



INVESTIGATE PHASE · CONSOLIDATED REPORT

Knowledge Revolution

AI virtual-human mentors and digital twins for every learner

“How can we make the best knowledge available to every learner, wherever and whenever they are?”

CHALLENGE

Allow every student, wherever they are, to ask questions and learn face to face from the best teachers and minds, in their own language.

01

Context & Problem

02

Data Study

03

Techniques, Tools & Limits

04

Guiding Activities

05

Legal, Ethical & Market

06

Synthesis & Objective Tree

TEAM · FBAI'S — FUTURE BUILDERS OF AI

Med Rayen Ayat · Youssef Jouini · Malek Tirellil · Yassmine Nouisser · Maryem Ouichka · Islem Tellili

SUPERVISION

Mr. Safouane Jebali · Mrs. Wided Askri
4th year · Artificial Intelligence

REFERENCE PLATFORM

BEAI™ Studio — Agent & Digital-Twin Creation
Partner: Fawzi Ben Messaoud

VERSION

v3 · 6 October 2026

Contents

OVERVIEW

About this report

Engage framework · Team and parts · Completeness dashboard · How to read

PART 1 · ALL MEMBERS

Context & Problem Structuring

1.1 Context · 1.2 Problem · 1.3 Causes · 1.4 Consequences · 1.5 Stakeholders · 1.6 Limitations · 1.7 Context → Problem → Impact schema · 1.8 Guiding Questions

PART 2 · MARYEM OUICHKA

Data Study

2.1 Types of data · 2.2–2.3 Availability and what we have · 2.4 What we do not have · 2.5 Data gaps · 2.6 Legal constraints · What can be collected · Sources

PART 3 · MED RAYEN AYAT & ISLEM TELLILI

Techniques, Tools & Limits

3A

Evidence review: 9 technique cards · 26-tool comparison · Needs × techniques matrix · What is missing · Caveats · 42 references —

3B

Hands-on testing: 7 tools tested · Technique notes · Feature coverage · Observed limits · Gaps

PART 4 · MALEK TIRELLIL

Guiding Activities & Resources

Introduction · Methodology · Investigation framework · 24 guiding activities · Guiding resources · Coverage matrix · Schedule · Outputs · Points of attention · References

PART 5 · YOUSSEF JOUINI

Legal, Ethical & Market Study

Scope and method · Partner input (BEAI) · 5.1 Legal · 5.2 Ethical · 5.3 State of the art · 5.4 Market · 5.5 Differentiation · Research gaps · Sources

PART 6 · YASSMINE NOUISSE

Synthesis, Objective Tree & Global Paragraph

6.1 Findings · 6.2 Revealed · 6.3 Emerging themes · 6.4 Structural gaps · 6.5 Conclusions · 6.6 Guiding Questions coverage · 6.7 Foundation for the solution · Alignment with BEAI · Objective tree · Global paragraph

ANNEX

Completeness check for each part

What is missing, against the Investigate plan, with the person responsible

00

About this report

How the investigation is organised and where each part stands

Engage framework

BIG IDEA

Knowledge Revolution

ESSENTIAL QUESTION

How can we make the best knowledge available to every learner, wherever and whenever they are?

CHALLENGE

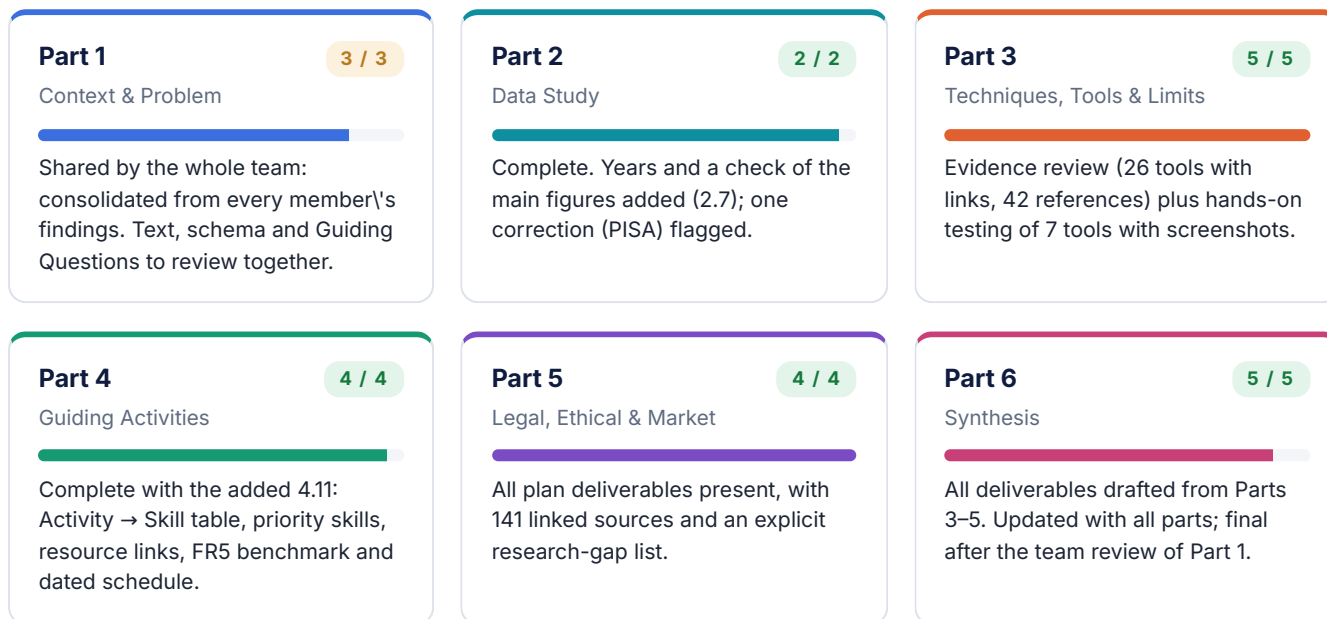
Allow every student to ask questions and learn face to face from the best teachers and minds, in their own language.

The investigation is framed by the reference document **BEAI™ Studio — Agent & Digital-Twin Creation Platform** (FBM V1A26): a four-layer platform (Presentation, Application/Services, Data, Infrastructure, plus Monitoring) with two paths, *Clone Myself* and *Build an AI Agent*, each leading to a pre-recorded video or a real-time interactive agent.

Team and parts

Part	Title	Responsible	Plan deliverable	Status
01	Context & Problem Structuring	All members	Structured text + Context → Problem → Impact schema + Guiding Questions	TEAM REVIEW
02	Data Study	Maryem Ouichka	Data table (Type · Source · Size · Scope · Availability · Limits) + data gaps	COMPLETE
03	Techniques, Tools & Limits	Med Rayen Ayat (3A) Islem Tellili (3B)	Comparison table + "What is missing today"	COMPLETE
04	Guiding Activities & Resources	Malek Tirellil	Activity · Skill · Why · Resource · Type · Link + priority skills	COMPLETE
05	Legal, Ethical & Market Study	Youssef Jouini	Solutions table + legal/ethical table + critical paragraph + sources	COMPLETE
06	Synthesis, Objective Tree & Global Paragraph	Yassmine Nouisser	Objective tree + synthesis + global paragraph + coverage table + requirements	PROVISIONAL

Completeness dashboard



Ratios count the deliverables listed in the Investigate plan for each part. The detailed check is in the Annex.

How to read this report

- Sections marked **COMPLETION** (Part 1 consolidation, 2.7, 3B references, 4.11) were added to fill the gaps of the plan; they are clearly separated from the members' original work.
- Parts 2, 3, 4 and 5 are reproduced in full (Part 2 translated from French), with every table, figure and screenshot of the original documents. Only the layout, the numbering and the visual charter have been harmonised.
- Each part keeps its own colour. Module badges **01** to **06** refer to the six modules of the objective tree.
- The working notes "To be used for Part 5" are identical to section 3B.1 of Part 3 (tool testing); they appear once, in Part 3B, and Part 5 refers to them.
- Part 6 separates documented facts from interpretations and cites the part each finding comes from.

01

RESPONSIBLE

All members

Shared part · October 2026

Context & Problem Structuring

Guiding Questions — Context → Problem → Causes → Consequences → Stakeholders → Limitations

A SHARED PART

Context and Problem is the responsibility of the whole team. This version consolidates the context and problem evidence brought by every member — data (Part 2), techniques and tools (Part 3), activities (Part 4), legal, ethical and market study (Part 5) — and the Engage phase. Each fact gives the part where its source is documented. To be reviewed together before the presentation.

1.1 Context

Knowledge has never been so available, yet access to the people who can explain it remains scarce, unequal and tied to place, money and language.

HOW LEARNERS REACH KNOWLEDGE TODAY

- **In class:** a teacher in front of a group, with limited time for individual questions.
- **After class:** office hours, e-mail, recordings, textbooks; in Tunisia, paid private lessons (Part 5, 5.4).
- **Online:** MOOCs, video platforms, open resources such as OpenStax, MIT OCW and Khan Academy (Part 2, 2.2).
- **With AI:** general chatbots and tutoring modes (ChatGPT Study Mode, Khanmigo, Gemini Guided Learning) used massively by students (Parts 3A, 5).

HOW IT EVOLVED

- **Internet era:** content becomes free and global, but one-way.
- **2012–2020:** MOOCs and video conferencing remove distance, not the teacher's time limit.
- **2023–2026:** generative AI answers instantly; talking avatars (D-ID, Tavus, HeyGen) give AI a face and a voice, mostly for business (Parts 3, 5).
- **2025–2026:** regulation arrives: EU AI Act bans emotion recognition in education and requires AI transparency (Part 5, 5.1).

Dimension	Situation	Source (part)
Teachers	44 million additional teachers needed worldwide by 2030; 15 million in sub-Saharan Africa; 4.3 million in Northern Africa	UNESCO 2024, 2023 (Part 5)
Learning outcomes	About 70% of 10-year-olds in low- and middle-income countries cannot read and understand a simple text ("learning poverty")	World Bank, 2022 (Part 2)

Dimension	Situation	Source (part)
Connectivity	About 68% of the world online in 2024; 2.2 billion people still offline; 36% of Africa online; rural use 58% vs urban 85%	ITU 2024–2025 (Parts 2, 3A, 5)
Geography (Tunisia)	Private tutoring paid by 67% of families, about 1,468 million TND a year, 15–20 TND per hour; costs peak in exam years (bac)	FTDES 2023; La Presse 2025–2026 (Part 5)
Language	Learners think and ask in Derja, Arabic, French and English, often mixed; AI quality for Tunisian Arabic is among the weakest measured	TounsiBench, ArabCulture-Dialogue (Part 3A)
Use of AI	95% of UK undergraduates use AI, but only 38% get AI tools from their institution	HEPI 2026 (Part 5)

Table 1.1 The context in figures

1.2 Problem

CENTRAL PROBLEM

Learners have access to content but not to understanding: when they are stuck, they cannot ask the best teachers and experts a question, privately, at the moment they need it, in their own language — unless they can pay for it.

The Engage phase illustrated the problem with six learner situations: attended but couldn't ask in front of everyone; attended and still lost; missed the lecture; rewatched the recording twice with the same confusion; tried an existing platform and stayed stuck; e-mailed and got no answer (Part 3A, Table 3.1).

Obstacle	What it means for the learner	Evidence (part)
Availability	The teacher's time does not multiply: office hours, e-mails and live sessions cannot answer every student	UNESCO 2024 (5); Engage
Cost	Personal help is bought privately (tutoring, mentoring platforms from about \$120/month)	FTDES 2023 (5); MentorCruise (3A)
Distance and access	Rural and low-income learners are the least connected and the furthest from experts	ITU 2025 (3A, 5)
Language	Explanations and AI tools work best in English and standard languages, not in the learner's dialect	NLLB, TounsiBench (3A)
Time	Questions arise at night, before exams, outside office hours	Engage; Part 4 survey IS1 (planned)
Psychological barrier	Students don't dare ask in public; fear of judgment	Engage; Part 5 (baseline)

Table 1.2 Obstacles to quality knowledge

Why existing resources fall short: recordings and MOOCs are one-way (edX completion about 3%); general chatbots answer fluently but can be wrong and give answers instead of building understanding; avatar platforms have a face but no pedagogy, no learner model and no link to the real teacher (Parts 3A, 3B, 5).

1.3 Causes

Type	Cause	Evidence (part)
Human	Shortage and uneven distribution of qualified teachers; experts' time is limited and expensive	UNESCO 2024 (5)
Economic	Personal help is a private market; families pay according to income; per-minute AI avatar pricing (about \$21 per live hour) is not affordable	FTDES 2023, Tavus (5)
Geographic	Experts concentrated in cities and universities; weak connectivity in rural areas	ITU 2025 (5); Part 2 (no regional data)
Linguistic	Few resources and little speech data in Tunisian Arabic; AI models under-trained on dialects and code-switching	Parts 2 (2.5), 3A
Technical	Current AI is fluent but unreliable (hallucinations, invented transcripts); tools are fragmented (tutor <i>or</i> face <i>or</i> translation)	Part 3A (12, 18)
Institutional	Institutions lag behind students' use of AI; no validated tools, no consent process for teachers' digital likeness	HEPI 2026, UNESCO (5); Part 3B
Pedagogical	One-way formats; unguarded AI gives answers and can harm learning	Bastani et al. 2025 (3A, 5)

Table 1.3 Causes

How the causes interact: the teacher shortage creates the demand; the private market answers it for those who can pay; AI could lower the cost, but its weakness in local languages and its reliability problems hit hardest the learners who already have the least access. The causes reinforce each other, so solving one alone (for example, translation) does not remove the problem.

1.4 Consequences

Type	Consequence	Evidence (part)
Educational	Confusion that is never resolved; passive learning; drop-out from online courses (edX completion 6% → 3%, second-year return 38% → 7%)	Reich & Ruipérez-Valiente 2019 (5)
Social	Success increasingly tied to family income; exam years become a financial burden	La Presse 2026 (5)
Economic	About 1,468 million TND a year spent on tutoring in Tunisia; up to 15,000 TND for some bac candidates	FTDES 2023; La Presse 2026 (5)
Skills	Learning poverty (about 70% in low- and middle-income countries); gaps in maths and sciences that persist into higher education	World Bank 2022; PISA, TIMSS (2)
Long-term	Unequal access to the best knowledge reproduces inequality between regions and countries (SDG 4 at risk)	UNESCO (5)
Risk of the wrong fix	Students turn to unsupervised AI that may be wrong or replace thinking (–17% exam grades without guardrails)	Bastani et al. 2025 (3A)

Table 1.4 Consequences

1.5 Stakeholders

Stakeholder	Role	Needs and interests	Responsibilities
Learners	Use the knowledge	Quick, private, reliable answers in their language; low cost; low bandwidth	Learn actively; use AI honestly
Teachers / professors	Produce and transmit knowledge	Reach more students without more hours; control over their likeness and content	Validate content; consent; supervise the mentor
Experts and mentors	Bring rare expertise	Recognition; fair use of their knowledge; payment	Accuracy; consent
Institutions (ESPRIT, universities, schools)	Organise learning, choose tools	Better results, retention, reputation; compliance	Validate AI tools before use (UNESCO); data protection
Families	Fund and support learners	Lower tutoring costs; protection of minors	Parental consent for minors
Platforms and providers	Supply the technology (D-ID, Tavus, LLM APIs)	Market growth	Consent safeguards; transparency; data processing contracts
Regulators	Set the rules (INPDP, EU, ministries)	Protection of personal and biometric data; fair education	Authorisations; enforcement
International bodies	UNESCO, Teacher Task Force, World Bank	SDG 4 progress	Guidance; funding of pilots

Table 1.5 Stakeholders (sources: Parts 4 and 5)

1.6 Limitations

HUMAN AND ECONOMIC

The human expert remains irreplaceable for judgment and care; professors may refuse to be cloned; GPU and per-minute costs limit free access.

TECHNICAL

Hallucinations, invented transcripts, weak dialect quality, latency and bandwidth; no evidence yet that an avatar teaches better than voice or text.

LEGAL AND ETHICAL

Biometric data and explicit consent; transfers abroad (INPDP); minors; ban on emotion recognition; AI label; risk of deepfakes and impersonation.

QUESTIONS THAT REMAIN HARD

Who is responsible when the avatar is wrong? How to keep trust when realism grows? How to measure real learning, not satisfaction? How to reach learners who are offline?

1.7 Schema: Context → Problem → Impact

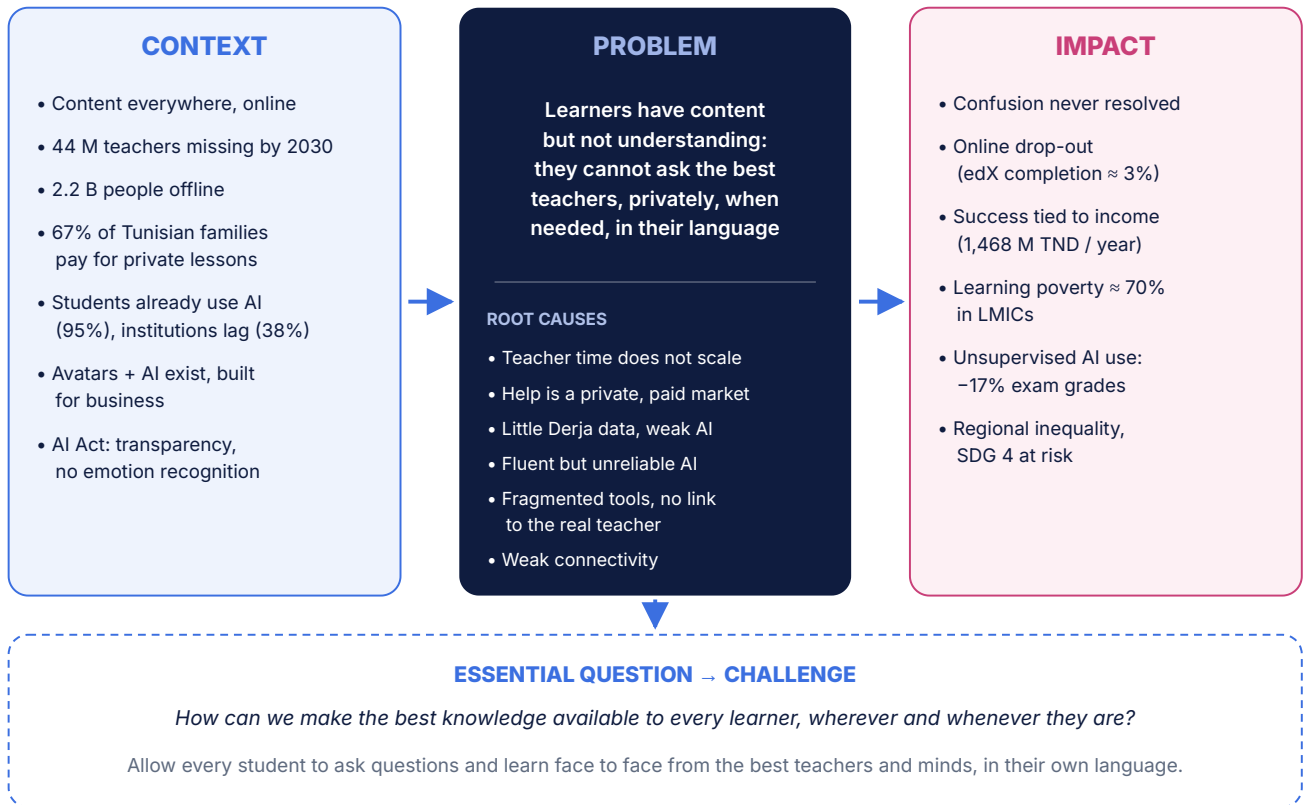


Figure 1.1 Context → Problem → Impact

1.8 Guiding Questions

The list below follows the structure of the Investigate plan. The module-level questions GQ1.1 to GQP.3 of Part 4 (Table 4.4) complete it for the technical modules.

