models agreed with teachers where template matching did not, and handwriting clusters were useful only as descriptions. Third, careful auditing mattered more than headline accuracy. Audits exposed an evaluation confound, a degenerate calibration and corrupted labels, and each would have produced misleading claims if left unexamined.

6.4 Stakeholder and user validation

Table 6.4 records who evaluated each component, how evidence was collected, the main findings and the status. Evaluation of the complete solution by a single group of stakeholders was not conducted; validation was component-level, with the shared Teacher Workspace and shared sensory layer being what teachers experienced across modules.

Table 6.4: Stakeholder and user validation

Component	Evaluators and method	Findings and changes	Status
1. Concept	Special education teachers using the Teacher and Student Workspaces; task-based assessment and guided discussion; supervised, observational child sessions	Three themes, each producing a design change: handover flag valued; summaries must describe, not interpret; more categories requested (a family category was added). No quantified instrument	Qualitative only; SUS not administered
2. Handwriting	Teachers; task-based usability evaluation with SUS, clarity and usefulness ratings, open feedback	Evaluation carried out; results being compiled	Pending
3. Pronunciation	Special education teachers reviewed 32 attempts through the review queue	Teacher scores are the reference standard for agreement results (87.5% pass/fail agreement). A formal teacher usability evaluation was planned but is not reported	Review complete; usability not reported
4. Dialogue	About 6 teachers and 10 children at 3 schools; 14-task usability survey, SUS, feedback questionnaire, observer sheets	Earlier test rounds with children led to removing the conversation level. Field results pending; study due to end in the third week of October 2026	Pending

The main remaining concern is therefore that no component yet has a measured usability result. This is the most important gap in the evidence for EQ4 and is the first item in the future work.

6.5 Achievement of common objectives

Table 6.5 sets the evidence against each group objective.

Table 6.5: Achievement of group objectives

Objective	Status	Justification
GO1 Four components in one deployed application	Achieved	All four modules run in one build with shared login, student profile, database and workspace; integration changed only frontend presentation for the dialogue module. Verification of the integrated whole was largely manual
GO2 Mastery-based adaptive progression by explicit rules	Achieved	Rule engines implemented and tested in all components: two-of-three tier gate and relearn; Attempt-3 mastery at 70; teacher-gated review; phase-based mastery with difficulty gate
GO3 Interpretable teacher evidence in one workspace	Partly achieved	Four module tabs with consistent layout and explanations are implemented; teacher validation is qualitative for one component and pending for others
GO4 Intelligence can never harm the child's session	Achieved, one case unevaluated	Timeouts, fallbacks, kill switches and "model never gates progress" rules verified or designed in all four; the handwriting Teacher Report with the ML service unavailable was not evaluated
GO5 Evaluation with autistic learners and teachers under ethical oversight	Partly achieved	Classroom deployment (21 children), controlled collection (7 learners), teacher review (32 attempts) completed; dialogue field study and usability evaluations pending; consent procedures differed by component
GO6 Candid evidence and an adoption pathway	Partly achieved	Negative and pending results reported openly; a licensing model is proposed but prices and costs are not validated

6.6 Integrated limitations and threats

The following limitations apply to the complete solution and should temper every claim in this report.