Code	Area	Guiding Question	Answered in
GQ-C1	Context	How do learners access teachers, mentors, experts and resources today?	1.1 · 3 · 5
GQ-C2	Context	How has access to knowledge changed with the Internet, platforms and AI?	1.1 · 3A
GQ-C3	Context	How does access differ by region, income, infrastructure and language?	1.1 · 2 · 5
GQ-P1	Problem	What prevents learners from getting answers from the best teachers when they need them?	1.2
GQ-P2	Problem	Why do recordings, MOOCs, chatbots and avatar platforms not solve it?	1.2 · 3A · 3B
GQ-K1	Causes	Which human, economic, geographic, linguistic and technical factors create the problem, and how do they interact?	1.3
GQ-K2	Consequences	What are the educational, social and economic effects of unequal access to experts?	1.4
GQ-S1	Stakeholders	Who produces, transmits, funds and regulates knowledge, and what does each need?	1.5
GQ-S2	Stakeholders	Under which conditions would professors agree to a digital version of themselves?	4 (IS2)
GQ-D1	Data	What data exists on learners, courses, languages and access, and what is missing?	2
GQ-T1	Techniques	Which techniques and tools exist, and why are they not enough?	3A · 3B
GQ-L1	Legal & ethics	Which rules apply to biometric data, minors, transfers abroad and AI transparency?	2.6 · 5
GQ-M1	Market	Who needs this most, and what already exists?	5
GQ-V1	Value	Does an interactive mentor help students learn better than a video or a text chatbot?	4 (EX7)
GQ-L2	Limitations	Which limits remain, and who is responsible when the mentor is wrong?	1.6 · 5

Table 1.6 Plan-level Guiding Questions (draft)

02

Data Study

Data availability, needs & gaps — professor avatars for every learner

RESPONSIBLE
Maryem Ouichka
October 2026

Project: an application in which a professor (or an expert) creates a digital avatar, with their own face and voice, to explain their courses to learners, in the learners' language and at any time.

PRINCIPLE OF THE STUDY

Building this product needs four blocks of data: **(1) the person** (face, voice, consent), **(2) the course** (content to explain), **(3) the language** (multilingual speech and text) and **(4) the learner and their context** (who learns, with what access).

JURISDICTIONS AND RELIABILITY

Reference jurisdictions: **Tunisia, France/EU, Nigeria, USA**. The author states that the figures come from their own knowledge and must be re-checked; they are marked "to verify" where relevant.

Translated from the French original for the consolidated report; all content is kept.

2.1 Types of data needed

Block	Data	Product stage	Why
A. The professor	Photo/video (1 min), voice sample (30 to 60 s), signed consent, proof of liveness	Avatar creation and voice cloning	Without these data there is no digital twin; they are the most sensitive legally
B. Course content	Official curricula, textbooks, courses, slides, transcripts, corrected exercises, glossaries	Knowledge base (RAG)	The avatar must explain exact, sourced content, not invent it
C. Multilingual speech	Annotated audio in Arabic (standard and Tunisian), French, English, African languages	ASR (listening) and TTS (voice)	The avatar must understand the student and answer with a natural voice
D. Text and language	Text corpora, translation pairs, scientific terminology per language	LLM, translation, 120+ language dubbing	Quality of explanations and translation in each language
E. Learners and contexts	Age, level, mother tongue, urban/rural setting, disability, school drop-out	Defining the audience and adapting	Knowing for whom the avatar explains, and at what level
F. Educational and school data	Pass rates, attainment by subject, pupil/teacher ratio, school calendar	Choosing priority subjects and levels	Targeting the courses where the learning gap is largest
G. Internet and technology	Coverage, bandwidth, data cost, devices, electricity, connected schools	Deployment (real-time video, latency < 500 ms)	Choosing between light pre-recorded video and a real-time agent
H. Teachers, mentors, experts	Numbers, geographic distribution, shortage by discipline	Recruiting partner professors	Measuring the lack of experts the avatar is meant to fill

Block	Data	Product stage	Why
I. Quality evaluation	Question–answer sets by subject, lip-sync tests, student feedback	Validation before going online	Checking that the avatar explains correctly and that students learn
J. Metadata	Licence, date, language, level, source, subject, author	All stages	Tracing the origin of content and respecting licences

Table 2.1 Data needed by the product, in ten blocks

2.2 Availability and what we have (2.2 and 2.3)

Deliverable table: Type · Source · Size · Scope/Population · Availability · Limits. A "Use for the project" column shows what each source brings; links are in the Source column.

Type	Source (link)	Size	Scope / Population	Availability	Limits	Use for the project
Course content	OpenStax openstax.org	Dozens of textbooks (higher ed, high school)	Maths, sciences, social sciences, in English (some Spanish editions)	Open (CC BY)	English only, not aligned with national curricula	Safe, sourced RAG knowledge base
Course content	MIT OpenCourseWare ocw.mit.edu	~2,500 courses	University level, videos and materials	Open (CC BY-NC-SA)	Non-commercial use only; English	Course model and test content
Course content	OER Commons / MERLOT oercommons.org , merlot.org	Tens of thousands of resources	All levels and subjects	Open (varied licences)	Uneven quality, licences to check one by one	Complete content by discipline
Course content	Khan Academy khanacademy.org	Thousands of lessons and exercises	Primary to early university	Free (limited reuse)	Protected content, partnership needed	Pedagogical reference
Course content	Wikipedia / Wikibooks dumps dumps.wikimedia.org	Millions of articles, 300+ languages	General knowledge	Open (CC BY-SA)	Uneven quality; "share alike" licence	Basic multilingual content
Course content (TN)	Centre National Pédagogique, Ministry of Education education.gov.tn (link to confirm)	Official Tunisian textbooks	Primary to bac, Arabic and French	Viewable; reuse to be authorised	State copyright, PDF formats	Only source aligned with the Tunisian curriculum
Course content (FR)	Éduscol, Éducation nationale eduscol.education.fr	Official curricula	Pre-school to final year	Open (curricula)	Not the textbooks (private publishers)	Align courses with French curricula
Transcribed video lectures	YouTube-Commons (Hugging Face, PlelAs)	~2 M CC-BY videos (to verify)	Multilingual transcripts	Open (CC BY)	Variable quality; non-educational content mixed in	Examples of courses explained orally
Multilingual speech	Mozilla Common Voice commonvoice.mozilla.org	100+ languages, tens of thousands of hours	Read, annotated voices	Open (CC0)	Uneven by language; weak for Tunisian dialect	Train/evaluate ASR and TTS
Multilingual speech	FLEURS (Google, Hugging Face)	~102 languages	Read speech aligned with FLORES	Open (CC BY)	Small volume per language (~12 h)	Multilingual ASR evaluation

Type	Source (link)	Size	Scope / Population	Availability	Limits	Use for the project
Multilingual speech	Multilingual LibriSpeech / VoxPopuli (Hugging Face)	Tens of thousands of hours	Audiobooks, parliamentary debates (European languages)	Open	Little Arabic; non-teaching style	Train ASR in French and English
Spoken dialectal Arabic	MGB-3/5, TARIC, MADAR (access on request)	Tens of hours to restricted sets	Egyptian and Tunisian Arabic, urban dialects	On request / research	Small sizes, restrictive licences	Only basis to understand spoken Tunisian Arabic
Talking faces	VoxCeleb, HDTF, CelebV-HQ, MEAD (university sites)	Hundreds of thousands of clips	Talking faces, celebrities and volunteers	Research only	Non-commercial licences; people have not consented to commercial use	Train/evaluate animation and lip-sync, not in production without a licence
Multilingual text	FLORES-200, OPUS, OSCAR/CC-100 (Hugging Face, opus.nlpl.eu , oscar-project.org)	200 languages (FLORES); massive corpora	Translation, pre-training	Open (varied licences)	Web bias, few dialects, thin educational domain	Translation and dubbing
African languages	Masakhane masakhane.io	Dozens of languages, evaluation sets	African NLP	Open	Small size, no educational domain	Extend to Hausa, Yoruba, etc.
Knowledge evaluation	MMLU, ARC, SciQ (Hugging Face)	Thousands of questions	School and university level	Open	Mostly English; not the Tunisian curriculum	Test the accuracy of explanations
Education (indicators)	UNESCO UIS data.uis.unesco.org	Thousands of indicators, ~200 countries	Enrolment, teachers, literacy	Open (API/CSV)	1 to 3 years' lag, aggregated	Size the market and the need
Education (inequalities)	UNESCO WIDE education-inequalities.org	Household surveys, ~160 countries	Gaps by wealth, gender, rural/urban	Open	Depends on survey cycles	Identify excluded audiences
Education	World Bank EdStats / Learning Poverty datatopics.worldbank.org/education	Annual series, ~200 countries	Learning poverty (~70% in low- and middle-income countries, 2022 estimate)	Open (CC BY 4.0)	Modelled estimates, no local level	Justify the scale of the need
Skills	OECD PISA oecd.org/pisa	~690,000 students, 81 systems (2022)	15-year-old students (Tunisia included)	Open (microdata)	Enrolled students only; 3-year cycle (PISA 2025 to verify)	Identify subjects with large gaps
Skills	TIMSS / PIRLS timssandpirls.bc.edu	Tens of thousands of students per cycle	Grades 4 and 8	Open	Maths, science, reading only	Choose priority subjects

Type	Source (link)	Size	Scope / Population	Availability	Limits	Use for the project
Skills (regional)	PASEC, ASER, Uwezo, SACMEQ (official sites)	Surveys of hundreds of thousands of children	Primary, rural, Africa and India	Open / on request	Tunisia not in PASEC; different methods	Benchmarks for Africa
Households	UNICEF MICS (Tunisia MICS6, 2018) mics.unicef.org ; DHS dhsprogram.com	A few thousand households per country	School attendance, ICT access	Open on registration	Old; DHS funding disrupted (to verify)	Family profiles and access to devices
Internet access	ITU DataHub datahub.itu.int	~190 countries	Internet users (~68% worldwide in 2024), subscriptions	Open	National level, no school detail	Estimate the reachable population
Internet access	Giga giga.global	More than 2 M schools (to verify)	School connectivity	Open (partial)	Uneven coverage	Choose light video or real-time agent
Internet access	GSMA Intelligence / DataReportal gsmaintelligence.com , datareportal.com	Annual series	Mobile, usage, cost	Partial; GSMA paid	Aggregated, commercial method	Assess smartphone access
Teachers	OECD TALIS oecd.org/talis ; Teacher Task Force teachertaskforce.org	~50 systems; ~44 M teacher shortfall by 2030	Teachers, training, needs	Open	Little MENA data; quality not measured	Justify the need for digital professors
Experts	OpenAlex openalex.org ; ORCID orcid.org	~250 M works	Researchers and publications	Open (CC0)	Academic experts, not mentor availability	Find partner professors
National data (TN)	INS Tunisia ins.tn ; data.gov.tn ; Ministries (Education, Higher Education)	2024 census; statistical yearbooks	Pupils, students, teachers, institutions	Partial	PDF, irregular updates, few microdata	Local reference data
National data (FR/EU)	data.gouv.fr / DEPP data.education.gouv.fr ; Eurostat ec.europa.eu/eurostat	Exhaustive	Pupils, teachers, digital uses	Open	Context different from MENA	European comparison
Learning analytics	OULAD analyse.kmi.open.ac.uk/open-dataset ; EdNet github.com/rriid/ednet	32,600 students; 784,000 users, 131 M interactions	Online paths, question–answers	Open (CC BY / research)	UK and Korean contexts	Model progression and drop-out
Legal texts	GDPR, AI Act eur-lex.europa.eu ; INPDP inpdp.nat.tn ; NDPC ndpc.gov.ng ; FERPA studentprivacy.ed.gov	Official texts	Data protection, AI	Open	Legal interpretation needed	Compliance framework

Table 2.2 Available data sources — Type · Source · Size · Scope · Availability · Limits

WHAT WE CAN OBTAIN THROUGH PARTNERS

- **Partner professors:** recordings (face, voice), courses, slides, exercises, under contract and with explicit consent.
- **Tunisian universities and schools:** course materials, curricula, anonymised pass rates.
- **Ministries and inspectorates:** official curricula, textbooks, regional statistics.
- **Telecom operators:** coverage and data cost by region.

ALREADY FREELY ACCESSIBLE

UIS, WIDE, EdStats, PISA/TIMSS, ITU, Giga, Common Voice, FLEURS, FLORES, OpenAlex, OpenStax, OER, Wikipedia.

RELIABILITY AND RECENCY

- **Reliable and recent:** UIS, World Bank, ITU, OECD (1 to 2 years' lag).
- **Ageing:** MICS Tunisia (2018), DHS, PASEC.
- **Use with caution:** web corpora and research sets (uneven quality and licences).

2.4 What we do not have yet

- **Professor data:** no database of consented recordings to train or validate an avatar. They will be collected professor by professor.
- **Tunisian Arabic speech** of good quality and volume, and local African languages, with school vocabulary.
- **Learners' real needs:** frequent questions, sticking points, preferences (short video, voice, language).
- **Effectiveness of a professor avatar:** no local data on learning, acceptability or trust.
- **Regional data:** learning level and connectivity by governorate or school.
- **Access to experts:** map of available professors and mentors by discipline and region.
- **Recent and representative data:** household surveys are more than 5 years old, and out-of-school children are absent.

2.5 Data gaps

Gap	Why it limits the analysis or the product
Local course content not available under licence	Without it, the avatar explains foreign or generic content, and the RAG lacks reliable sources (risk of hallucination)
No educational Tunisian Arabic speech corpus	ASR and TTS less accurate, so the avatar misunderstands students and sounds wrong
Professor face/voice data to be collected one by one	Each avatar requires consent, recording quality and verification, which slows scaling
Talking-face datasets limited to research	They cannot legally be used for a commercial product without another licence
No regional granularity	Impossible to target the areas where access to knowledge is weakest
Students' needs not measured	Risk of designing courses around school results rather than what students are trying to understand
No data on AI tutors' effectiveness in the local context	No comparison with current methods, so no proof of impact
Old data and selection bias	They reflect neither the post-COVID period nor the rise of mobile, and excluded learners are invisible

Table 2.3 Main data gaps

2.6 Legal constraints on data

Subject	Tunisia	France / EU	Nigeria	USA
General framework	Organic Law 2004-63; INPDP; Convention 108	GDPR	NDPA 2023 (NDPC)	No general federal law; state laws (e.g. CCPA)
Consent	Mandatory; INPDP authorisation for some processing	Free, informed, specific, revocable	Same	Variable
Biometrics (professor's face, voice)	Sensitive data, strict regime	Special category (Art. 9), explicit consent	Sensitive data	Specific laws (e.g. BIPA, Illinois)
Image and voice rights	Personality right (to confirm)	Image right (Art. 9 French Civil Code)	Privacy; image regulated	"Right of publicity" by state
Minors (students)	Guardian consent; authorisation for children's data	13–16 threshold depending on the state (Art. 8)	Enhanced protection	COPPA (< 13), FERPA (school records)
Copyright on courses	Intellectual property: the course belongs to the professor or institution (to confirm by contract)	Copyright; limited teaching exception	Copyright Act	Copyright; fair use uncertain for AI
Transfers abroad	INPDP authorisation	Standard clauses / adequacy decision	NDPA conditions	Free, except by contract
AI / synthetic content	No dedicated law (to confirm)	AI Act: transparency of synthetic content; education as high risk (timeline to verify)	In progress	State laws on deepfakes

Table 2.4 Data constraints in the four jurisdictions

What can be collected, used or shared

Level	Data and conditions
FREE	Aggregated public statistics (UIS, ITU, PISA...) and openly licensed content (CC BY: attribution; CC BY-SA: share alike; CC BY-NC: no commercial use).
WITH CONDITIONS	Survey microdata (usage agreement, no re-identification); partners' courses and data (contract, purpose, anonymisation).
HIGH RISK	Professor's face and voice (biometrics), data of minor students, health/disability data. They require explicit and revocable consent, a legal guardian for minors, minimisation, limited retention and secure hosting.
PUBLIC ≠ FREE	A professor's video or voice online authorises neither cloning nor training without their written agreement, and a school textbook remains protected by copyright.
PRODUCT RULE	Consent and liveness check before any avatar creation, with a clear notice that the content is AI-generated.

Sources used

UNESCO UIS, WIDE, GEM Report; World Bank EdStats; OECD PISA, TALIS; IEA TIMSS/PIRLS; PASEC; ASER; Uwezo; SACMEQ; UNICEF MICS; DHS; ITU DataHub; Giga; GSMA; DataReportal; INS Tunisia; data.gov.tn; data.gouv.fr/DEPP; Eurostat; Éduscol; OpenStax; MIT OpenCourseWare; OER Commons; MERLOT; Khan Academy; Wikipedia/Wikibooks; YouTube-Commons; Mozilla Common Voice; FLEURS; Multilingual LibriSpeech; VoxPopuli; MGB/TARIC/MADAR; VoxCeleb; HDTF; CelebV-HQ; MEAD; FLORES-200; OPUS; OSCAR; Masakhane; MMLU; ARC; SciQ; OULAD; EdNet; OpenAlex; ORCID; Teacher Task Force; texts: Tunisian Law 2004-63, GDPR, AI Act, NDPA 2023, COPPA, FERPA, BIPA.

2.7 Years of the sources and check of the figures COMPLETION

The author asked for the figures of Table 2.2 to be verified, and the plan requires the year of each source. This section adds both. Each figure was compared with what the publisher reports as of mid-2026. "Consistent" means the order of magnitude matches. It does not replace opening each link before the presentation.

Source	Publisher	Year / edition	Check of the figure in Table 2.2	Status
OpenStax	Rice University	Continuously updated (2012–)	Size consistent (~70 titles); licence CC BY confirmed	CONSISTENT
MIT OpenCourseWare	MIT	Continuously updated (2001–)	~2,500 courses: consistent with MIT's published figure	CONSISTENT
OER Commons / MERLOT	ISKME / Cal State	Continuously updated	Order of magnitude plausible; licences vary per resource	CONSISTENT
Khan Academy	Khan Academy	Continuously updated	Reuse restricted: CC BY-NC-SA for most videos, partnership needed for platform content	CONSISTENT
Wikimedia dumps	Wikimedia Foundation	Monthly dumps	300+ language editions: consistent	CONSISTENT
Tunisian textbooks	CNP / Ministry of Education	Editions per school year	Link to confirm, as noted by the author	TO VERIFY
YouTube-Commons	PleIAs (Hugging Face)	2024	About 2 million CC-BY videos: consistent with the dataset card	CONSISTENT
Mozilla Common Voice	Mozilla Foundation	Releases every ~3 months	100+ languages, tens of thousands of validated hours: consistent	CONSISTENT
FLEURS	Google	2022	102 languages, about 12 h each: consistent	CONSISTENT
Multilingual LibriSpeech / VoxPopuli	Meta AI	2020 / 2021	Sizes consistent; Arabic absent from MLS	CONSISTENT
MGB-3/5, TARIC, MADAR	QCRI, academic teams	2017–2019	Sizes small; access on request: consistent	CONSISTENT
VoxCeleb, HDTF, CelebV-HQ, MEAD	Oxford VGG; academic teams	2017–2022	Research-only licences confirmed	CONSISTENT
FLORES-200, OPUS, OSCAR	Meta; Univ. Helsinki; Inria	2022; continuous; 2019–2023	200 languages (FLORES): consistent	CONSISTENT
Masakhane	Masakhane community	Continuous (2019–)	Consistent	CONSISTENT
MMLU, ARC, SciQ	Academic benchmarks	2020; 2018; 2017	Consistent	CONSISTENT
UNESCO UIS	UNESCO Institute for Statistics	Data releases twice a year	1–3 years' lag: consistent	CONSISTENT
UNESCO WIDE	UNESCO GEM Report	Updated with each GEM report	~160 countries: consistent	CONSISTENT
World Bank Learning Poverty	World Bank	2022 update	About 70% in low- and middle-income countries (2022 simulation): consistent	CONSISTENT

Source	Publisher	Year / edition	Check of the figure in Table 2.2	Status
OECD PISA	OECD	PISA 2022 (published Dec 2023); PISA 2025 results expected late 2026	About 690,000 students, 81 systems: consistent. Correction: Tunisia did not take part in PISA 2022; its last participation was PISA 2015	CORRECTION
TIMSS / PIRLS	IEA / Boston College	TIMSS 2023 (published Dec 2024); PIRLS 2021	Check Tunisia's latest participation before citing Tunisian results	TO VERIFY
PASEC, ASER, Uwezo, SACMEQ	CONFEMEN; Pratham; Twaweza; SACMEQ	PASEC 2019; ASER yearly	Tunisia not in PASEC: consistent	CONSISTENT
UNICEF MICS / DHS	UNICEF; ICF (USAID-funded)	MICS6 Tunisia 2018; DHS disrupted in 2025	Consistent; DHS was stopped by the US funding freeze of early 2025 and only partly resumed	CONSISTENT
ITU DataHub	ITU	Facts and Figures 2024 and 2025	About 68% online in 2024; 2.2 billion offline in 2025: consistent	CONSISTENT
Giga	UNICEF / ITU	Continuous (2019–)	More than 2 million schools mapped: consistent with Giga's public map	CONSISTENT
GSMA / DataReportal	GSMA; Kepios	Yearly (Digital 2026)	Consistent	CONSISTENT
TALIS / Teacher Task Force	OECD; UNESCO	TALIS 2024 (results 2025); Global Report on Teachers 2024	44 million teachers needed by 2030: consistent	CONSISTENT
OpenAlex / ORCID	OurResearch; ORCID	Continuous	About 250 million works: consistent	CONSISTENT
INS / data.gov.tn	Institut National de la Statistique	Census 2024; yearly statistical yearbooks	Consistent	CONSISTENT
data.gouv.fr / DEPP / Eurostat	French State; EU	Yearly	Consistent	CONSISTENT
OULAD / EdNet	Open University; Riiid	2017; 2020	32,593 students; 784,309 users and 131 million interactions: consistent	CONSISTENT
GDPR, AI Act, INPDP, NDPC, FERPA	EU; Tunisia; Nigeria; USA	GDPR 2016/679; AI Act 2024/1689; Law 2004-63; NDPA 2023; FERPA 1974	Consistent; status of Tunisian bill PLO 095/2025 still to confirm (Part 5)	TO VERIFY

Table 2.5 Year of each source and check of the figures — 27 consistent, 3 to verify, 1 correction

CORRECTION TO APPLY IN TABLE 2.2

OECD PISA: Tunisia is listed as "included". Tunisia took part in PISA up to 2015 but not in PISA 2018 or 2022. For recent Tunisian learning data, use national exam statistics (bac, brevet) and UNESCO UIS indicators instead.

03

PART 3 RESPONSIBLE
Member 3 — Med Rayen Ayat
Hands-on testing: Islem Tellili
October 2026

Techniques, Tools & Limits

A conceptual, evidence-based review of the technologies that could bring knowledge to every learner

STRUCTURE OF PART 3

3A — Evidence-based review (Med Rayen Ayat): nine technique cards, a comparative table of 26 tools, why current tools are not enough, what is missing, caveats and 42 references.

3B — Hands-on testing (Islem Tellili): seven tools tested with screenshots, technique notes, feature coverage matrix, observed limits and gaps.

Scope note: this part stays at the conceptual level (high-level, not implementation). It does not design the team's solution.

3A · Executive summary

Today's techniques can each deliver one part of "the best knowledge, for every learner, anywhere, anytime": a fluent explanation, a translation, a voice, a face, or a live session. None of them reliably combines accuracy, sound pedagogy, the learner's own language or dialect, and real human accountability. The evidence shows these gaps are measurable, not hypothetical.

FINDING 1 — THE TECHNIQUES WORK, BUT ONLY UNDER CONDITIONS

A carefully designed AI tutor gave 194 Harvard physics students median learning gains more than double those of an in-class active-learning group [1]. In a large high-school trial, an unguarded GPT-4 tutor raised practice scores, but students scored worse than the control group once access was removed [3]. The difference is pedagogical design and guardrails, not raw model power.

FINDING 2 — RELIABILITY IS THE WEAKEST LINK AT EVERY LAYER

Retrieval-grounded legal AI tools still hallucinated 17–33% of the time [12]. OpenAI Whisper invented sentences in about 1% of audio clips, 38% of them harmful [18]. LLMs perform measurably worse in Tunisian Arabic than in Modern Standard Arabic or English [14, 15]. Watermarks and C2PA metadata can be removed, and platforms routinely strip them [39, 42].

FINDING 3 — THE BIGGEST GAP IS HUMAN, NOT TECHNICAL

The strongest results so far come from AI that supports or is supervised by a human expert: Tutor CoPilot (+4 percentage points in mastery) [4] and Google LearnLM with Eedi (66.2% vs 60.7% transfer) [2]. Yet human mentors stay scarce and costly, 2.2 billion people are still offline [23], and no tool links an AI version of a teacher to the real teacher's consent, oversight and accountability.

3A.1 · Introduction and scope

The Engage phase framed the problem as follows: students have access to knowledge, but access does not always mean understanding. Six situations illustrate it:

Before / during class	After class	Seeking help
Attended, couldn't ask in front of everyone	Missed the lecture	Tried an existing platform, still stuck
Attended, still lost	Rewatched the recording twice, same confusion	Nobody answered my email

Table 3.1 Six learner situations from the Engage phase

Part 3 surveys the existing techniques and tools that could address these situations and judges each one against evidence. Following the work plan, it:

- **stays conceptual:** how each technique works in principle, not how to implement it;
- **covers nine techniques:** the seven required by the plan plus two added because they are central to the project direction (digital humans and AI transparency);
- **does not invent limits:** every limit is tied to a named study, test, official document or credible news report;
- **does not design the solution:** the BEAI™ Studio direction (teacher avatars, RAG-grounded agents, translation, speech, lip-sync, transparency) is used only as a lens for judging relevance.

3A.2 · Existing techniques: nine technique cards

1 AI Tutoring / Conversational Learning

Principle. A large language model (LLM) generates replies in a dialogue. A "tutor" layer, usually a system prompt or fine-tuning, steers it to ask questions, give hints and check understanding instead of handing over answers (e.g., Khanmigo, ChatGPT Study Mode, Google LearnLM).

STRENGTHS

- Patient, judgment-free questioning at any hour — directly addresses "I couldn't ask in front of everyone".
- Harvard RCT (n = 194, physics): a custom tutor built on active-learning principles produced median learning gains more than double those of in-class active learning, in less time, with higher engagement [1].
- UK RCT (165 students, five schools): human-supervised LearnLM matched expert tutors on mistake correction; its students were 5.5 points more likely to solve novel problems on the next topic (66.2% vs 60.7%) [2].

LIMITS (EVIDENCE)

- **Unguarded AI can harm learning.** In a preregistered RCT with nearly a thousand Turkish high-school students, GPT-4 access raised practice grades by 48% ("GPT Base") and 127% ("GPT Tutor"), but once removed, GPT Base students scored about 17% lower than students who never had it. The guardrailed version largely removed the harm [3].
- **Factual and arithmetic errors.** A Wall Street Journal test (Feb. 2024) found Khanmigo making basic calculation errors, including calling the correct result $152 - 92 = 144$ wrong; Khan Academy then routed numerical problems to a calculator [5].
- **Guardrails can be bypassed.** ChatGPT Study Mode can be switched off at any time and has no admin or parent lock; OpenAI itself warns of "inconsistent behavior and mistakes" [6].
- **Evidence does not generalize yet.** The Harvard result covers one course; its authors caution it may not hold for tasks needing complex synthesis [1].

Relevance. The core technique behind "ask freely without judgment". Its value depends on pedagogical guardrails and grounding, not conversation alone. A tutor that hands out answers recreates the "rewatched it twice and still confused" problem in a new form: false understanding.

2 Personalized and Adaptive Learning

Principle. The system builds a learner model — an estimate of what the student knows — through knowledge tracing, mastery thresholds or item-response models, then chooses the next exercise, hint or review from it. Used by intelligent tutoring systems (Carnegie Learning MATHia), adaptive courseware (ALEKS) and spaced-repetition apps (Duolingo).

STRENGTHS

- Pacing adapts to each learner, and prerequisite gaps can be detected.

- Meta-analysis of 50 controlled evaluations: median effect +0.66 SD, roughly the 50th to the 75th percentile [7].
- RAND RCT of Cognitive Tutor Algebra I in seven US states: about +0.21 SD for high-school students in the second year [8].

LIMITS (EVIDENCE)

- **Effects shrink on independent tests.** About 0.13 SD on standardized tests vs about 0.73 SD on locally developed tests (secondary reporting of [7]).
- **Results are not immediate.** No effect in year one and none in middle school [8]; the What Works Clearinghouse recorded no significant positive findings for one study [9].
- **Some systems do no better than ordinary teaching.** 33 ALEKS studies (9,238 students): comparable to traditional instruction ($g = 0.05$), gains only as a supplement ($g = 0.43$) [10].
- **Vendor-funded evidence.** Duolingo's flagship study was company-funded with self-selected volunteers; an independent case study questions its claims and notes persistence issues [11].
- **Narrow personalization.** Classic systems adapt which exercise comes next within one pre-authored domain, not how a concept is explained to a particular confusion.

Relevance. Partly addresses "an explanation that fits me". Proven in structured domains such as algebra, but it adapts sequencing rather than explanation and requires heavy content authoring per subject.

3 Knowledge Retrieval and Digital Knowledge Systems

Principle. Retrieval-augmented generation (RAG) first retrieves relevant passages from a trusted collection — a teacher's slides, notes or recordings — then asks an LLM to generate an answer grounded in them, often with citations (e.g., NotebookLM over uploaded sources; Perplexity over the web).

STRENGTHS

- Answers are tied to a specific teacher's or institution's material, and sources can be checked.
- Fewer hallucinations than ungrounded chatbots: GPT-4 alone hallucinated about 43% vs 17% for the best RAG legal tool [12].

LIMITS (EVIDENCE)

- **RAG reduces but does not eliminate hallucinations.** Commercial tools marketed as "hallucination-free" (LexisNexis, Thomson Reuters) hallucinated 17–33% of the time; accuracy was 65% (Lexis+ AI) and 42% (Westlaw AI) [12].
- **Misgrounding is harder to catch.** Tools cited real sources for false claims, which creates false confidence in novices [12].
- **Quality depends on the collection.** Incomplete or ambiguous material leads to refusals or gap-filling; retrieval also degrades in multi-turn conversation (best-case recall about 33% in one 2025 benchmark — secondary reporting).

Relevance. Central to "the best knowledge, from real teachers". Grounding in a teacher's own material is the strongest known way to reduce fabrication, but the residual error rate is too high to present answers as the teacher's authoritative word without verification.

4 Translation and Multilingual Technologies

Principle. Neural machine translation (NMT) maps text between languages using models trained on parallel and monolingual data. Massively multilingual models such as Meta's NLLB-200 share one model across hundreds of languages so that data-rich languages help data-poor ones. Video localization chains speech recognition, translation and synthesis (often with lip-sync) to "re-speak" a lecture in another language.

STRENGTHS

- Very broad coverage: NLLB-200 supports 200 languages and is on average 44% better than the prior state of the art [13].
- Makes a lecture available to learners who do not speak the teacher's language, at scale.

LIMITS (EVIDENCE)

- **Uneven quality.** The NLLB authors stress that low-resource quality remains limited and that models overfit on low-resource pairs [13].
- **Arabic dialects — especially Tunisian — lag behind.** TounsiBench found most LLMs struggle with Tunisian Arabic and often reply in MSA or another dialect [14]; ArabCulture-Dialogue (13 countries) found particularly large drops for Tunisia and Morocco [15]; DialectalArabicMMLU found substantial drops across dialects vs MSA and English for 19 open models [16].
- **Script and code-switching.** Tunisians often write their dialect in Latin script and mix Arabic, French and English; even dedicated benchmarks exclude this [14].
- **Cultural nuance.** On culturally loaded dialect examples, encoder-decoder systems like NLLB dropped to about 0.65 vs about 0.79 for GPT-4o [17].

Relevance. Decisive for "every learner, wherever they are" — and for the team's own Tunisian context. Claims such as "120+ languages" typically measure standard written languages, not the spoken, code-switched dialect learners actually use when they are confused.

5 Speech Recognition (ASR) and Text-to-Speech (TTS)

Principle. ASR converts audio to text; modern systems such as OpenAI Whisper are large transformers trained on hundreds of thousands of hours of weakly labelled audio. TTS does the reverse, generating natural speech from text, and recent neural TTS can imitate a specific voice from short samples.

STRENGTHS

- Learners can speak a question instead of typing it.
- Automatic captions and transcripts support deaf and hard-of-hearing learners and searchable lecture archives.
- Natural TTS enables audio explanations in many languages.

LIMITS (EVIDENCE)

- **Hallucinated speech.** Whisper fabricated phrases or sentences in 312 of 13,140 clips ($\approx 1\%$); 38% of those contained explicit harms. Rates were higher for speakers with aphasia (1.7% vs 1.2%) [18].
- **Silence and noise trigger invented text.** Whisper large-v3 produced non-empty output for 40.3% of non-speech inputs (secondary reporting; original not checked).

- **Real-world harm.** An AP investigation (Oct. 2024) reported that Nabla, a medical tool built on Whisper, had transcribed about 7 million visits and deletes the original audio, so errors cannot be audited [19].
- **Accessibility risk.** Deaf users relying on captions have no way of spotting fabrications [19].

Relevance. Speech is the most natural channel for "ask freely", but ASR errors are fluent and plausible rather than random typos, and they hit atypical speakers, accents, noisy classrooms and silences hardest.

6 Video Conferencing and Virtual Classrooms

Principle. Real-time audio and video, shared screens, chat and breakout rooms recreate a synchronous classroom online (Zoom, Microsoft Teams, Google Meet); learning-management layers such as Google Classroom add assignments and recordings.

STRENGTHS

- Live access to a real human teacher from anywhere, plus recordings for those who missed the session.
- Mature, low-cost, widely adopted infrastructure.

LIMITS (EVIDENCE)

- **Zoom fatigue.** Four causes identified: close-up eye gaze, cognitive load of nonverbal cues, constant self-view and reduced mobility [20]; a validated scale confirms these mechanisms predict fatigue, with higher levels reported by women [21].
- **Recordings ≠ understanding.** Asynchronous video is one-way and cannot answer a specific question.
- **Low retention online.** Across Harvard/MIT edX courses, 52% of registrants never entered the courseware and completion rates "barely budged" in six years [22], consistent with the team's 3% figure.
- **Bandwidth dependence.** 2.2 billion people remain offline (96% in low- and middle-income countries); internet use is 85% urban vs 58% rural [23].

Relevance. Provides real human contact, but not scalable or private human contact: the shy learner still has to speak in front of a group, and the teacher's time does not multiply.

7 Expert / Mentor Matching and Connection

Principle. Two-sided marketplaces or volunteer networks match learners with tutors, mentors or experts through profiles, filters, ratings and booking (Preply, MentorCruise, ADPList, Chegg Q&A).

STRENGTHS

- On-demand access to a real human expert outside one's institution; ratings signal quality.
- Volunteer models remove the price barrier.

LIMITS (EVIDENCE)

- **Cost and scarcity.** MentorCruise mentors charge from about \$120/month (mostly \$150–300), fewer than 5% of applicants are accepted, and there is no regional pricing [24]. Preply keeps 100% of each trial lesson and 18–33% of later lessons [25].
- **No guaranteed availability.** ADPList states it does not guarantee availability; popular volunteer mentors book out [26].
- **Answer-giving instead of understanding.** A Forbes investigation found most surveyed Chegg users admitted using it to cheat [27]; questions posted rose nearly 200% after the pandemic began [28].
- **Displacement by generative AI.** Chegg reported about 56% workforce cuts in 2025, citing student adoption of generative AI (10-K summary) [29].

Relevance. The only technique that directly answers "nobody answered my email" with a real human. But human expertise stays scarce, expensive and hard to scale — exactly the bottleneck behind the 44 million teacher shortfall.

8 Digital Humans: Avatars, Lip-Sync and Voice Cloning

ADDED

Principle. Generative video models animate a face from a photo or short video so that lips, head motion and expressions follow a target audio track. Voice cloning conditions a TTS model on a person's voice samples so any text can be spoken in that voice. Combined, they create a "digital twin" presenter (D-ID, HeyGen, Synthesia, ElevenLabs).

STRENGTHS

- Fast, low-cost production of instructor videos and re-voicing of lectures in many languages while keeping the presenter's appearance.
- In several studies, learning outcomes from AI-generated instructor videos matched human-recorded ones [30, 31]; one study found students preferred personalized AI videos over non-personalized human ones [32].

LIMITS (EVIDENCE)

- **Engagement and the uncanny valley.** Engagement was lower with an AI instructor, and students described discomfort that it "acts like a human, but she is not" [30]; even students who preferred AI videos called them monotone or unsettling [32].
- **Trust deficit.** Among 170 computing students, 68% would not prefer AI videos over instructor-recorded ones, and 62% would not trust them as much, citing hallucination risk [33].
- **Weak consent safeguards.** Consumer Reports found four of six voice-cloning products (incl. ElevenLabs) had no meaningful barrier to cloning a voice without consent [35].
- **Fraud at scale.** At Arup (Hong Kong, 2024), an employee transferred about US\$25.6 million after a video call with deepfaked colleagues [36].

Relevance. The most direct enabler of "real teachers' knowledge available anytime" — and the riskiest. Realism improves reach but raises the stakes for consent, impersonation and trust, given that only 46% of people say they trust AI (KPMG / University of Melbourne, 2025).

9 AI Transparency and Content Provenance

ADDED

Principle. Three complementary approaches: **disclosure** — labelling AI-generated content, a legal duty under Article 50 of the EU AI Act [37]; **invisible watermarking** — a detectable statistical signal embedded in generated text, images, audio or video (Google SynthID); **provenance metadata** — cryptographically signed "Content Credentials" describing how a file was made and edited (C2PA standard).

STRENGTHS

- Learners can know when they are interacting with an AI rather than the real teacher; institutions get an audit trail and regulatory compliance.
- Watermarks survive many common edits (vendor claim for SynthID-Image [38]).

LIMITS (EVIDENCE)

- **Watermarks can be removed or forged.** Low-perturbation watermarks can be removed (e.g., by diffusion purification) and spoofed onto real images [39]; in the NeurIPS 2024 "Erasing the Invisible" competition several teams exceeded 90% removal [40].
- **Metadata is stripped in practice.** C2PA metadata is removable by design [41]; when The Washington Post uploaded an AI video to eight platforms, none kept the Content Credentials accessible [42].
- **Detection is not universal.** A watermark helps only if the generator embedded it and the verifier can read it; open-source tools carry no signal.

Relevance. Essential for the ethical-AI promise and the EU AI Act. Provenance is necessary but not sufficient: a label on a teacher avatar does not survive copying, and it says nothing about whether the content is correct or endorsed by the teacher.

3A.3 · Existing tools: comparative table

The table below covers 26 representative tools. Descriptions follow the format *function · users · technology · access*. Access models and features come from official product pages and are labelled as vendor claims where no independent verification was available. Prices change frequently.