- **Small, local samples.** Evaluations used between 7 and 21 children, mostly at one or a few schools, with heterogeneous autistic learners. Results describe those settings and do not generalise to autistic children broadly.
- **Observational designs.** The concept data cannot separate the effect of adaptive mechanisms from repeated exposure, teacher effects or maturation, and the gain measure uses re-attempted sequences, which regress to the mean. The learning claim is an association, not a causal effect.
- **Reviewer and label bias.** Pronunciation reviews were chosen by uncertainty rather than at random, 25 of the 32 came from students with learning disabilities rather than ASD, and teachers could see system scores. The dialogue pace model's real dataset contains test accounts and it is not yet better than a synthetic-trained model.
- **Dependencies between components.** All four modules share the backend, database and authentication, so a shared fault would affect all. The planned use of concept, speaking and writing mastery in dialogue Level 2 was not evaluated.
- **Unequal maturity of evidence.** Usability, dialogue field results, handwriting manual test results and the pronunciation adaptation model are pending.
- **Inconsistent build practices.** Some research builds used local HTTP, development secrets and unauthenticated internal calls, while the production deployment is described as TLS-fronted with internal-only services. The integrated platform has not been security-tested as a whole.
- **Coverage and generalisability.** Fruits account for 64.2 per cent of concept attempts, three concept categories have no data, and only English letters and Grade 1 to 3 content are covered.
- **Unresolved integration risks.** Consent coverage for secondary use of data, a documented retention and deletion policy, backup procedures and cross-module reporting under load remain open.

7 COMMERCIALISATION SUSTAINABILITY AND RESPONSIBLE INNOVATION

7.1 Value proposition and beneficiaries

Auriva's value proposition is that a classroom teacher with no specialist technology background can deliver personalised, curriculum-aligned English instruction to autistic children and can see, for four skill areas, how each child is progressing and why the system made each decision. It approximates support that would otherwise need one-to-one specialist time, which the sector lacks. Table 7.1 distinguishes the buyer, the users and the beneficiaries.

Table 7.1: Buyers, users and beneficiaries

Stakeholder	Role	Value received
Principals and managers of special education schools, units and centres	Buyer, approver and administrator	One system for all modules, children and teachers; documented evidence of progress; predictable per-institution cost
Special education teachers	Daily user	Ready-to-use curriculum content; mastery evidence; explanations; review queue, handover flag and worksheets that keep professional judgement in control
Autistic children	Non-paying beneficiary	Structured practice with graduated support, gentle feedback, sensory personalisation and a non-verbal pathway
Parents and guardians	Secondary beneficiary	Worksheets for guided home practice; indirect gains as communication and independence improve
Society and education system	Indirect beneficiary	Inclusive access to English instruction for a group frequently excluded from mainstream support

Stakeholder evidence supports specific parts of this proposition: teachers valued the bounded-failure handover and asked for descriptive rather than interpretive summaries [7], and teacher judgement is the reference standard for pronunciation scoring [9]. It does not yet include measured teacher satisfaction, because usability results are pending.

7.2 Market alternatives and differentiation

The initial market is Sri Lanka's special education sector, which the group identified as about 27 special schools and 704 special education units or institutes [7], [8]. Official figures show 7,574 students in government special education units and 1,956

students in 27 government-assisted special schools in 2024 [5]. The market is small and concentrated, which favours institutional licensing over consumer sales.

Competing and substitute solutions fall into four groups: autism-focused learning applications such as Otsimo, Speech Blubs, Mita and AutoSpark; augmentative communication tools such as Proloquo2Go and Avaz; handwriting applications such as Writing Wizard and LetterSchool; and local research platforms that target general skills rather than curriculum-aligned English [13]. None of the audited products follows the Sri Lankan syllabus. Table 7.2 states the differentiation that the evidence supports, and where Auriva does not yet differ.

Table 7.2: Differentiation of the integrated product

Capability	Typical alternatives	Auriva
Sri Lankan curriculum alignment	Global content in the audited applications	Content traced to the National Institute of Education Grade 1 and 2 materials
Teacher-mediated model	Parent-child model	Teacher runs the session; handover flag; review queue
Mastery tiers and graduated remediation	Sequential activities; distractor removal only or identical repetition	Tier gate and a three-level relearn pathway that varies the stimulus
Assessment suited to autistic and Sinhala-influenced speech	General speech recognition; fuzzy string matching	Phoneme-level scoring with partial credit for transfer errors; echolalia handling
Process-level handwriting evidence	Tracing completion or final form	Trajectory measures and exact mastery evidence
Explanations for the teacher	Raw numbers behind a subscription	Plain-language summaries, mastery evidence and model explanations
Offline operation	Continuous connectivity required	Not yet offered; listed as future work

The honest competitive position is a local fit and transparency, not artificial-intelligence personalisation. The concept thesis notes that describing the graph layer as AI-driven personalisation would overstate what its results support; the defensible claim is an adaptation grounded in an empirically validated model of which concepts children confuse [7]. The main competitive risks are international applications adding local languages and the limited budgets of government schools.