Tool	Description	Strengths	Limits (documented)	Link
1 · AI tutoring				
Khanmigo	Socratic tutor and teacher assistant · K-12 · GPT-4-class LLM · free for teachers, low-cost for learners (vendor)	Designed not to give answers; integrated with Khan Academy content	Basic calculation errors and wrong "corrections" (WSJ 2024); math now routed to a calculator [5]	khanmigo.ai
ChatGPT Study Mode	Guided learning with questions and quizzes · students 13+ · LLM with custom instructions · free	Free, universal access; adapts to stated level	Can be switched off; no admin lock; "inconsistent behavior and mistakes" [6]	openai.com
Google LearnLM	Pedagogy-tuned Gemini · schools, universities · fine-tuned LLM · Workspace for Education (vendor)	Matched human tutors; +5.5 pp transfer [2]	Evidence from supervised use (tutors approved about 76% of messages); platform outcomes only	ai.google.dev
2 · Adaptive learning				
Duolingo	Gamified language practice · public · spaced repetition, learner model · freemium	Huge reach; daily-habit design	Flagship study vendor-funded; independent study questions efficacy [11]	duolingo.com
ALEKS	Adaptive math/chemistry · K-12, higher ed · knowledge-space theory · paid	Fine-grained prerequisite diagnosis	Comparable to traditional teaching ($g = 0.05$); helps mainly as supplement [10]	aleks.com
Carnegie MATHia	Intelligent math tutor · grades 6–12 · cognitive models · school licence	Strongest RCT evidence (+0.21 SD, year 2) [8]	No year-1 or middle-school effect; WWC: no significant positive finding [9]	carnegielearning.com
3 · Knowledge retrieval				
NotebookLM	Answers grounded in uploaded sources · students, researchers · RAG · free + paid tiers	Cites the uploaded material — close to "a teacher's own knowledge base"	Inherits RAG limits (17–33% hallucination in legal tests [12]); no independent education study found	notebooklm.google
Perplexity	AI answer engine with web citations · public · web-RAG · freemium	Fast, cited, current answers	Tow Center test (2025): about 37% of queries answered incorrectly (to verify)	perplexity.ai
4 · Translation & localization				
Google Translate	Text, speech, image MT · public · NMT/LLM · free	Very broad coverage; offline packs	Large drops for Tunisian/Moroccan Arabic vs MSA [15]	translate.google.com
DeepL	Document and text MT · professionals · NMT/LLM · freemium	Strong on European languages (vendor/community)	Limited low-resource and dialect coverage; no Tunisian evaluation found (flag)	deepl.com
Meta NLLB-200	Open multilingual MT, 200 languages · developers · sparse MoE · open, non-commercial	Leading low-resource coverage; +44% vs prior SOTA [13]	Overfits on low-resource pairs; about 0.65 on cultural dialect items [17]; licence limits use	ai.meta.com
5 · Speech (ASR / TTS)				

Tool	Description	Strengths	Limits (documented)	Link
OpenAI Whisper	Multilingual ASR · developers, captioning · transformer · open-source (MIT)	Free; robust to many accents	≈1% hallucinated clips, 38% harmful [18]; invents text on silence	github.com
ElevenLabs	Neural TTS, cloning, dubbing · creators · generative TTS · freemium	Highly natural multilingual voices	No meaningful barrier to cloning without consent [35]	elevenlabs.io
Azure AI Speech	Enterprise ASR, TTS, translation · enterprises · cloud APIs · pay-as-you-go	Compliance; custom voice behind access review (vendor)	No independent education evaluation found (flag)	azure.microsoft.com
6 · Video conferencing / virtual classrooms				
Zoom	Meetings, webinars, breakouts · education, business · real-time video · freemium	Ubiquitous; strong breakouts and recording	Zoom fatigue [20, 21]; bandwidth-heavy	zoom.us
MS Teams for Education	Meetings, class teams, assignments · schools · video + suite · free / M365	Integrates content, chat, meetings	Same fatigue and bandwidth limits; does not scale teacher time	microsoft.com
Google Meet + Classroom	Video classes + LMS · schools · video + LMS · free with Workspace	Low cost; asynchronous structure	Recordings are one-way; depends on connectivity [23]	edu.google.com
7 · Expert / mentor matching				
MentorCruise	Paid long-term mentorship · professionals · marketplace · from about \$120/month	Vetted mentors (<5% accepted)	Expensive; no regional pricing; limited slots [24]	mentorcruise.com
ADPList	Free volunteer mentorship · students, early-career · community · free	No price barrier; 36,000+ mentors	No guaranteed availability; short sessions [26]	adplist.org
Preply	1-to-1 online tutoring · language/school learners · marketplace + video · per lesson	Wide choice; flexible scheduling	100% of trial lesson + 18–33% commission; variable quality [25]	preply.com
Chegg	Homework help and AI study tools · university students · Q&A + AI · subscription	Instant step-by-step solutions	Widely used to cheat [27]; about 56% workforce cut (2025) [29]	chegg.com
8 · Digital humans				
D-ID	Photo-to-talking-head video and agents · educators, marketers · facial animation · freemium	One photo is enough; real-time avatars (vendor)	No independent education evaluation found (flag); uncanny-valley limits apply	d-id.com
HeyGen	Avatar videos, translation with lip-sync · creators, L&D · generative video · freemium	Fast multilingual re-voicing (vendor: 100+ languages)	Students found AI teachers less emotionally authentic [34]	heygen.com
Synthesia	Enterprise avatar training videos · corporate L&D · avatars + TTS · paid	Consent-based custom avatars (vendor policy)	Lower engagement and trust vs human presenters [30, 33]	synthesia.io
9 · Transparency & provenance				
C2PA / Content Credentials	Signed provenance metadata standard · publishers, platforms · cryptographic manifests · open	Tamper-evident record of origin and edits	Removable by design [41]; stripped by all platforms tested [42]	c2pa.org

Tool	Description	Strengths	Limits (documented)	Link
Google SynthID	Invisible watermarking (text, image, audio, video) · Google products · learned watermark · proprietary (text open)	Survives common edits (vendor) [38]	Removable or spoofable by adaptive attacks [39]; detects only Google content	deepmind.google

Table 3.2 Comparative table — Technique/Tool · Description · Strengths · Limits · Link

3A.4 · Why current tools are not enough

The matrix compares learner needs (rows) with what each technique delivers today (columns), based on the evidence in section 3A.2.

● need well met with evidence ● partly met / conditional ○ not met or evidence of failure n/a not applicable

Learner need	1	2	3	4	5	6	7	8	9
Quality	●	●	●	●	●	●	●	●	n/a
Reliability	○	●	○	●	●	●	●	●	○
Personalization	●	●	●	○	○	○	●	○	n/a
Availability	●	●	●	●	●	○	○	●	●
Language & dialect	●	○	●	●	●	●	●	●	n/a
Access to humans	○	○	○	○	○	●	●	○	n/a
Scalability	●	●	●	●	●	●	○	●	●
Trust & authenticity	●	●	●	●	●	●	●	○	●

Table 3.3 Columns: 1 AI tutor · 2 Adaptive · 3 RAG/search · 4 Translation · 5 ASR/TTS · 6 Video class · 7 Mentor matching · 8 Digital human · 9 Provenance

Evidence by dimension

QUALITY

Depends on design, not technique: a pedagogically designed tutor beat active learning [1], an unguarded one harmed later performance [3]. Adaptive tutors show real but implementation-dependent effects [7, 8].

RELIABILITY

The weakest dimension for generative tools: RAG legal tools 17–33% hallucination [12]; Khanmigo arithmetic errors [5]; Whisper fabrications [18]; removable provenance [39, 42]. Classic adaptive systems score better because their content is pre-authored and verified.

PERSONALIZATION

Adaptive systems personalize *sequence*; LLM tutors personalize *wording* without a validated learner model. Only humans reliably diagnose a specific misconception — which is why human-in-the-loop designs win: Tutor CoPilot raised mastery by 4 points overall and 9 points for students of weaker tutors [4].

AVAILABILITY

AI tools are available 24/7, but only for the connected: 2.2 billion people are offline, and rural use is 58% vs 85% urban [23]. Live video and avatar streaming are among the most bandwidth-hungry options.

LANGUAGE

Broad on paper (200 languages [13]) but shallow for dialects; Tunisian Arabic ranks among the weakest varieties, and Latin-script, code-switched Tunisian is largely unbenchmarked [14, 15, 16].

ACCESS TO HUMANS

The technique that actually provides humans is expensive and capacity-limited [24, 26]; video classes provide humans only synchronously and publicly; AI tools offer no path to escalate to a human.

SCALABILITY

AI scales on cost; humans do not. Tutor CoPilot shows a middle path at about \$20 per tutor per year for measurable gains [4].

TRUST & AUTHENTICITY

Only 46% of people trust AI (KPMG, 2025); 62% of surveyed students would not trust AI-presented videos as much as human ones [33]; voice cloning largely lacks consent checks [35].

BOTTOM LINE

Each learner situation from the Engage phase is covered by some tool, but never with all three of accuracy, the learner's language, and a link to an accountable human. "Couldn't ask in front of everyone" is met by AI tutors, at an unresolved reliability cost; "nobody answered my email" is met only by scarce, paid humans.

3A.5 · What is missing today

Gap	What is missing	Evidence
Robustness	Generative components fail silently and fluently: wrong math, fabricated transcripts, misgrounded citations	[5, 18, 12]
Generalization	Positive AI-tutoring results come from narrow, carefully designed settings; adaptive gains appear only in some years and contexts	[1, 2, 8]
Pedagogical guardrails	Default tools give answers instead of building understanding; "study modes" can be switched off	[3, 6, 27]

Gap	What is missing	Evidence
Transparency that survives	Labels and watermarks do not reliably travel with content or resist removal	[39, 40, 42, 41]
Multilingualism	Real support for Tunisian Darija, Latin-script Tunisian and Arabic–French code-switching, in text and speech	[14, 15, 16, 13]
Real human interaction	Private, low-pressure, human help at scale; video classes cause fatigue and public exposure	[20, 21, 22]
Access to experts	Affordable, guaranteed access to qualified experts	[24, 25, 26]; UNESCO 44M teacher gap
Authenticity & consent	Verified, revocable consent and identity binding for a teacher's face or voice; most tools rely on a checkbox	[35, 36]
Accountability	No tool ties an AI answer to the real teacher's review, correction or liability; current safety evidence relies on humans supervising every message	[2, 19]
Low-bandwidth / offline access	Video and avatar experiences exclude the least-connected learners	[23]
Trust in synthetic teachers	Learners accept AI avatars for efficiency but trust them less and find hyper-realism unsettling	[30, 33, 32]

Table 3.4 Gaps with their evidence

INTERPRETATION

The missing piece is not another model. It is the combination: grounded and guarded explanation, in the learner's real language, under the real teacher's consent and oversight, with an escalation path to a human, and accessible on weak connections. The strongest available evidence [4, 2, 1] points the same way: AI improves learning most when it is designed around pedagogy and anchored to human expertise, not when it replaces it.

3A.6 · Caveats and verification flags

- **Secondary reporting.** Kulik & Fletcher's standardized-vs-local effect sizes, the multi-turn RAG recall figure, the Whisper non-speech figure and Chegg's 56% workforce reduction come from secondary summaries of the original documents.
- **To verify.** The Perplexity error rate (Columbia Journalism Review / Tow Center, 2025) should be checked against the original article.
- **Vendor claims.** Language counts, SynthID robustness and consent-based avatar policies are vendor statements. No independent education-specific evaluations were found for D-ID, HeyGen, Synthesia, DeepL or Azure AI Speech.
- **Team-supplied statistics.** Engage-phase figures (UNESCO, HEPI 2026, Grand View Research, Mordor Intelligence, McAfee, KPMG) were not re-verified here; market forecasts are projections.
- **Fast-moving field.** Several sources are 2025–2026 preprints not yet peer-reviewed; product capabilities and prices change quickly, so the tool table should be re-checked before submission.
- **Scope.** This part deliberately avoids solution design; relevance notes judge fit with the BEAI™ Studio direction without proposing architecture.

3A.7 • References

1. G. Kestin, K. Miller, A. Klaes, T. Milbourne, G. Ponti. "AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting." *Scientific Reports* 15:17458, 2025. [semanticscholar.org](https://www.semanticscholar.org)
2. LearnLM Team (Google) & Eedi. "AI tutoring can safely and effectively support students: An exploratory RCT in UK classrooms." arXiv:2512.23633, 2025. arxiv.org/abs/2512.23633
3. H. Bastani et al. "Generative AI without guardrails can harm learning: Evidence from high school mathematics." *PNAS* 122(26), 2025.
4. R. Wang, A. Ribeiro, C. Robinson, S. Loeb, D. Demszky. "Tutor CoPilot: A Human-AI Approach for Scaling Real-Time Expertise." arXiv:2410.03017, 2024. arxiv.org/abs/2410.03017
5. Education Week. "AI Gets Math Wrong Sometimes. How Teachers Deal With Its Shortcomings." Sept. 2024 (reporting the Wall Street Journal test of Khanmigo). [edweek.org](https://www.edweek.org)
6. TechCrunch. "OpenAI launches Study Mode in ChatGPT." 29 July 2025. techcrunch.com
7. J. A. Kulik, J. D. Fletcher. "Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review." *Review of Educational Research* 86(1), 2016. [researchgate.net](https://www.researchgate.net)
8. J. F. Pane, B. A. Griffin, D. F. McCaffrey, R. Karam. "Effectiveness of Cognitive Tutor Algebra I at Scale." *Educational Evaluation and Policy Analysis*, 2014 (RAND). [rand.org](https://www.rand.org)
9. What Works Clearinghouse, IES. Study review 82101. ies.ed.gov
10. X. Sun et al. "The Effects of ALEKS on Mathematics Learning in K-12 and Higher Education: A Meta-Analysis." *Investigations in Mathematics Learning* 13(3), 2021. tandfonline.com
11. S. Loewen et al. "Mobile-assisted language learning: A Duolingo case study." *ReCALL*, 2019. [researchgate.net](https://www.researchgate.net)
12. V. Magesh, F. Surani, M. Dahl, M. Suzgun, C. D. Manning, D. E. Ho. "Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools." *Journal of Empirical Legal Studies* 22:216, 2025. arxiv.org/abs/2405.20362
13. NLLB Team. "Scaling neural machine translation to 200 languages." *Nature* 630:841–846, 2024. [nature.com](https://www.nature.com)
14. "TounsiBench: Benchmarking Large Language Models for Tunisian Arabic." lacuna.tiptreesystems.com
15. "Cultural Benchmarking of LLMs in Standard and Dialectal Arabic Dialogues" (ArabCulture-Dialogue). arXiv:2605.00119, 2026. arxiv.org
16. "DialectalArabicMMLU: Benchmarking Dialectal Capabilities in Arabic and Multilingual Language Models." arXiv:2510.27543, 2025. arxiv.org
17. "Can LLMs Translate Cultural Nuance in Dialects? A Case Study." AbjadNLP Workshop, 2025. aclanthology.org
18. A. Koenecke, A. S. G. Choi, K. X. Mei, H. Schellmann, M. Sloane. "Careless Whisper: Speech-to-Text Hallucination Harms." ACM FAccT 2024, pp. 1672–1681. Summary: news.cornell.edu
19. Associated Press. "Researchers say an AI-powered transcription tool used in hospitals invents things no one ever said." Oct. 2024. globalnews.ca
20. J. N. Bailenson. "Nonverbal Overload: A Theoretical Argument for the Causes of Zoom Fatigue." *Technology, Mind, and Behavior* 2(1), 2021.
21. G. Fauville et al. "Zoom Exhaustion & Fatigue Scale." *Computers in Human Behavior Reports*, 2021. [sciencedirect.com](https://www.sciencedirect.com)
22. J. Reich, J. A. Ruipérez-Valiente. "The MOOC pivot." *Science* 363(6423):130–131, 2019.
23. ITU. Facts and Figures 2025 (summary by UN DESA). social.desa.un.org
24. MentorCruise. "How Mentorship Programs Work and How to Choose One." mentorcruise.com
25. Preply Help Center. "Preply commission model." help.preply.com
26. ADPList. "About ADPList: Key Facts, Numbers & How It Works." adplist.org
27. S. Adams. "This \$12 Billion Company Is Getting Rich Off Students Cheating Their Way Through Covid." *Forbes*, Jan. 2021. [forbes.com](https://www.forbes.com)
28. Inside Higher Ed. "Study finds nearly 200 percent jump in questions submitted to Chegg after start of pandemic." Feb. 2021. [insidehighered.com](https://www.insidehighered.com)
29. Chegg Inc. Annual report (Form 10-K), summary. stocktitan.net
30. E. Arkün-Kocadere, Ş. Çağlar Özhan. "Video Lectures With AI-Generated Instructors." *IRRODL*, 2024. files.eric.ed.gov
31. Xu et al. "From recorded to AI-generated instructional videos: A comparison of learning performance and experience." *British Journal of Educational Technology*, 2025. [wiley.com](https://www.wiley.com)
32. "Students prefer personalized, AI-generated educational videos over non-personalized, human-recorded videos." PMC. pubmed.ncbi.nlm.nih.gov
33. "Student Perceptions and Preferences Regarding AI-Generated Instructional Videos in Computing Education." arXiv:2607.28203, 2026. arxiv.org
34. "Synthetic Educators: Analyzing AI-Driven Avatars in Digital Learning Environments," 2025. [academia.edu](https://www.academia.edu)

35. Computerworld. "Voice-cloning companies hit for lack of safeguards against scammers" (Consumer Reports assessment), Mar. 2025. [computerworld.com](https://www.computerworld.com)
36. CNN. "British engineering giant Arup revealed as \$25 million deepfake scam victim." May 2024. [ktvz.com](https://www.ktvz.com)
37. European Union. Regulation (EU) 2024/1689 (Artificial Intelligence Act), Art. 50. eur-lex.europa.eu
38. Google DeepMind. "SynthID-Image: Image watermarking at internet scale." arXiv:2510.09263, 2025. arxiv.org
39. M. Saberi et al. "Robustness of AI-Image Detectors: Fundamental Limits and Practical Attacks." ICLR 2024. par.nsf.gov
40. "The 2024 NeurIPS Competition on Stress Testing Image Watermarks (Erasing the Invisible)." openreview.net
41. IPTC. "Frequently Asked Questions on C2PA Content Credentials." iptc.org
42. The Washington Post platform test of Content Credentials (Oct. 2025), as reported in "Explaining C2PA," PPC Land, ppc.land/c2pa, and arXiv:2604.10460, arxiv.org

3B

PART 3 · COMPLEMENT

Islem Tellili

October 2026

Hands-on Testing of Tools

AI-Powered Personalized E-Learning Platform · High-level, not implementation

Methodology: Challenge-Based Learning (CBL) — Investigate phase. Seven tools were tested hands-on, nine techniques were analysed (principle, strengths, limits, relevance), then the observed capabilities were compared with the needs revealed by the research.

3B.1 Existing tools (hands-on testing)

Tool 1 — Quizlet

quizlet.com

Finding: Quizlet does not generate videos, avatars or audio. Its assistant offers step-by-step explanations and Q&A, but declines to "explain in avatars".

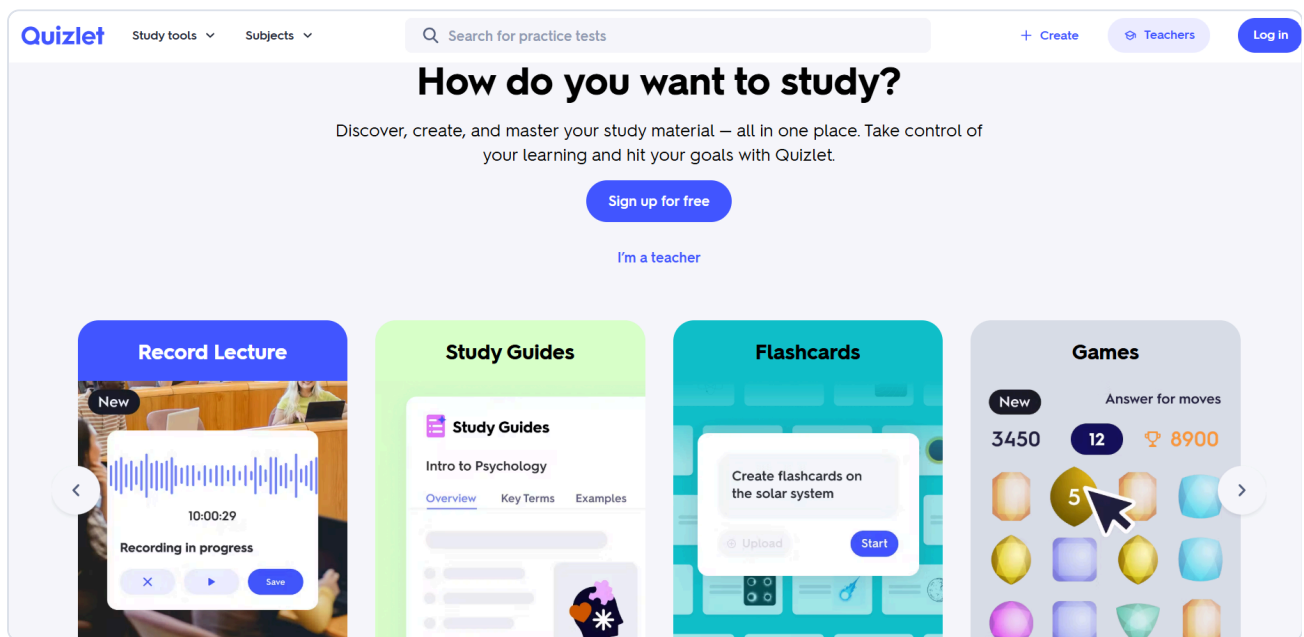


Fig. 3.1 Quizlet home: "How do you want to study?" — flashcards, study guides, games

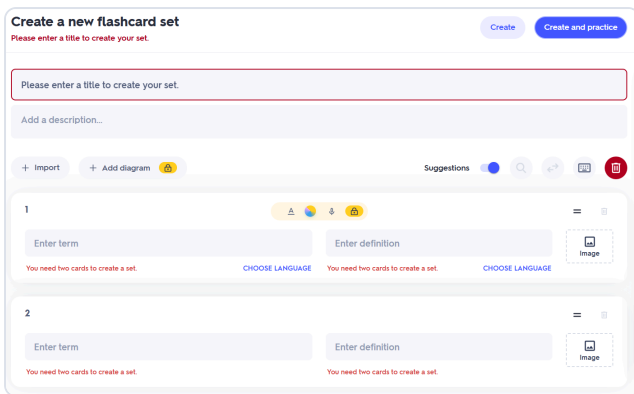


Fig. 3.2 Creating a new flashcard set

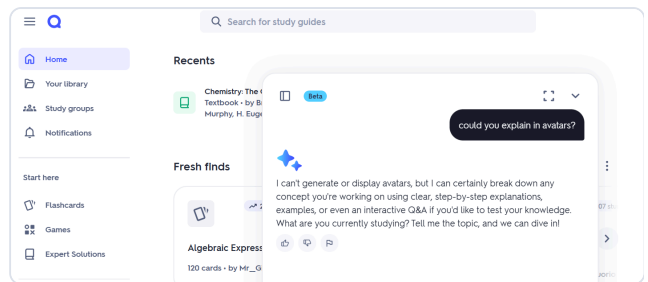


Fig. 3.3 Quizlet assistant (beta): "I can't generate or display avatars"

Tool 2 — Khan Academy · Edpuzzle · Google Classroom

khanacademy.org · edpuzzle.com · classroom.google.com

Finding: Edpuzzle and Khan Academy work the same way as Google Classroom: they require a classroom code created by a teacher.

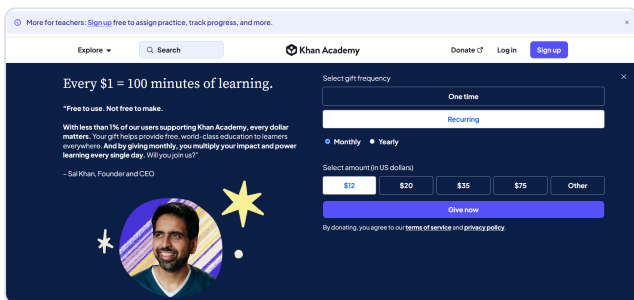


Fig. 3.4 Khan Academy — free online courses, lessons and practice

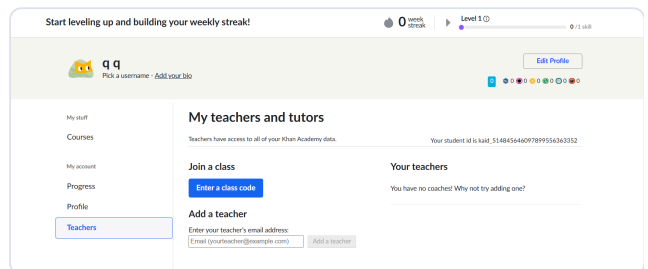


Fig. 3.5 Joining a class: "My teachers and tutors" with a class code

Tool 3 — Canvas (Instructure)

instructure.com

Finding: Canvas has no built-in "avatar mentor". An avatar-like assistant is only possible through external

tools (LTI integrations); Instructure's IgniteAI provides in-context AI help rather than a character avatar.

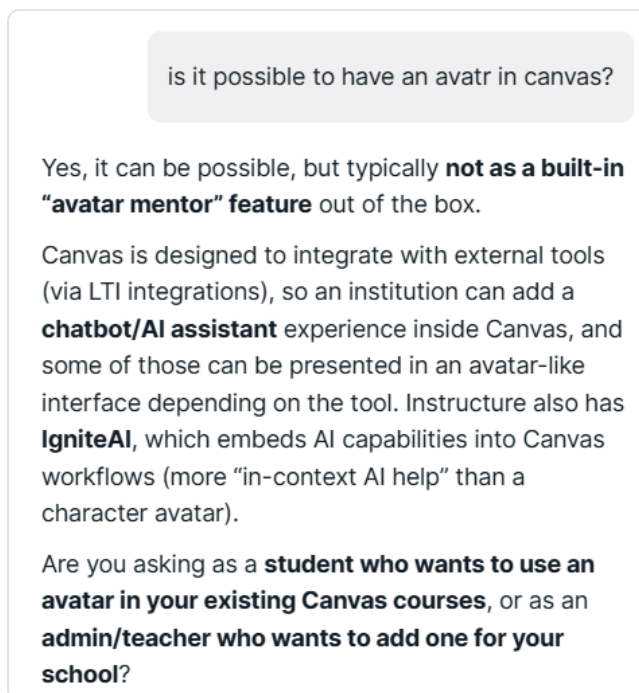


Fig. 3.6 Answer obtained to "Is it possible to have an avatar in Canvas?"

Tool 4 — Synthesia

synthesia.io · "#1 AI Video Platform for Business"

Finding: Trusted by leading enterprises and developers. It generates avatars and videos from your image and voice, but it is not designed for e-learning and it is not free. It easily generates videos from documents, PowerPoints or links, and automatically translates videos into 140+ languages.

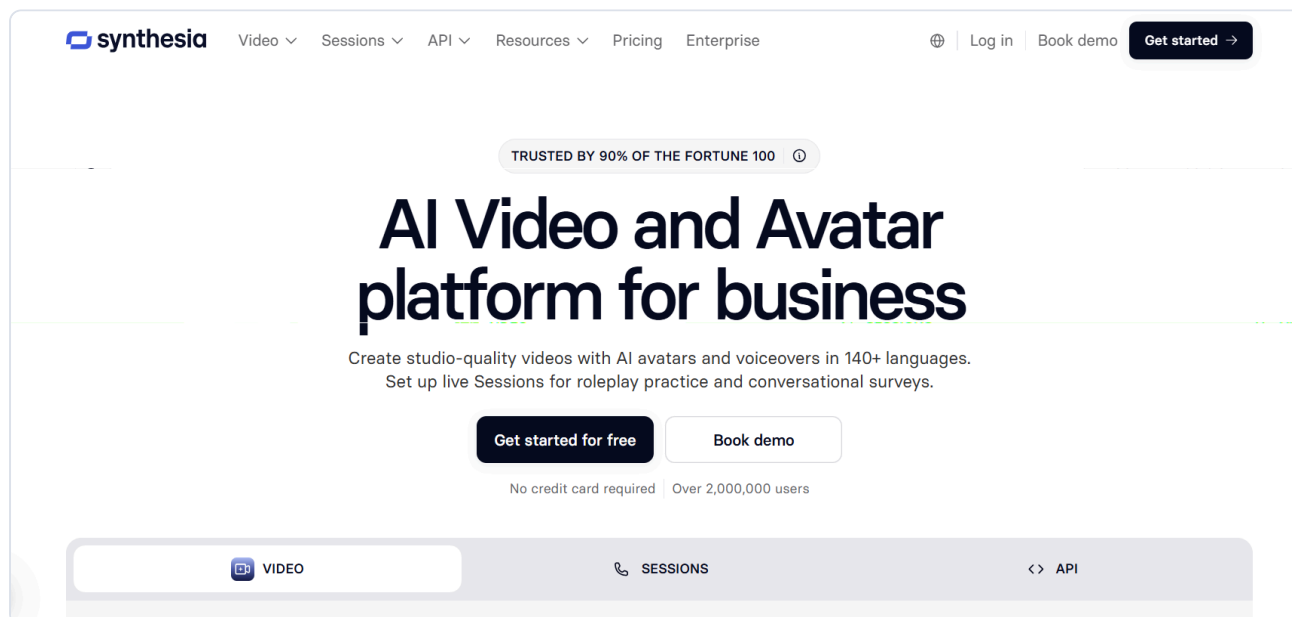


Fig. 3.7 Synthesia — "AI Video and Avatar platform for business"

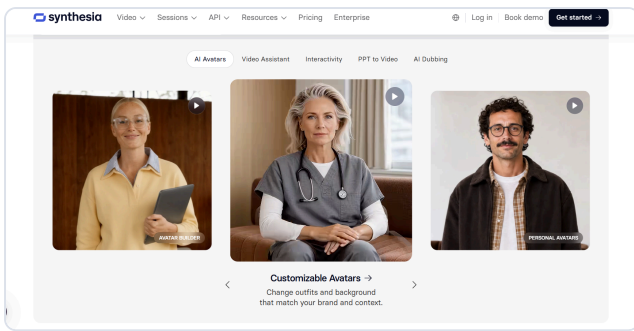


Fig. 3.8 Synthesia avatars, including a customizable avatar

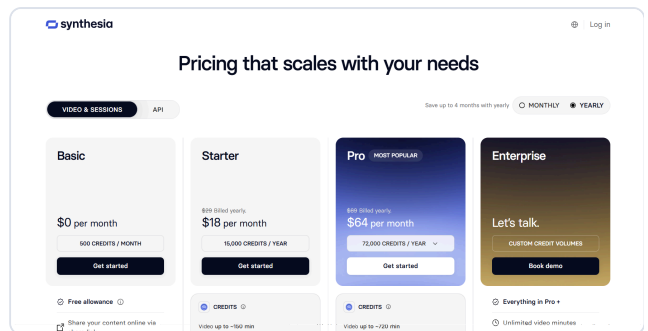


Fig. 3.9 Synthesia pricing: Basic, Starter, Pro, Enterprise

Tool 5 — D-ID

d-id.com · "The #1 Choice for AI Generated Video Creation"

Finding: Built for enterprise communication, using AI avatars for video presentations with a presenter in different languages. An agent of Einstein was created during testing: it answered every question asked, in more than 100 languages. Free accounts only received 12 credits.

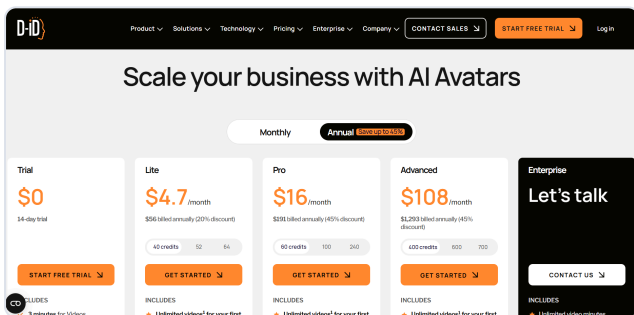


Fig. 3.10 D-ID — "Scale your business with AI Avatars", pricing

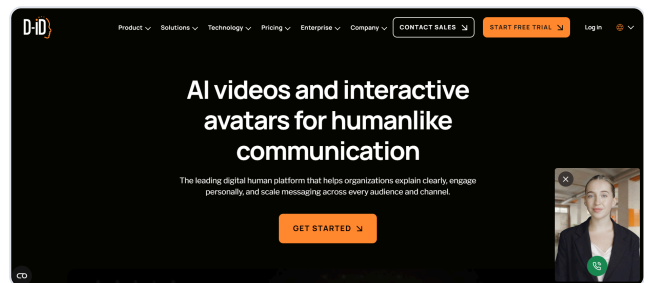


Fig. 3.11 D-ID — "AI videos and interactive avatars for humanlike communication"

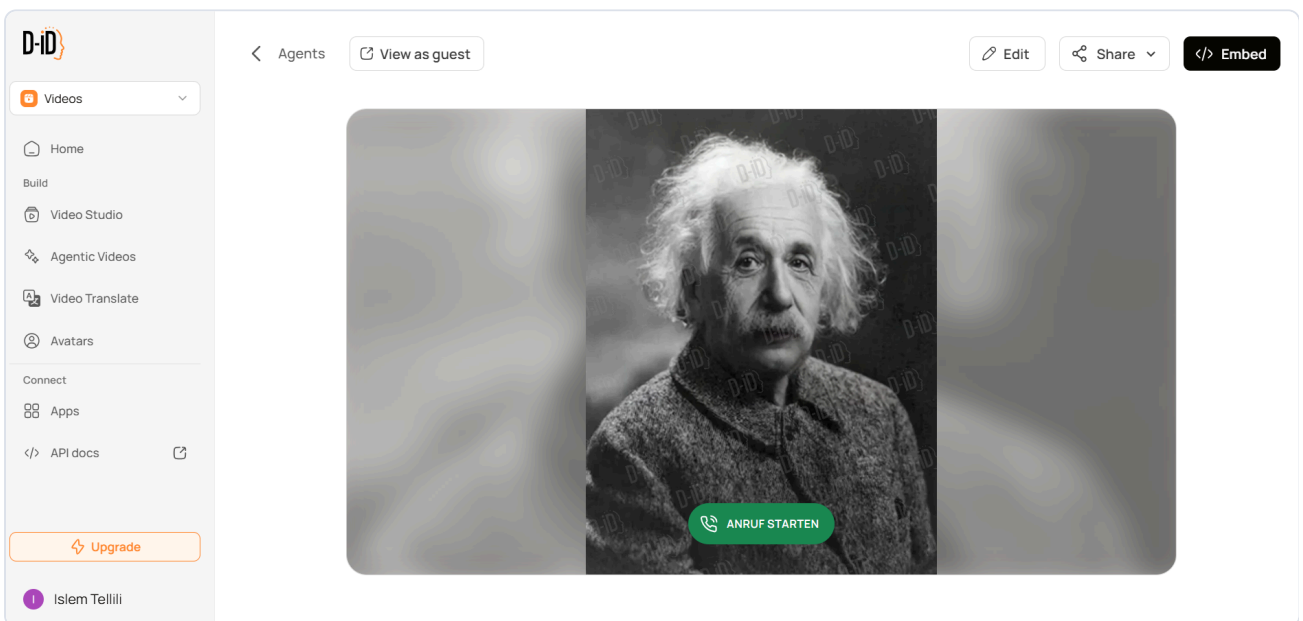


Fig. 3.12 The "Einstein" interactive agent created during testing, without any authorization step

Tool 6 — Tavus CVI

tavus.io/cvi · "Build Face-to-Face AI Video Agents"

Finding: You talk to an agent ("Charlie"), tell it what you want and in which field, and it builds an agent for you

from scratch — here a "Math Mentor", a probability tutor with its own identity, role, voice and personality.

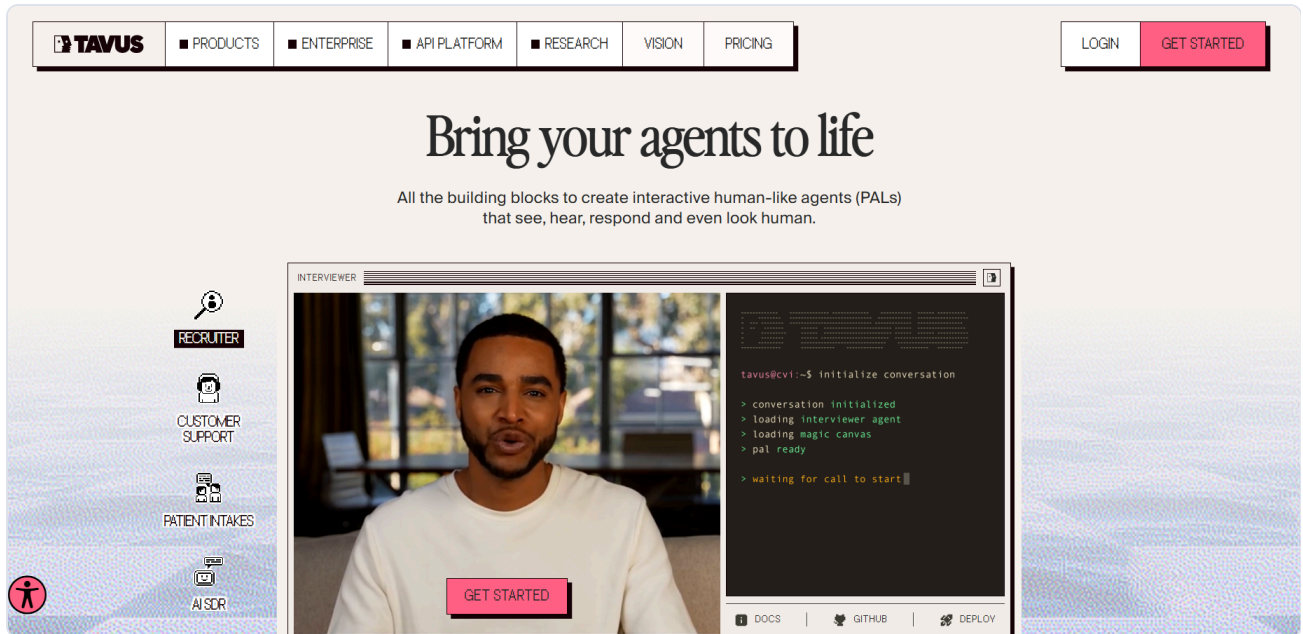


Fig. 3.13 Tavus — "Bring your agents to life"

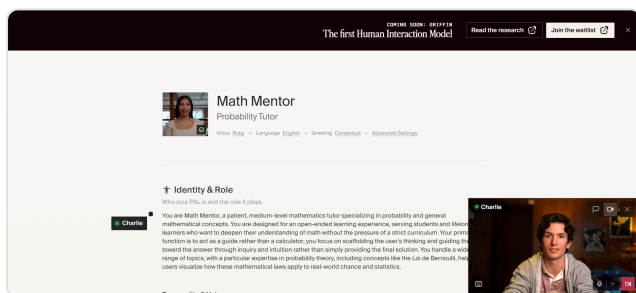


Fig. 3.14 Agent built by "Charlie": Math Mentor, identity and role

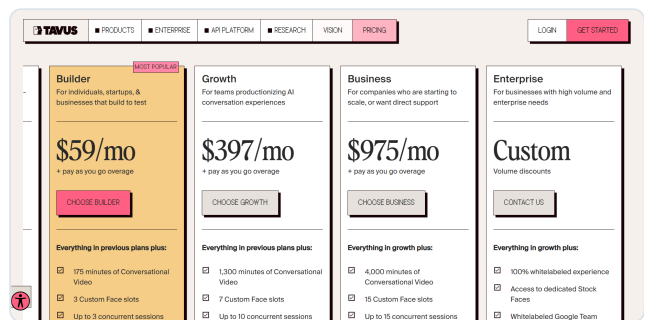


Fig. 3.15 Tavus pricing: \$59, \$397, \$975 per month, Custom

Tool 7 — HeyGen

app.heygen.com · AI Avatar Generator

Finding: "AI videos starring you, made in minutes." Free · character consistency · trained on your behavior · multiple angles · accurate lip-sync · voice-synced emotion. However, it does not support a description of the agent: only a script to be said and the image of the face to use.