7.3 Business model and feasibility

The platform is offered as software as a service, licensed per institution rather than per student so that costs are predictable under fixed allocations. A freemium trial lets schools gather the trial evidence that public procurement usually requires. Tier I covers up to 30 students, Tier II covers 30 to 120 students, and the concept thesis

proposes a third tier for 120 or more students or multi-campus organisations. Prices have not been set or validated with customers and no return-on-investment estimate is claimed. For government units, adoption could be supported through education authorities, non-governmental organisations and inclusive-education programmes, and the Ministry of Education is a secondary route should pilot-scale outcomes be demonstrated. Short, teacher-led onboarding is intended to keep the cost of introducing the platform low.

Table 7.3 lists the main cost areas that the business model must cover.

Table 7.3: Main cost areas

Cost area	Description
Cloud hosting	Backend, database, Python services and graph database; grows with active institutions
External services	Speech-to-text fees per minute of child speech and text-to-speech for personalised sentences (reducible by caching); language model API usage
Storage	Media assets, raw trajectories, worksheet images and, for pronunciation, recordings under a retention limit
Devices	Tablets for schools without suitable devices
Maintenance and security	Bug fixes, dependency updates, monitoring, backups, recalibration and retraining as data grow
Onboarding and content	Teacher training, user guides, support, and updates when the curriculum changes

Technical and operational feasibility is supported by the deployed classroom use described in Chapter 6. Running cost is moderated by design choices: the concept graph model was not promoted, so the production container carries no machine-learning serving stack, and the pronunciation models run on the project server on an ordinary machine without a graphics processor or paid speech services [7], [9]. The concept thesis reports development costs of approximately Rs. 34,500 for the six-month prototype period. With consent, subscribing schools' data could be used to retrain models, and word sets and scenarios are stored as data so they can be updated without code changes [8].

7.4 Adoption intellectual property and compliance

Adoption barriers. Expected barriers, based on design and context rather than measured findings, are limited availability of suitable tablets and reliable connectivity, tight budgets in government units, teacher time for training and for reviewing reports, English-only letter templates, and parental and school concern

about collecting children's data. They can be reduced through the free trial, student-based tiers, short onboarding, planned Sinhala and Tamil extensions and a clear data-protection position.

Intellectual property. The main assets are the source code, the learning-flow and mastery logic, the scoring and sequencing rules, the Sinhala error taxonomy and echolalia rule, trained model artefacts and original educational content. The product would be delivered as a hosted service rather than by distributing source code. Third-party services are used under their own licences and kept separate from the product's logic. Teaching media produced with generative tools was reviewed before use. Ownership and commercialisation arrangements would follow the policy of the Sri Lanka Institute of Information Technology.

Data protection, ethics and safety. Commercial use would require compliance with Sri Lanka's Personal Data Protection Act, No. 9 of 2022 [20], a lawful basis for processing children's data, data-processing agreements with institutions, defined retention periods and access controls. Current practices differ and should be harmonised. Research data in all components are pseudonymised and no video of children is recorded. The concept module sends only anonymised statistics to a language model API. The dialogue component sends spoken responses to Google Speech-to-Text with data logging disabled, so audio is processed in memory and only the transcript is kept. The pronunciation component keeps recordings in PostgreSQL, restricted to authorised teachers, with a 60-day deletion script, and processes them on its own server. Consent also differed: written parental or guardian consent and teacher consent were obtained for the concept study, parental consent for pronunciation, and school consent on behalf of guardians for handwriting data collection, which that thesis acknowledges as a limitation, while the dialogue study operates under Ministry and school permission with teachers present at every session. Ethical clearance and school approvals were obtained for the handwriting study. For all components, secondary use of data to train models needs explicit confirmed consent, and a policy for models trained on a withdrawn child's data is required. The platform must be presented as an educational support tool and not as a diagnostic or medical device, and teachers remain the final decision-makers.

Security prerequisites. Before wider deployment the research builds must close the gaps recorded in the individual theses: HTTPS and TLS everywhere, managed and rotated secrets, restricted cross-origin settings, authenticated calls between the backend and each internal service, and a backup and restore procedure.

7.5 Sustainable Development Goal contribution

The platform's plausible contribution to SDG 4 is greatest for targets 4.5 and 4.a [10]. It provides disability-sensitive, curriculum-aligned English practice on general-purpose tablets, reduces dependence on scarce specialist time and gives teachers evidence that supports individualised instruction. It contributes indirectly to SDG 10 target 10.2 by supporting the inclusion of learners with additional needs. These remain plausible contributions: the evidence shows system behaviour and short-term learning signals in small samples, not changes in educational attainment or access.

Possible negative effects include widening a digital divide where schools lack tablets or connectivity, over-reliance on screen-based practice, privacy risks to a vulnerable group, and over-trust in automated scores. The design measures against these are the use of general-purpose tablets, paper worksheets for home practice, teacher final authority, teacher review of uncertain scores, non-diagnostic positioning and data-minimising practices. To sustain the benefits the platform needs sustained funding or grant-supported pricing for low-resource schools, teacher onboarding, regular content updates and Sinhala and Tamil localisation, and a longer study to show whether practice transfers to everyday classroom use.

8 CONCLUSION AND FUTURE DIRECTIONS

8.1 Overall conclusion

The common problem was the lack of a curriculum-aligned, adaptive, teacher-controlled tool that teaches autistic children in Sri Lankan special education the foundational skills of English and gives teachers interpretable evidence. The integrated evidence shows that a solution to this problem is feasible and, in important respects, achieved, but that the evidence for its effect on learning remains preliminary.

What the evidence supports is the following. A single application with one login, one student record, one database and one Teacher Workspace now delivers four adaptive English modules. It has been deployed and used in live classrooms, where the concept module showed mastery and within-concept gains with 19 children, pronunciation scores agreed closely with teacher judgement, handwriting measures were reproducible, and the dialogue pace model beat a formula baseline on verbal attempts. Design safeguards, namely rule-based decisions, timeouts, fallbacks and teacher authority, were verified so that optional intelligence cannot interrupt a child's session. What the evidence does not yet support is a claim of improved long-term learning, teacher usability at a measured level, or readiness for unsupervised production use. These gaps are stated, not hidden.

8.2 Principal integrated contribution

Beyond the four components, the team has produced a reusable pattern and a body of lessons. The pattern is a teacher-controlled, rule-governed platform in which machine learning is advisory, bounded and switchable, which allowed four research strands to run in live classrooms without risk to children. The lessons are transferable to similar work with small special education populations.

- **Computed, interpretable features can beat learned models at small scale.** A similarity feature and a pretrained phoneme scorer produced reliable signal where a graph attention network and a word-level template matcher did not.

- **Evaluation must mirror production and be audited.** Auditing exposed a self-labelling confound, a degenerate calibration, corrupted labels and an evaluation script that did not match production encoding.
- **Child-safe design is a system property.** No scores or failures for the child, errorless and non-verbal pathways, bounded failure with teacher handover, and sensory preferences shared across modules proved compatible with adaptive instruction.
- **Local data and methods.** The work yields curriculum-aligned content, datasets from Sri Lankan classrooms, a method for building labelled child-speech data from ordinary prompt-and-repeat sessions, and Sinhala transfer-aware assessment rules.