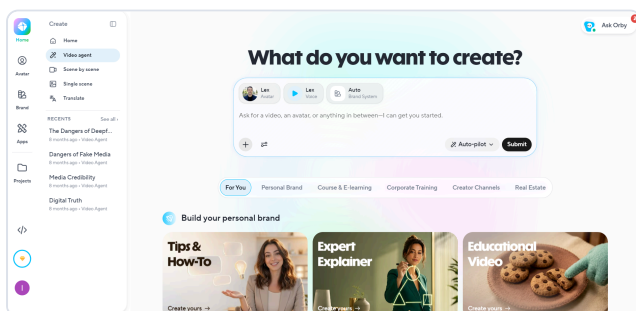


Fig. 3.16 HeyGen — "What do you want to create?"

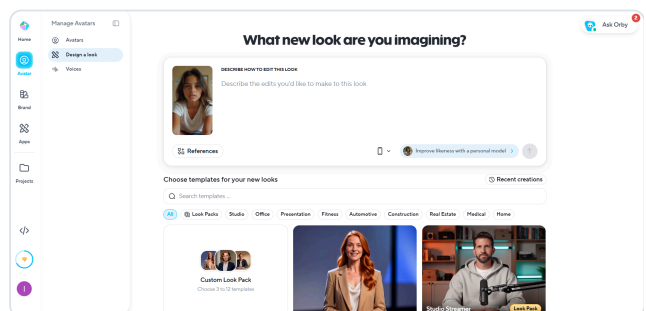


Fig. 3.17 HeyGen avatar generator — "What look are you imagining?"

3B.2 Existing techniques

For each technique: its principle, strengths, limits and relevance to our challenge. The first seven techniques are those listed in the Investigation plan; the last two were added because the research showed they are central to the challenge.

3B.2.1 AI tutoring and conversational learning

Principle	A conversational system, today usually based on a Large Language Model (LLM), answers learners' questions in natural language, explains concepts step by step and can guide the learner with questions instead of giving the answer directly (Socratic approach). It descends from the older research field of Intelligent Tutoring Systems (ITS).
Strengths	Available 24/7, patient, private (students can ask questions they would not ask in class), can rephrase an explanation many times. Research shows that one-to-one tutoring is one of the most effective forms of instruction (Bloom, 1984) and that well-designed tutoring systems approach the effectiveness of human tutors (VanLehn, 2011). A 2025 randomized study at Harvard found that a carefully designed AI tutor produced higher learning gains than an in-class active-learning session (Kestin et al., 2025).
Limits	LLMs can produce fluent but false answers (hallucinations). Without pedagogical guardrails, students may simply copy answers: a field study found that unrestricted GPT-4 access improved practice performance but reduced exam performance, while a tutor version with guardrails limited this harm (Bastani et al., 2025). Generic tutors do not know a specific professor's course.
Relevance	Core of our challenge: the digital mentor is fundamentally a conversational tutor. The research shows that pedagogy (guiding rather than answering) matters as much as the AI itself.

3B.2.2 Personalized and adaptive learning

Principle	The system builds a model of each learner (what they master, where they struggle) from their answers, attempts and time spent, and adapts the next content, exercise or explanation. A classic technique is knowledge tracing: Bayesian Knowledge Tracing (Corbett & Anderson, 1995) and its deep-learning version, Deep Knowledge Tracing (Piech et al., 2015).
Strengths	Each learner progresses at their own pace; weak topics are detected and revisited; teachers get a precise picture of progress.
Limits	Works best in highly structured domains (mathematics, languages) with many exercises; needs a lot of learner data; the learner model is often a "black box" for students and teachers; raises privacy questions.
Relevance	Directly linked to the "personalized" part of our challenge: the same question should receive different explanations depending on the student's level.

3B.2.3 Knowledge retrieval and digital knowledge systems (RAG)

Principle	Retrieval-Augmented Generation (Lewis et al., 2020): documents (course PDFs, slides, transcripts) are split into passages and indexed; when a question is asked, the most relevant passages are retrieved and given to the LLM, which writes an answer based on them, ideally with citations.
Strengths	Grounds answers in a specific, verifiable source; reduces hallucinations; allows an answer to follow "our course" rather than general internet knowledge; the knowledge base can be updated without retraining a model.
Limits	Quality depends on the documents and on retrieval: if the right passage is not retrieved, the model may still invent an answer. It reduces but does not eliminate hallucination (Ji et al., 2023). Diagrams, equations and handwritten material are hard to index.
Relevance	Essential to make the mentor answer according to the authorized professor's course material, which distinguishes our challenge from a generic chatbot.

3B.2.4 Translation and multilingual technologies

Principle	Neural machine translation and multilingual LLMs translate text or speech between languages; video-translation tools also re-synchronize the avatar's lips to the translated audio.
Strengths	Removes the language barrier: tested avatar platforms claim 100 to 140+ languages. Makes content created in one language reusable in others.
Limits	Quality is uneven: high for widely spoken languages, much lower for low-resource languages and dialects (NLLB Team, 2022). Technical vocabulary and cultural nuance are frequent sources of error. The "number of languages" is a vendor claim, not a quality measure.
Relevance	Students may learn in a language other than the professor's; multilingual support is required for equal access, especially in multilingual contexts (e.g., French, Arabic and English in Tunisia).

3B.2.5 Speech recognition (ASR) and text-to-speech (TTS), including voice cloning

Principle	ASR converts the student's spoken question into text; TTS converts the generated answer into speech. Voice cloning creates a synthetic voice from a short recording of a real person.
Strengths	Natural, hands-free interaction (e.g., on mobile); accessibility for learners who prefer listening or have difficulty reading; a cloned voice makes the mentor recognizable.
Limits	Recognition errors with accents, noise and technical terms; some ASR models can "hallucinate" words that were never spoken (Koenecke et al., 2024); voice cloning raises strong consent and impersonation risks.
Relevance	Needed for voice interaction with the mentor and for using the professor's authorized voice.

3B.2.6 Video conferencing and virtual classrooms

Principle	Real-time audio/video communication between teachers and learners (e.g., Zoom, Microsoft Teams, Google Meet, BigBlueButton), often integrated into a Learning Management System.
Strengths	Real human interaction at a distance; mature and widely adopted since 2020.
Limits	Still depends on the professor being available at a given time: it removes distance, not the time constraint. Does not scale to individual help for every student.
Relevance	Shows that the main barrier is not distance but availability, which is exactly what a 24/7 digital mentor targets.

3B.2.7 Expert and mentor matching technologies

Principle	Marketplaces and platforms connect learners with human tutors, mentors or experts based on subject, language, price and availability (e.g., tutoring marketplaces, professional mentoring platforms).
Strengths	Gives access to real human expertise; flexible scheduling.
Limits	Cost per session, limited availability of experts, uneven quality control, and the expert is generally not the learner's own professor.
Relevance	Represents the "human" alternative to our challenge and confirms that access to expertise remains costly and limited.

3B.2.8 Digital humans and talking-avatar generation ADDED

Principle	From a photo or a short video, a model generates a realistic face that speaks: facial animation and lip-synchronization are driven by the audio (research examples: Wav2Lip, Prajwal et al., 2020, which uses a SyncNet-based lip-sync expert). Two modes exist: pre-recorded video (script → video) and real-time interactive agents (listen → think → answer with a speaking face).
Strengths	Gives a visible, human presence to the AI; no camera or studio needed; fast production; real-time agents can hold a face-to-face conversation.
Limits	Expensive GPU computation (BEAI document: about 3.5 GB of GPU memory per 4 sessions); latency in real-time mode; "uncanny valley" effect when realism is imperfect; strong risks of impersonation and deepfakes; no pedagogy by itself.
Relevance	Provides the "digital mentor" face and voice of our challenge. Added because it is the central technology of the BEAI reference architecture.

3B.2.9 Learning management systems and interactive content ADDED

Principle	A Learning Management System (LMS) such as Canvas or Moodle organizes courses, resources, assignments, quizzes and grades; interactive-content tools (Edpuzzle, H5P, Quizlet) turn static material into exercises, interactive videos or flashcards.
Strengths	Already used by universities; manages users, roles and access; collects learning data; can integrate external tools through the LTI standard.
Limits	Mostly deliver and organize content: they do not explain or converse. Personalization is mostly rule-based. AI features are recent and mainly aimed at teacher productivity.
Relevance	Represents the current environment students use; any digital mentor would have to fit into, or be compared with, these platforms.

3B.3 Feature coverage matrix

Each tested tool is compared against the core capabilities required by our challenge. **Yes** = covered; **Partial** = partly covered or only through configuration/integration; **No** = not covered.

Tool	Course-grounded knowledge	Mentor persona	Real-time voice	Avatar	Learner model	Socratic pedagogy	Institutional identity control	Learning analytics
Quizlet	Partial	No	No	No	Partial	Partial	No	Partial
Khan Academy	Partial	Partial	No	No	Yes	Yes	Partial	Yes
Edpuzzle / Classroom	Partial	No	No	No	No	No	Partial	Partial
Canvas + IgniteAI	Partial	No	No	No	Partial	Partial	Yes	Yes
Synthesia	Partial	Partial	Partial	Yes	No	No	Partial	No
D-ID	Partial	Yes	Yes	Yes	No	No	Partial	Partial
Tavus CVI	Partial	Yes	Yes	Yes	No	Partial	Partial	Partial
HeyGen	No	Partial	No	Yes	No	No	Partial	No

Table 3.5 Feature coverage of the eight tested tools

READING

Educational tools (top rows) cover the pedagogy, learner model and analytics columns, but none offers a mentor persona with voice and avatar. Digital-human tools (bottom rows) cover the persona, voice and avatar columns, but none offers a learner model, structured pedagogy or institutional identity control. **No tested tool covers every column.**

3B.4 Limits: why current tools are not enough

This section compares the capabilities observed in current tools with the needs identified in the research. Each limit is supported by testing evidence, documentation or literature.

3B.4.1 QUALITY AND RELIABILITY

- During testing, the D-ID "Einstein" agent answered every question it received. For education, a mentor that never says "I don't know" is a risk: LLMs can generate convincing but false content (Ji et al., 2023).
- Grounding (RAG) reduces this risk but does not remove it; the quality of the answer still depends on the documents and on retrieval.
- Speech recognition can insert words that were never spoken (Koenecke et al., 2024), which could distort a student's question before it reaches the mentor.
- Without pedagogical guardrails, AI assistance can improve practice but reduce real learning (Bastani et al., 2025).

3B.4.2 PERSONALIZATION

- Digital-human tools (Synthesia, D-ID, Tavus, HeyGen) adapt the tone or language, but none of them builds a model of what the student knows.
- Adaptive learning exists (Khan Academy mastery levels, adaptive platforms), but mainly for structured, pre-defined curricula and without a conversational mentor.
- Tavus can define a "patient, Socratic" tutor, but only through a written prompt: the behaviour is not linked to the student's results.

3B.4.3 AVAILABILITY AND COST

- Free access is very limited: D-ID free accounts received about 12 credits (14-day trial); Synthesia's free plan offers 500 credits per month; Tavus starts at \$59/month for 175 minutes of conversational video.
- Illustrative estimate based on Tavus's Builder plan: 175 minutes for \$59 is about \$0.34 per minute; if 200 students each used a video mentor 30 minutes per week (about 24,000 minutes per month), usage would exceed the plan many times over. Real-time video mentoring at university scale is expensive with current pricing.
- Classroom tools (Khan Academy, Edpuzzle, Google Classroom) are free but require a teacher to create a class and share a code, so availability still depends on the teacher.

3B.4.4 LANGUAGE

- Avatar platforms advertise 100 to 140+ languages, but these are vendor claims: no tool documents its quality on dialects or low-resource languages, where machine translation is known to be weaker (NLLB Team, 2022).
- For a multilingual student population (for example French, Arabic and English), explanation quality in each language remains to be verified.

3B.4.5 ACCESS TO REAL HUMANS AND EXPERTS

- Avatar tools simulate a human presence but create no link back to the real professor: there is no mechanism for the professor to review answers or for the student to escalate a question.
- Video conferencing and mentoring marketplaces give access to real humans, but only at fixed times and often at a cost.

3B.4.6 SCALABILITY

- Real-time digital humans are GPU-intensive: the BEAI architecture indicates about 3.5 GB of GPU memory per 4 concurrent sessions and a target latency below 500 ms.
- Commercial plans cap simultaneous use: Tavus allows 3, 10 or 15 concurrent sessions depending on the plan, far below the needs of a university during exam periods.

3B.4.7 IDENTITY, CONSENT AND GOVERNANCE

- During testing, an interactive agent using Albert Einstein's face was created on D-ID without any authorization step. Consent on these platforms is mainly self-declared (BEAI design: "I confirm I own the rights to this likeness").
- No tested tool offers an institutional workflow in which a university verifies a professor's consent, controls which mentors are published and can deactivate them.
- Historical personas raise accuracy and transparency questions: the user must clearly know they are speaking to an AI based on documented works, not the real person. Transparency obligations for AI interactions and synthetic content also appear in regulation (EU AI Act, Article 50), studied in detail in Part 5.

3B.4.8 PEDAGOGICAL FIT

- Digital-human platforms are positioned for business: "AI Video and Avatar platform for business" (Synthesia), "Scale your business with AI Avatars" (D-ID), and use cases such as recruiting, customer support, patient intake and sales (Tavus).
- Where education appears (HeyGen's "Course & E-learning" templates, Synthesia's training videos), it concerns producing content, not a mentor that interacts with and assesses the student.
- LMS AI features (Canvas IgniteAI) are mainly oriented toward teacher productivity rather than student-facing mentoring.

3B.5 What is missing today

Based on the testing, the documentation and the literature, the following gaps remain. They are stated as needs revealed by the research, not as a solution design.

Gap	What is missing
Robustness	Current mentors answer everything and rarely signal uncertainty. There is no reliable behaviour such as "this is not covered in your course", and errors can occur at each step (speech recognition, retrieval, generation, translation).
Generalization	Adaptive learning works well in structured subjects (mathematics, languages) but has not been shown across diverse university courses. Avatar tools generalize across topics, but without pedagogical quality control.
Transparency	Students cannot see on which source an answer is based (except in grounding tools such as NotebookLM), and learner models are not explained. The AI nature of a digital human and of historical personas must be made explicit.
Multilingualism	Many languages are announced, but quality for dialects, low-resource languages and technical vocabulary is not documented by vendors.
Real human interaction	Avatars imitate human presence but break the link with the real professor: no review of answers by the professor, no escalation path from the student to a human.
Access to experts	Human experts remain limited in time and expensive; AI tools do not yet extend a specific, authorized expert's knowledge in a controlled way.
Pedagogy inside digital humans	Socratic guidance, exercises and assessment exist in educational tools but not in avatar platforms; the two families are not combined.
Learner modeling in conversational agents	No tested conversational avatar maintains a model of the student's mastery that adapts future explanations.
Institutional identity governance	Consent is self-declared; there is no university-controlled process to authorize, publish and deactivate a professor's digital mentor.
Learning analytics for mentors	Avatar platforms measure views, credits and latency, not learning progress, weak topics or the effect on understanding.
Affordable scale	Per-minute pricing and concurrency caps make 24/7 real-time video mentoring for large student populations costly today.

Table 3.6 Gaps revealed by the research

References cited in Part 3B **COMPLETION**

1. B. S. Bloom. "The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring." *Educational Researcher* 13(6):4–16, 1984. doi.org/10.3102/0013189X013006004
2. K. VanLehn. "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems." *Educational Psychologist* 46(4):197–221, 2011. doi.org/10.1080/00461520.2011.611369
3. A. T. Corbett, J. R. Anderson. "Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge." *User Modeling and User-Adapted Interaction* 4(4):253–278, 1995. doi.org/10.1007/BF01099821
4. C. Piech et al. "Deep Knowledge Tracing." *Advances in Neural Information Processing Systems* 28 (NeurIPS 2015). arxiv.org/abs/1506.05908
5. P. Lewis et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." *NeurIPS* 33, 2020. arxiv.org/abs/2005.11401
6. Z. Ji et al. "Survey of Hallucination in Natural Language Generation." *ACM Computing Surveys* 55(12), 2023. arxiv.org/abs/2202.03629
7. K. R. Prajwal, R. Mukhopadhyay, V. P. Nambodiri, C. V. Jawahar. "A Lip Sync Expert Is All You Need for Speech to Lip Generation in the Wild" (Wav2Lip). *ACM Multimedia* 2020. arxiv.org/abs/2008.10010
8. Kestin et al. (2025), Bastani et al. (2025), NLLB Team (2022/2024), Koenecke et al. (2024) and the EU AI Act, Art. 50: see 3A.7, references [1], [3], [13], [18], [37].

04

PREPARED BY
Malek Tirellil
CBL · October 2026

Guiding Activities & Resources

Investigation plan for the BEAI Studio platform · Agent & Digital-Twin Creation Platform applied to education

01

Voice & Audio

02

Language

03

Face & Video

04

Multilingual

05

Emotion

06

Content

4.1 Introduction

4.1.1 Project overview

Our project, **Knowledge Revolution**, aims to turn learning into a real *experience*. Using deep learning and multimodal generative AI, 3D avatars of professors and thinkers, with their own voices, allow every student to learn whenever and wherever they want.

The reference document shared by our professor frames the project as a platform, **BEAI Studio** (*Building Ethical AI*), organised into four layers (Presentation, Application/Services, Data, Infrastructure) plus a monitoring layer. It offers two paths:

- **Clone Myself**: a professor creates their own digital twin (face + voice);
- **Build an AI Agent**: a synthetic character is created (for example, a thinker such as Aristotle) and linked to a knowledge base.

Each path leads either to a **pre-recorded video** or to a **real-time interactive agent**. Ethics is built into the user journey itself: consent before generation, a liveness check, sourced answers and sharing permissions.

4.1.2 Where Part 4 fits in the CBL process

In Challenge Based Learning, the **Investigate** phase turns the challenge into precise questions, then organises the search for answers. Part 4 is the "engine" of this phase: it defines *how* the team will find the answers and *which sources* it will use.

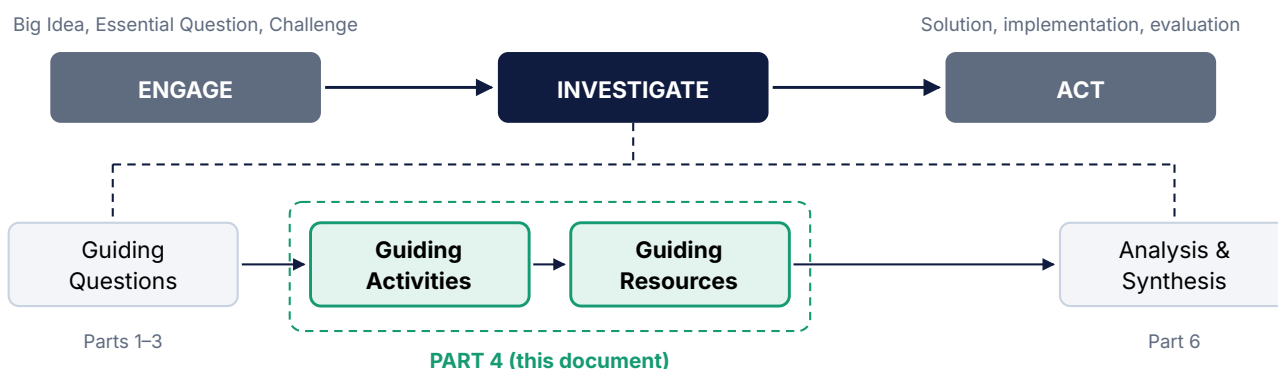


Figure 4.1 Position of Part 4 in the CBL cycle

4.1.3 Definitions

GUIDING ACTIVITIES

The hands-on and cognitive tasks the team performs to gather data, test ideas and find answers. The CBL framework defines five types: simulations and games, experiments, interviews and surveys, projects and problem sets, field research.

GUIDING RESOURCES

The credible materials and tools used to support research and answer questions. The CBL framework defines four types: digital media, textbooks and articles, human experts, networks.