8.3 Future directions

The following work is prioritised from the most to the least urgent.

1. **Complete the pending evidence.** Finish the dialogue field study, compile the handwriting teacher usability results, run a quantified usability study for the concept and pronunciation components, and report SUS scores and the manual test results.
2. **Harden for production.** Enforce HTTPS and TLS, managed secrets, restricted cross-origin settings and authenticated internal services across all builds; add database-backed integration tests and automated device end-to-end tests; document backup and restore.
3. **Settle consent, retention and deletion.** Confirm consent for secondary use of data, harmonise retention and logging practices across components, and define the procedure for withdrawn children's data and models.
4. **Fix known evaluation limits.** Log the options presented for every scored concept attempt and introduce randomised distractor exposure before repeating the ranking evaluation; collect blind and random-sample teacher reviews for pronunciation; remove test accounts and retrain the dialogue model on real data.

5. **Strengthen the models only when data allow.** Revisit graph models only when stated data gates are met, train the child-speech adaptation model once labelling is complete, and calibrate handwriting scores and thresholds against teacher or occupational-therapist ratings.
6. **Run a longer, controlled study.** Measure learning gain and transfer to classroom use over time with more children, more schools and a comparison condition.
7. **Evaluate cross-component use.** Test whether concept, speaking and writing mastery can usefully shape dialogue activities.
8. **Extend and scale.** Add Sinhala and Tamil support and templates, offline operation, additional curriculum categories and a Sinhala-English instruction toggle, and pilot in more schools with the Ministry of Education.

REFERENCES

- [1] World Health Organization, "Autism Spectrum Disorder," 17 Sep. 2025. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders> (accessed 11 Mar. 2026).
- [2] GBD 2021 Autism Spectrum Collaborators, "The global epidemiology and health burden of the autism spectrum: Findings from the Global Burden of Disease Study 2021," *Lancet Psychiatry*, vol. 12, no. 2, pp. 111-121, 2025, doi: 10.1016/S2215-0366(24)00363-8.
- [3] H. Perera, K. Wijewardena, and R. Aluthwelage, "Screening of 18-24-month-old children for autism in a semi-urban community in Sri Lanka," *J. Trop. Pediatr.*, vol. 55, no. 6, pp. 402-405, 2009.
- [4] "Increase in autism among children: Over 9,000 identified in Sri Lanka," *Newswire.lk*, Apr. 10, 2025. [Online].
- [5] Statistics Branch, Ministry of Education, Sri Lanka, "Annual School Census of Sri Lanka - Summary Report 2024," Jan. 2025. [Online]. Available: https://www.statistics.gov.lk/Resource/en/Education/School_Census/Annual_School_Census_SummaryReport_2024.pdf
- [6] National Institute of Education, *English Language Teacher's Guide: Grade 1 and Grade 2*. Maharagama, Sri Lanka: National Institute of Education.
- [7] H. H. Jayasundera, "Auriva: Adaptive English language learning platform for autistic children in Sri Lanka - GNN-enhanced adaptive concept learning module," individual thesis (IT22084590, group R26-IT-006), B.Sc. (Hons) Inf. Technol., Sri Lanka Inst. Inf. Technol., Sri Lanka, Oct. 2026.
- [8] M. C. Gunawardena, "Auriva: Adaptive English language learning platform for autistic children in Sri Lanka - Adaptive dialogue learning and progression tracker," individual thesis (IT22166838, group R26-IT-006), B.Sc. (Hons) Inf. Technol., Sri Lanka Inst. Inf. Technol., Sri Lanka, Oct. 2026.
- [9] N. G. Samaranayake, "Adaptive English language learning platform for autistic children in Sri Lanka - Phoneme-aware adaptive pronunciation support component," individual thesis (IT22266378, group R26-IT-006), B.Sc. (Hons) Inf. Technol., Sri Lanka Inst. Inf. Technol., Sri Lanka, Oct. 2026.
- [10] United Nations General Assembly, "Transforming our world: The 2030 Agenda for Sustainable Development," A/RES/70/1, Oct. 2015.
- [11] R. S. Liluksha, "Auriva: Adaptive learning platform for English language development in autistic children in Sri Lanka - Motor-informed adaptive handwriting learning for children with autism spectrum disorder," individual

thesis, draft v5 (IT22216700, group R26-IT-006), B.Sc. (Hons) Inf. Technol., Sri Lanka Inst. Inf. Technol., Sri Lanka, Oct. 2026.

- [12] A. Kushki, T. Chau, and E. Anagnostou, "Handwriting difficulties in children with autism spectrum disorders: A scoping review," *J. Autism Dev. Disord.*, vol. 41, no. 12, pp. 1706-1716, 2011, doi: 10.1007/s10803-011-1206-0.
- [13] K. G. Ranaweera, J. Praneetha, W. P. S. De Silva, K. R. B. N. Kempitiya, I. Gamage, and T. Thilakarthna, "Personalized and interactive learning platform for students with autism spectrum disorder," in *Proc. 5th Int. Conf. Adv. Comput. (ICAC)*, 2023, pp. 828-833, doi: 10.1109/ICAC60630.2023.10417404.
- [14] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph Attention Networks," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2018.
- [15] S. M. Witt and S. J. Young, "Phone-level pronunciation scoring and assessment for interactive language learning," *Speech Commun.*, vol. 30, no. 2-3, pp. 95-108, 2000.
- [16] A. Baevski, H. Zhou, A. Mohamed, and M. Auli, "wav2vec 2.0: A framework for self-supervised learning of speech representations," in *Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 12449-12460.
- [17] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," in *Proc. 40th Int. Conf. Mach. Learn. (ICML)*, 2023, pp. 28492-28518.
- [18] M. Jones, J. Bradley, and N. Sakimura, "JSON Web Token (JWT)," *IETF RFC 7519*, May 2015.
- [19] OWASP Foundation, "OWASP Top 10:2021," 2021. [Online]. Available: <https://owasp.org/Top10/> (accessed Oct. 7, 2026).
- [20] Parliament of the Democratic Socialist Republic of Sri Lanka, *Personal Data Protection Act, No. 9 of 2022*, 2022.

APPENDICES

The appendices contain the material needed to verify the integrated solution: the interface specification (Appendix A) and the agreed record of team contributions (Appendix B). Both are cited from Chapter 4.

Appendix A Integrated interface specification

Table A.1 lists the interfaces of the integrated platform as documented in the four individual theses. Endpoint paths and full schemas are held in the respective repositories.