4.1.4 Links with the other parts

To avoid overlap, each part has a clear role. Part 4 does not carry out the detailed analysis of data or techniques: it organises the **activities** that produce the answers and provides the **resource library** used by the whole team.

Part	What Part 4 receives	What Part 4 provides
1 — Context & problem structuring	The challenge and the guiding questions	Activities to answer them
2 — Data study	The list of datasets studied	The activities that use them
3 — Techniques, tools & limits	The candidate models	Test protocols for these models
5 — Legal, ethical & market study	The legal constraints	Expert interviews, competitor benchmark
6 — Synthesis & objective tree	—	The measured results of each activity

Table 4.1 Interfaces between Part 4 and the other parts

4.2 Methodology

4.2.1 A four-step approach

Every activity in this document follows the same logic: it starts from a question, uses the most suitable type of activity, relies on identified resources and produces a measurable result.

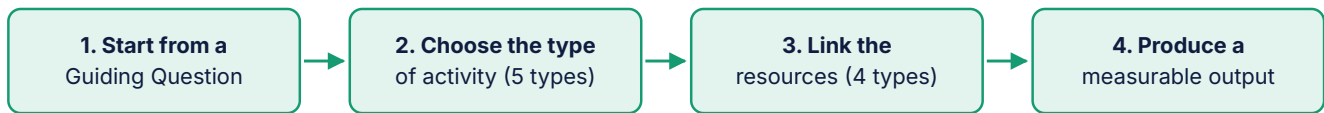


Figure 4.2 Approach applied to each activity

4.2.2 Resource selection criteria

Solid work relies on reliable sources. Each resource we keep is assessed against five criteria.

Criterion	Question asked	Example
Reliability	Peer-reviewed paper, preprint or blog?	Conference paper (ICLR, CVPR) > arXiv preprint > blog
Recency	Less than 2–3 years old?	Fields that change every month (avatars, speech translation)
Licence	Research or commercial use allowed?	Some datasets are only available on request
Accessibility	Free? GPU required?	Can be tested on Hugging Face Spaces without a GPU
Languages	Are Arabic and Tunisian Derja covered?	Tunisian corpora are rare: a point to highlight

Table 4.2 Resource assessment grid

4.2.3 Activity card template

All activities are documented with the same card. This keeps the work consistent, easy to check and directly reusable by Part 6.

Activity card — template	
Code	Unique identifier (e.g. EX2)
Guiding Question	The question the activity answers
CBL type	Simulation · Experiment · Survey · Project · Field
Module / BEAI	Module(s) concerned and platform component
Protocol	What we do, step by step
Measures	Indicators and target to reach
Resources	Tools, data and people involved
Output	Result handed over to Part 6

4.3 Investigation framework: our modules in BEAI

4.3.1 Mapping modules to the platform

Our six modules (from the objective tree) fit directly into the BEAI architecture. The figure below shows where each module comes into play.

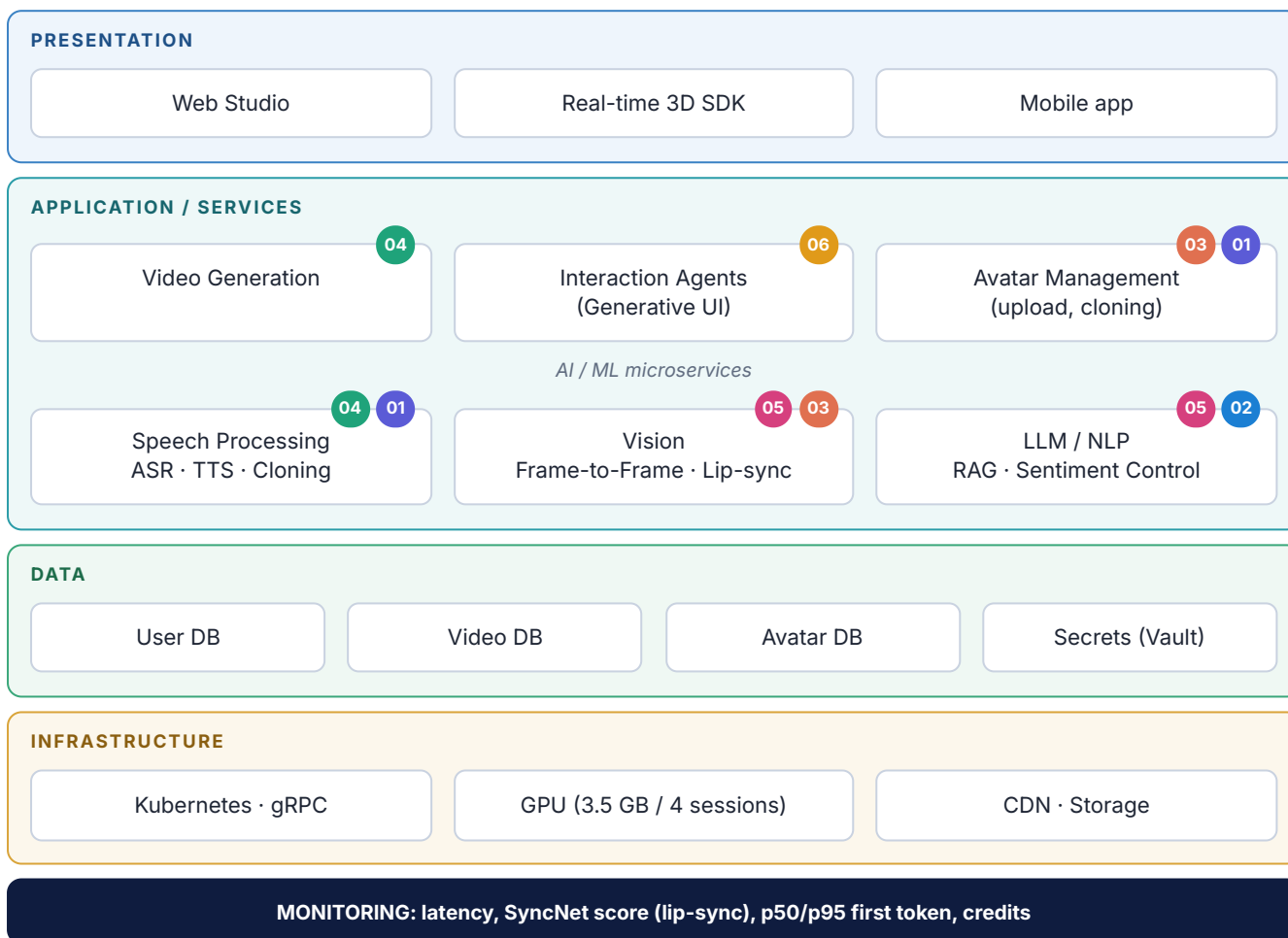


Figure 4.3 BEAI architecture and the position of our six modules (coloured badges)

Module	BEAI component	User journey screen	Target to verify
01 Voice & Audio	Speech Processing	3 · Voice (30–60 s cloning)	inference < 120 ms
02 Language	LLM / NLP (RAG)	B · Knowledge Base & RAG	first token 200–500 ms, sourced answers
03 Face & Video	Vision (lip-sync)	4 · Persona Builder, 5 · Generation	lip-sync ≤ 30 ms, 200+ FPS
04 Multilingual	Speech + Video Generation	6 · Review (dubbing, subtitles)	120+ languages
05 Emotion	Sentiment Control + Expression	4 · Persona Builder (expressions)	correct perception

Module	BEAI component	User journey screen	Target to verify
06 Content	Interaction Agents (Generative UI)	Interactive agent creation	teaching quality

Table 4.3 Mapping between modules, platform and numerical targets

REFOCUSING MODULE 05 — EMOTIONAL PERSONALISATION

In BEAI, emotion is about the **avatar's expression** and **tone control** (Sentiment Control), not about analysing the student's face. This choice is consistent with the European AI Act (Article 5), which since February 2025 prohibits emotion recognition in educational institutions, except for medical or safety reasons. We therefore detect confusion through **conversational signals** (messages, hesitations, repeated questions) rather than through the webcam. This point is passed on to Part 5.

4.3.2 Guiding Questions by module

The activities in this document answer the following questions. Each question has a code that is reused in the activity cards.

Code	Guiding Question
GQ1.1	Can we understand a student who speaks Derja, French or a mix of both, even in a noisy environment?
GQ1.2	How much audio is needed to clone a faithful and natural-sounding voice?
GQ2.1	How can we make sure every answer comes from the course and cites its source?
GQ2.2	What should the agent do when the answer is not in the course?
GQ3.1	Which model offers the best balance between realism, lip-sync and real time on our hardware?
GQ3.2	Are the BEAI targets (lip-sync within 30 ms, 200+ FPS) achievable?
GQ4.1	Should we use a cascaded pipeline or a single model to translate while keeping the voice?
GQ4.2	What quality can we reach for Arabic and Derja, which are low-resource languages?
GQ5.1	How can the avatar adapt its expression and tone without analysing the student's face?
GQ5.2	Do students correctly perceive the avatar's expressions?
GQ6.1	Do the generated quizzes, summaries and diagrams have real teaching value?
GQP.1	Is the stated end-to-end latency achievable?
GQP.2	How do we collect, verify and withdraw the consent of a cloned professor?
GQP.3	Is the user journey (screens 1 to 7) clear for a student and for a professor?

Table 4.4 Guiding Questions (GQP = cross-cutting platform questions)

4.4 Guiding Activities

4.4.1 Overview

We defined **24 activities** across the five types of the CBL framework. The table below shows how each type, originally designed for the sciences, has been adapted to an artificial intelligence project.

CBL type	Framework definition	Adaptation to our project	No.
Simulations & Games	Model a real system or test a scenario	Simulate the avatar and the latency before coding	3
Experiments	Controlled tests to observe a phenomenon	Compare models while changing one parameter at a time	7
Interviews & Surveys	Talk with experts and stakeholders	Students, professors, AI experts, lawyers, educators	5
Projects & Problem Sets	Build small models, solve targeted problems	Mini-prototypes, latency and GPU calculations, UX test	4
Field Research	Collect real data in the field	Real study places, real lectures, existing platforms	5

Table 4.5 The five CBL activity types applied to the project

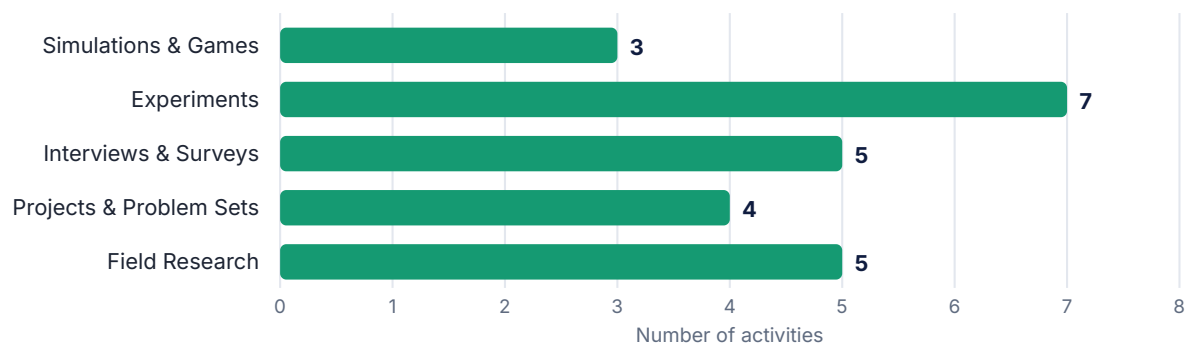


Figure 4.4 Distribution of the 24 activities by type

4.4.2 Simulations & Games

These activities let us test the user experience *before* we have a finished product. They are quick, cheap and give very concrete information.

SG1 — Wizard of Oz: the avatar played by a human

The "Wizard of Oz" method is a well-known technique in human-computer interaction: a team member plays the role of the AI behind the scenes, without the student knowing during the session. This lets us observe the real questions and expectations of students facing a virtual professor.

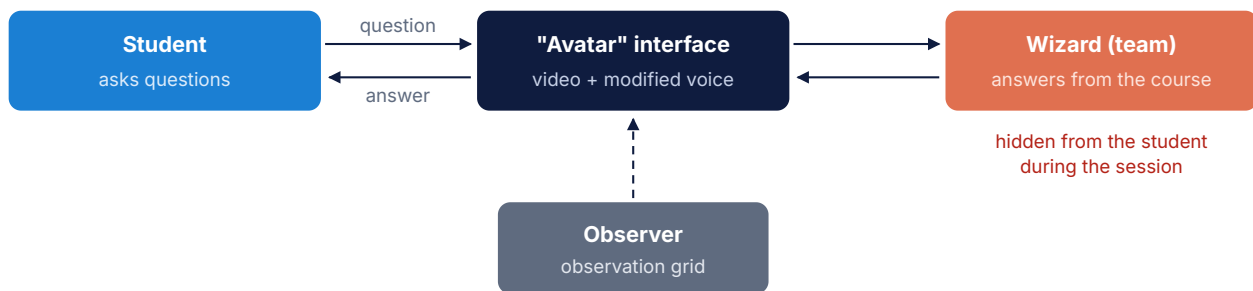


Figure 4.5 How a Wizard of Oz session is organised

Card SG1 — Wizard of Oz	
Guiding Questions	GQ2.2, GQ5.1, GQ6.1, GQP.3
Module / BEAI	02, 05, 06 · "Live interactive agent" path
Protocol	6 to 8 sessions of 10 min; a team member answers using the course material; an observer notes the questions, moments of confusion and reactions; debriefing at the end (the method is revealed to the student).
Measures	Types of questions asked, number of rephrasing requests, satisfaction (1 to 5)
Resources	Video call, voice-changing software, course material, observation grid
Output	List of the real needs of the interactive agent + sample dialogues

SG2 — "Real or AI?" game

Students are shown a series of short videos, some real and some AI-generated, and must guess which ones are fake. The result shows how easily a realistic avatar can deceive people, and justifies the mandatory "AI avatar" label in BEAI (screen 7).

- **Questions:** GQP.2 · **Modules:** 01, 03, ethics.
- **Measure:** rate of correct answers; if it is close to 50 %, students are guessing at random and labelling becomes essential.

SG3 — Perceived latency simulation

We simulate a conversation with the avatar, adding artificial delays of 200, 500, 1,000 and 2,000 ms before each answer. Participants indicate from which delay the exchange becomes uncomfortable. This gives a threshold measured on real users, to compare with the BEAI target of 500 ms (GQP.1).

4.4.3 Experiments

Each experiment changes only one parameter at a time (independent variable) and measures precise indicators. The targets in the BEAI document are used as success criteria.

Code	Experiment	What changes	What is measured	BEAI target
EX1	Voice cloning	2–3 models; 30 s then 60 s sample	Voice similarity, naturalness (1–5 rating by 10 students), latency	inference < 120 ms
EX2	Lip-sync comparison	2–3 models, same photo and audio	SyncNet score, FPS, GPU memory per session	30 ms, 200+ FPS, 3.5 GB / 4 sessions

Code	Experiment	What changes	What is measured	BEAI target
EX3	RAG on a real course	Chunk size, top-k, out-of-course questions	% sourced answers, % correct "I'm not sure"	sourced answers
EX4	Speech translation	Cascade vs single model; FR → AR / EN	Fidelity, voice preserved, sync, SRT export	120+ languages
EX5	Avatar expressions	4 expressions (Neutral, Happy, Serious, Surprised)	% of expressions recognised by students	correct perception
EX6	Generated content	Type: quiz, summary, diagram	Rating by 2 professors (Bloom grid)	teaching quality
EX7	Teaching A/B test	Classic video vs interactive avatar	Score on the same quiz, engagement, satisfaction	project value

Table 4.6 Experimental protocols

Card EX7 — Teaching A/B test (key experiment)	
Guiding Question	Does an interactive avatar help students learn better than a classic video?
Protocol	Two groups of at least 15 students. Group A: a classic 10-minute lecture video. Group B: the same content with the interactive avatar (prototype or Wizard of Oz). Both groups then take the same 10-question quiz.
Measures	Average quiz score, attention time, number of questions asked, satisfaction (1 to 5)
Precautions	Groups of similar level, same duration, written consent from participants
Output	First measured evidence of the project's teaching value, for Part 6

4.4.4 Interviews & Surveys

IS1 — Student survey

Target: at least 100 responses, shared through class groups and student clubs.

Section	Sample questions	Answer type
Study habits	When and where do you study outside class?	Multiple choice
Difficulties	What do you do when you don't understand a lecture?	Choice + free text
Languages	In which language(s) do you prefer explanations?	Derja, French, Arabic, English
Trust	Would you trust an AI professor? Why?	1–5 scale + text
Ethics	Would you allow the platform to use your microphone? Your webcam?	Yes / No / Under conditions
Features	Rank: quizzes, summaries, translation, 24/7 availability	Ranking

Table 4.7 Structure of the student survey

IS2 to IS5 — Interviews

Code	Interviewee	Objective	Sample questions	No.
IS2	Professors	Acceptance of cloning	Would you agree to be cloned? Under which conditions (control, withdrawal, payment)? Which courses are suitable?	3–5

Code	Interviewee	Objective	Sample questions	No.
IS3	AI expert	Feasibility of BEAI targets	Is 500 ms latency realistic? Which models do you recommend?	1–2
IS4	Lawyer / data protection	Legal framework	Image and voice rights? Data retention period? Withdrawal of consent?	1–2
IS5	Educator	What makes a good tutor	How should a tutor react when a student is lost? What is a good rephrasing?	1–2

Table 4.8 Interview plan

4.4.5 Projects & Problem Sets

PP1 — Mini-prototypes for each module

A small prototype (*proof of concept*) is built for each module on Google Colab, Kaggle or Hugging Face Spaces. It supports experiments EX1 to EX6.

PP2 — Problem: the latency budget

The BEAI document gives several latency targets. Put side by side, they raise a real investigation question:

Indicator in the document	Value	Document section
End-to-end latency	< 200 ms	Monitoring & Observability
End-to-end conversational latency	< 500 ms	Key Performance Metrics
LLM first token	200–500 ms (ideal)	Key Performance Metrics
Core model inference	< 120 ms	Key Performance Metrics
API response time	< 800 ms	Key Performance Metrics

Table 4.9 Latency targets quoted in the BEAI document

If the LLM's first token alone can take 500 ms, a total target of 500 ms becomes very hard to meet. The exercise is to split the budget across the links of the chain, then measure each link on the prototype. The split below is a **starting hypothesis**, to be checked through experiments and validated with our professor.

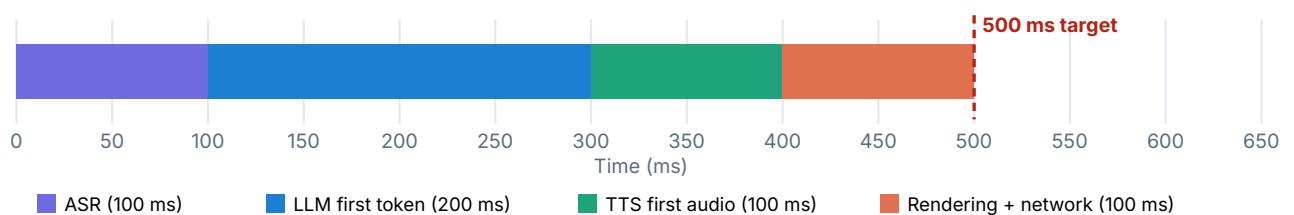


Figure 4.6 Hypothetical split of the latency budget (to be measured)

PP3 — Problem: GPU sizing

The document states that 3.5 GB of GPU memory are needed for 4 concurrent sessions. From this we can estimate the order of magnitude of the number of GPUs required:

$$\text{Sessions per GPU} = \lfloor M_{\text{GPU}} / 3.5 \rfloor \times 4 \quad \text{GPUs needed} = \lceil N_{\text{students}} / \text{Sessions per GPU} \rceil$$

GPU memory	Sessions per GPU	100 students	1,000 students
24 GB	$[6.86] \times 4 = 24$	5 GPUs	42 GPUs
80 GB	$[22.86] \times 4 = 88$	2 GPUs	12 GPUs

Table 4.10 Estimate based on memory only

This calculation is an **optimistic lower bound**: it does not account for the computing power needed to reach 200+ FPS, nor for the memory used by the voice and language models. The activity consists in comparing this calculation with the measurements from EX2, then estimating the monthly cost.