Table A.1: Interface catalogue

ID	Interface	Payload and flow	Timeout, fallback and owner
I1	Tablet to backend (all modules)	REST/JSON with JWT bearer token; role guard and student-ownership check on every route	Shared; rate limiting and validation apply
I2	Concept attempt event	Student, category, concept, tier, event type and a flexible payload (response, selected distractor when wrong, difficulty, elapsed time, options presented for activity events), written to the interaction log	Synchronous to PostgreSQL; owner: Concept
I3	Concept to GKB service (write)	Scored block summaries projected to Neo4j; confusion edge weights incremented; append-only attempt nodes	Fire-and-forget; failure logged and ignored; graph rebuildable from the log
I4	Concept to GKB service (read)	Confusion-informed distractor candidates	Hard 400 ms timeout; fallback to interaction log, then curriculum sequence; environment flag disables inference
I5	Concept to language model API	Anonymised session statistics and graph-derived analytics, on demand	Isolated from the child session; dashboard renders without the summary on failure
I6	Handwriting attempt submission	Learner, activity, letter, case, attempt number, raw trajectory and client-computed letter features	Authoritative Motor Score computed in the backend; idempotent finalisation
I7	Handwriting Teacher Report, Writing Check and worksheets	Assessment summary, motor performance, mastery evidence, word progress; start, progress and complete requests; worksheet plan and image upload	JWT with teacher ownership enforced; owner: Handwriting
I8	Handwriting backend to motor-pattern service	Three letter features in; cluster identifier and descriptive summary out	Internal only; no operational decision depends on it
I9	Pronunciation scoring	Recording (mono AAC, 16 kHz) with hesitation time; server quality checks; Layer 1, Layer 2 and Layer 3 scores; result stored with result_id	Endpoint rate limit 30 per 5 minutes per teacher; low confidence goes to review queue
I10	Pronunciation teacher review	Review score submitted for an attempt; used for calibration only when validated	Owner: Pronunciation
I11	Dialogue Phase 2 attempt	Recording with avatar-audio-end and recording-start times; WAV conversion; Google Speech-to-Text transcript with word timings; accepted-form check; /score-phoneme; latency and echolalia flag stored	2 s limit on phoneme service; fuzzy-matching fallback

ID	Interface	Payload and flow	Timeout, fallback and owner
I12	Dialogue /predict-trajectory and /explain-trajectory	13 first-session features per word in; pace label, vote-share confidence and explanation out	Confidence gate (0.60), Tier 1 formula fallback, feature flag; model cannot change mastery
I13	Dialogue Level 2 session build	Child profile for the topic to target sentences, word tiles, gap options and Azure Text-to-Speech audio	Owner: Dialogue; audio can be cached per profile

Table A.2 lists the entities shared across components.

Table A.2: Shared entities

Entity	Key fields	Used by
Student	Identifier, assigned teacher, population group, avatar theme, reduced-stimulation setting, stored sensory preferences	All modules
Teacher and role	Identifier, role (teacher, principal), student assignment	All modules (authentication and ownership checks)
Concept records	Interaction log, tier progress, relearn state, report artefacts; graph projection in Neo4j	Concept
Handwriting records	Attempts, trajectories, features, scores, mastery and word progress, worksheets, evidence	Handwriting
Pronunciation records	pronunciation_session_results with scores, evidence, teacher reviews and audio	Pronunciation
Dialogue records	Word progress, phase records, sessions, echolalia rate, Level 2 profile data	Dialogue

Appendix B Team contribution declaration

This record summarises each member's individual contribution as reported in the individual theses, and the shared contributions made by the group. The shared platform (login, student profiles, integration of the four components and deployment) was provided by the group, with integration and deployment carried out by the group leader. All four members are asked to confirm the record below.

Table B.1 records the agreed contribution of each member for confirmation.

Table B.1: Team contribution declaration

Member	Individual contribution	Shared contribution	Confirmed (signature and date)
H. H. Jayasundera (IT22084590)	Concept Learning Module: three-tier engine, relearn pathway, Graph Knowledge Base service, similarity layer, Graph Attention Network evaluation, teacher analytics	Shared sensory, authentication and Teacher Workspace conventions; coordination of the common reporting layout	

Member	Individual contribution	Shared contribution	Confirmed (signature and date)
R. S. Liluksha (IT22216700)	Handwriting component: trajectory capture, DTW and motor features, Motor Score, practice workflow, Writing Check, worksheets, motor-pattern service, data-collection mode	Shared backend, database and authentication; integration of handwriting routes and screens	
N. G. Samaranayake (IT22266378)	Pronunciation component: recording, three-layer scoring, review queue, adaptive next word, child-speech adaptation pipeline	Shared backend, student records and design system; integration of pronunciation routes and screens	
M. C. Gunawardena (IT22166838)	Dialogue component: Level 1 and Level 2 flows, microservice, assessment, echolalia and trajectory models, reports, audits, field-study instruments	Shared backend and database; frontend restyling and Level 2 report extension during integration	