PP4 — Usability test of the wireframes

Five students and two professors go through screens 1 to 7 of the BEAI document on a clickable (Figma) or paper prototype, with a precise task: "create your digital twin" or "create an agent for your course". We measure the success rate, the blocking points and the **SUS** score (*System Usability Scale*, a standard 10-question questionnaire).

4.4.6 Field Research

For an AI project, "the field" means the real places where students will use the platform, the real lectures it must reproduce, and the competing platforms that already exist.

Code	Activity	Description	Used by
FR1	Real study environments	Record audio in a café, a bus, a dorm and a noisy home; measure the noise level	EX1, module 01
FR2	Connection quality	Measure network speed and latency (4G, university Wi-Fi) from these places	PP2, 500 ms target
FR3	Lecture observation	Observe how a professor explains, rephrases and reacts when a student is lost	Modules 05 and 06
FR4	Course material collection	Gather real lectures (slides, PDFs, videos) with the teachers' agreement	EX3, EX4, EX6
FR5	Platform benchmark	Go through D-ID, Tavus and HeyGen with trial accounts and fill in a common grid	Part 5, BEAI screens

Table 4.11 Field research activities

The BEAI platform is explicitly inspired by D-ID (several screens mirror how it works) and refers to the Tavus API. Benchmark FR5 therefore uses the same grid for each platform:

Criterion observed	D-ID	Tavus	HeyGen
Consent and verification steps	...	...	...
Time to create an avatar	...	...	...
Lip-sync quality (rating 1-5)	...	...	...
Languages and quality in Arabic	...	...	...
Interactive agent and knowledge base	...	...	...

Criterion observed	D-ID	Tavus	HeyGen
Pricing model (credits, minutes)	...	...	...

Table 4.12 Benchmark grid to fill in

4.5 Guiding Resources

Resources are grouped according to the four types of the CBL framework. Each resource is linked to the modules and activities it supports.

4.5.1 Digital Media

Resource	Type	Use in the project	Access
arXiv	Preprints	Latest papers for each module	Free
Google Scholar, Semantic Scholar	Academic search engines	Finding papers and tracking citations	Free
Connected Papers	Paper mapping	Finding work related to a survey	Free (limited)
Hugging Face (Hub, Spaces)	Models, datasets, demos	Testing models without a GPU, downloading corpora	Free
GitHub	Source code	Checking project activity and licences	Free
Google Colab, Kaggle	Cloud GPUs	Mini-prototypes and experiments	Free (quotas)
D-ID and Tavus documentation	Product documentation	Reference platforms behind the BEAI design	Free
LiveKit Agents, Pipecat	Real-time frameworks	Voice agent prototype and latency measurement	Open source
MediaPipe	Computer vision	Blink detection (liveness check)	Open source
Hugging Face Audio Course	Online course	Audio processing basics (modules 01, 04)	Free
DeepLearning.AI short courses	Online courses	RAG and agents (module 02)	Free
Prometheus + Grafana	Monitoring	p50/p95 measurements, as in the Monitoring layer	Open source

Table 4.13 Digital media

4.5.2 Textbooks & Articles

Module	Key references	Activities
01 Voice	Whisper (Radford et al., 2022); AV-HuBERT (Shi et al., 2022); LRS3 dataset; Tunisian corpora TunSwitch, LinTO, TuniSpeech-21h	EX1, FR1
02 Language	RAG (Lewis et al., 2020); ColPali (Faysse et al., 2024); RAGAS (Es et al., 2024)	EX3
03 Face	Survey <i>From Pixels to Portraits</i> ; Wav2Lip (2020); SadTalker (2023); MuseTalk (2024); Hallo-Live (2026); SyncNet (Chung & Zisserman, 2016)	EX2, SG2
04 Translation	Seamless (Meta, 2023); UniSS (ICLR 2026); STEB benchmark (2026); TUNIFRA and TEDxTN corpora	EX4
05 Emotion	DAiSEE (Gupta et al., 2016); MEAD (Wang et al., 2020); AI Act, Article 5	EX5, IS4
06 Content	Revised Bloom's taxonomy (Anderson & Krathwohl, 2001); UNESCO guidance on generative AI in education (2023)	EX6, IS5

Module	Key references	Activities
Platform	Apple CBL Classroom Guide (2011); AI Act, Regulation (EU) 2024/1689; C2PA; AudioSeal (2024); SUS (Brooke, 1996); Nielsen's heuristics	PP2, PP4, SG2

Table 4.14 Key papers and reference texts by module

Core textbooks: *Speech and Language Processing* (Jurafsky & Martin) for speech and language; *Deep Learning* (Goodfellow, Bengio & Courville, MIT Press, 2016) for the fundamentals. **Official texts:** the European AI Act (Articles 5 and 50) and Tunisian Organic Law No. 2004-63 on the protection of personal data.

4.5.3 Human Experts

Expert	Expected contribution	Activities
Supervising professor	Validate the BEAI targets, clarify the latency question	PP2, PP3
School teachers (AI, signal processing, vision)	Model choices, technical advice, GPU access	EX1-EX4, IS3
Teachers (all subjects)	Acceptance of cloning, course material	IS2, FR3, FR4
Researchers in Derja NLP	Corpora and models for Tunisian Arabic	EX1, EX4
INPDP (Tunisian National Authority for Personal Data Protection)	Tunisian legal framework for voice, image and student data	IS4
Digital law specialist	Image rights, consent, AI Act	IS4
Educator / teaching adviser	How a good tutor behaves	IS5, EX6

Table 4.15 Human experts

4.5.4 Networks

Network	What it brings	Activities
Classmates	Participants in tests, the survey and the Wizard of Oz	IS1, SG1-SG3, EX7
Alumni	Outside perspective, professional contacts	IS3
Student clubs (IEEE, Google Developer Groups, AI clubs)	Technical profiles to test the prototypes	PP1, PP4
Online communities (Hugging Face, LiveKit)	Answers to technical questions	PP1, EX2
Family and teachers we know	Views of outside teachers on cloning	IS2

Table 4.16 Networks involved

4.5.5 Standard resource card

Every resource added to the team library is recorded with the same card:

Name	Type	Module	Link	Year	Licence	Limits
e.g. DAISEE	Dataset	05	IIT Hyderabad website	2016	On request	Highly imbalanced classes

Table 4.17 Resource card template

4.6 Coverage matrix

The matrix below checks that every module and the platform are covered by several activities of different types.

Type	Activity	01	02	03	04	05	06	Platf.
Simulations	SG1 Wizard of Oz		✓			✓	✓	✓
	SG2 Real or AI	✓		✓				✓
	SG3 Perceived latency	✓		✓				✓
Experiments	EX1 Voice cloning	✓						
	EX2 Lip-sync			✓				✓
	EX3 RAG		✓					
	EX4 Translation	✓			✓			
	EX5 Expressions					✓		
	EX6 Generated content						✓	
	EX7 A/B test	✓	✓	✓		✓	✓	✓
Interviews	IS1 Student survey	✓			✓	✓		✓
	IS2 Professors	✓		✓			✓	✓
	IS3 AI expert	✓	✓	✓				✓
	IS4 Lawyer / INPDP					✓		✓
	IS5 Educator					✓	✓	
Projects	PP1 Mini-prototypes	✓	✓	✓	✓	✓	✓	
	PP2 Latency budget	✓	✓	✓				✓
	PP3 GPU sizing			✓				✓
	PP4 UX test (SUS)							✓
Field	FR1 Environments	✓						
	FR2 Connection							✓
	FR3 Lecture observation					✓	✓	
	FR4 Course material		✓		✓		✓	
	FR5 Benchmark			✓	✓			✓

Table 4.18 Module coverage by the 24 activities

HOW TO READ IT

Every module is covered by at least three activities, and always by at least two different CBL types. Our results therefore never rely on a single source of information.

4.7 Schedule

The schedule below covers four weeks and will be adjusted to the presentation date of the Investigate phase.



Figure 4.7 Planned schedule for Part 4

4.8 Outputs and link with Part 6

Each activity produces a concrete result that feeds the overall synthesis and the objective tree.

Activities	Output	Contribution to Part 6
EX1-EX6, PP1	Comparison tables of the models, with measurements	Justified technical choice for each module
PP2, PP3, SG3, FR2	Latency budget and GPU estimate	Feasibility and cost of the platform
IS1, SG1, FR3	Summary of student needs	Feature prioritisation
IS2, IS4, SG2	Conditions for accepting cloning	Ethical rules of the platform
EX7, EX6, IS5	First measure of teaching value	Core argument of the project
PP4, FR5	SUS score and benchmark grid	Improvements to the BEAI user journey

Table 4.19 Part 4 outputs

4.9 Points of attention

- **Participant consent:** anyone recorded or filmed (tests, Wizard of Oz, A/B test) signs a consent form; data is anonymised and deleted at the end of the project.
- **Professors' voices and faces:** no cloning without written agreement, in line with screen 2 (Capture & Consent) of BEAI.
- **Dataset licences:** some corpora (for example DAiSEE) are only available on request and for research; the licence is checked before any use.
- **Emotions:** no analysis of students' faces (AI Act, Article 5); confusion is detected through conversational signals only.
- **Latency targets:** the gap between the figures in the document is presented as a question to clarify with our professor, not as an error.
- **Computing resources:** free GPU quotas are limited; experiments are planned to stay within them.

4.10 Conclusion

Part 4 turns the project's questions into a concrete investigation plan: 24 activities across the five types of the CBL framework, and a resource library organised into its four categories. Each activity is linked to a module, to a component of the BEAI platform and to a measurable target. This approach ensures that the choices presented in the Act phase will be backed by evidence: measurements, user feedback and expert opinion.

Clear questions, measurable activities, reliable sources: the foundation of a solution people can trust.

Main references

- Apple Inc. (2011). *Challenge Based Learning: A Classroom Guide*. apple.com/br/education/docs/CBL_Classroom_Guide_Jan_2011.pdf
- BEAI Studio – Agent & Digital-Twin Creation Platform. Reference document provided by the professor.
- Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (AI Act), Articles 5 and 50.
- UNESCO (2023). *Guidance for generative AI in education and research*.
- Radford, A. et al. (2022). Robust Speech Recognition via Large-Scale Weak Supervision (Whisper).
- Lewis, P. et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks.
- Faysse, M. et al. (2024). ColPali: Efficient Document Retrieval with Vision Language Models.
- Prajwal, K. R. et al. (2020). A Lip Sync Expert Is All You Need for Speech to Lip Generation In the Wild (Wav2Lip).
- Seamless Communication et al. (2023). Seamless: Multilingual Expressive and Streaming Speech Translation. arxiv.org/abs/2312.05187
- Cheng et al. (2026). UniSS: Unified Expressive Speech-to-Speech Translation with Your Voice. ICLR 2026.
- Gupta, A. et al. (2016). DAiSEE: Towards User Engagement Recognition in the Wild. arxiv.org/abs/1609.01885
- Tunisian corpora: TunSwitch (2023), LinTO (2025), TEDxTN (2025), TUNIFRA (2025), TuniSpeech-21h (2026).
- Brooke, J. (1996). SUS: A quick and dirty usability scale.
- Anderson, L. W. & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing*.

4.11 Completion of the plan deliverables COMPLETION

The Investigate plan asks Part 4 for a table **Activity · Skill · Why needed · Learning resource · Type · Link** and a list of priority skills. This section adds them, using only the activities and resources defined above. It also adds the missing resource links, the FR5 benchmark grid filled from the team's evidence, and a dated schedule.

4.11.1 Activity → Skill table

Activity	Skill to acquire	Why it is needed	Learning resource	Type	Link
SG1 Wizard of Oz	Running a Wizard-of-Oz study and observing users	Reveals real student questions before any code exists	Dahlbäck, Jönsson & Ahrenberg, "Wizard of Oz studies — why and how" (1993)	Article	doi.org/10.1016/0950-7051(93)90017-N
SG2 Real or AI?	Deepfake literacy; AI labelling and provenance	Justifies the mandatory AI label (Art. 50)	Detect Fakes (Northwestern Kellogg) · C2PA explainer	Online test · Standard	detectfakes.kellogg.northwestern.edu
SG3 Perceived latency	Response-time perception in HCI	Sets the latency threshold from real users	NN/g, "Response Times: The 3 Important Limits"	Article	www.nngroup.com/articles/response-time
EX1 Voice cloning	Speech synthesis and voice cloning; MOS rating	Module 01: faithful teacher voice	Hugging Face Audio Course; XTTS-v2 model card	Online course · Model	huggingface.co/learn/audio-course
EX2 Lip-sync	Talking-face generation; SyncNet, FPS and GPU profiling	Module 03: realism vs real time	Wav2Lip repository; SyncNet (Oxford VGG)	Code · Paper	github.com/Rudrabha/Wav2Lip
EX3 RAG	Retrieval-augmented generation and its evaluation	Module 02: sourced answers, "not in your course"	LangChain RAG tutorial; RAGAS documentation	Tutorial · Docs	docs.ragas.io
EX4 Speech translation	Speech-to-speech translation (cascade vs end-to-end)	Module 04: lecture in the learner's language	Meta Seamless Communication repository	Code · Paper	github.com/facebookresearch/seamless_c
EX5 Avatar expressions	Perception studies; descriptive statistics	Module 05: expressions read correctly	A. Field, <i>Discovering Statistics</i> (companion site)	Textbook	www.discoveringstatistics.com
EX6 Generated content	Pedagogical evaluation with Bloom's taxonomy	Module 06: teaching value of quizzes and summaries	Vanderbilt Center for Teaching, Bloom's Taxonomy guide	Guide	cft.vanderbilt.edu/guides-sub-pages/bl
EX7 Teaching A/B test	Controlled experiments; t-test; sample size	First measured evidence of teaching value	Kohavi, Tang & Xu, <i>Trustworthy Online Controlled Experiments</i>	Textbook	experimentguide.com

Activity	Skill to acquire	Why it is needed	Learning resource	Type	Link
IS1 Student survey	Questionnaire design; sampling	Real needs, languages and trust of students	Pew Research Center, "Writing survey questions"	Guide	www.pewresearch.org/our-methods/u-s-su
IS2 Professors	Semi-structured interviews; consent discussion	Conditions for professors to be cloned	NN/g, "User Interviews 101"	Article	www.nngroup.com/articles/user-intervie
IS3 AI expert	Technical questioning; reading model cards	Feasibility of BEAI targets	Hugging Face model cards and leaderboards	Platform	huggingface.co/docs/hub/model-cards
IS4 Lawyer	AI and data-protection law (AI Act, GDPR Art. 9, Law 2004-63)	Image, voice and consent framework	EUR-Lex AI Act text; CNIL AI guidance	Official texts	eur-lex.europa.eu/eli/reg/2024/1689/oj
IS5 Educator	Tutoring pedagogy; Socratic questioning	What a good tutor does when a student is lost	UNESCO, <i>Guidance for generative AI in education and research</i> (2023)	Report	www.unesco.org/en/articles/guidance-ge
PP1 Mini-prototypes	Rapid ML prototyping and demo apps	Proofs of concept for EX1-EX6	Gradio; Hugging Face Spaces documentation	Docs	huggingface.co/docs/hub/spaces
PP2 Latency budget	Real-time voice pipelines; p50/p95 measurement	Checks the < 500 ms target link by link	LiveKit Agents; Pipecat documentation	Framework docs	docs.livekit.io/agents
PP3 GPU sizing	GPU memory estimation and cloud cost	Cost per student at scale	Hugging Face Accelerate model memory estimator	Tool	huggingface.co/spaces/hf-accelerate/mo
PP4 Usability test	Usability testing; SUS questionnaire; Figma prototyping	Clear user journey for screens 1-7	NN/g, "Usability Testing 101"; Figma	Article · Tool	www.nngroup.com/articles/usability-tes
FR1 Study environments	Audio data collection; noise measurement (dB)	Real noise conditions for module 01	Audacity (recording and spectrum analysis)	Tool	www.audacityteam.org
FR2 Connection quality	Network measurement (bandwidth, latency, jitter)	Feasibility of real-time video vs low-bandwidth mode	Ookla Speedtest; M-Lab speed test	Tool	www.speedtest.net
FR3 Lecture observation	Field observation with a structured grid	How a teacher rephrases and reacts (modules 05, 06)	NN/g, "Field Studies"	Article	www.nngroup.com/articles/field-studies
FR4 Course collection	Licensing of course material (Creative Commons)	Legal knowledge base for EX3, EX4, EX6	Creative Commons licence guide	Guide	creativecommons.org/licenses

Activity	Skill to acquire	Why it is needed	Learning resource	Type	Link
FR5 Platform benchmark	Structured product benchmarking	Compares D-ID, Tavus and HeyGen on one grid	D-ID, Tavus and HeyGen developer docs	Docs	docs.tavus.io

Table 4.20 Activity · Skill · Why needed · Learning resource · Type · Link (24 activities)

4.11.2 Priority skills for the team

#	Skill	Modules	Activities	Priority
1	Speech processing: ASR, TTS, voice cloning	01, 04	EX1, EX4, FR1	HIGH
2	RAG and LLM evaluation (sourced answers, hallucination tests)	02	EX3, IS3	HIGH
3	Talking-face generation and lip-sync evaluation	03	EX2, SG2	HIGH
4	Real-time pipelines and latency measurement	Platform	PP2, SG3, FR2	HIGH
5	Experimental design and statistics (A/B test, sample size)	All	EX5–EX7	HIGH
6	Survey and interview design; qualitative analysis	All	IS1–IS5, SG1	MEDIUM
7	AI and data-protection law (AI Act Art. 5 and 50, GDPR Art. 9, Law 2004-63)	All, enablers	IS4, SG2	MEDIUM
8	Pedagogical design (Bloom's taxonomy, Socratic tutoring, knowledge tracing)	05, 06	EX6, IS5, FR3	MEDIUM
9	UX research and usability testing (SUS, prototyping)	Platform	PP4	LOWER
10	GPU sizing and cloud cost estimation	Platform	PP3	LOWER

Table 4.21 Priority skills, ranked by how many modules and key experiments depend on them

4.11.3 Links for the digital media of Table 4.13

Resource	Link
arXiv	arxiv.org
Google Scholar	scholar.google.com
Semantic Scholar	www.semanticscholar.org
Connected Papers	www.connectedpapers.com
Hugging Face Hub and Spaces	huggingface.co
GitHub	github.com
Google Colab	colab.research.google.com
Kaggle	www.kaggle.com
D-ID documentation	docs.d-id.com
Tavus documentation	docs.tavus.io
LiveKit Agents	docs.livekit.io/agents/
Pipecat	github.com/pipecat-ai/pipecat
MediaPipe	ai.google.dev/edge/mediapipe
Hugging Face Audio Course	huggingface.co/learn/audio-course
DeepLearning.AI short courses	www.deeplearning.ai/short-courses/
Prometheus	prometheus.io
Grafana	grafana.com

HOW TO USE THE RESOURCE CARD

Each resource in Tables 4.13–4.16 should get a card (Table 4.17) when it is used. The links on the left complete the "Access" column; papers in Table 4.14 are linked in Part 3A.7 and in the Part 3B reference list.

Table 4.22 Links for the digital media

4.11.4 FR5 benchmark grid, pre-filled

Pre-filled from the hands-on tests of Part 3B and the market study of Part 5. Cells marked "to measure" must be completed with trial accounts during FR5 and EX2.

Criterion	D-ID	Tavus	HeyGen
Consent and verification steps	Self-declared: an agent with Einstein's face was created without any authorization (3B)	On-camera consent statement required to train a personal replica (5)	Consent video required for custom video avatars (vendor policy, to verify)
Time to create an avatar	Minutes, from a single photo (3B)	Replica trained from a short recorded video; to measure	Photo avatar in minutes; video avatar needs a recording; to measure
Lip-sync quality (1–5)	To measure (EX2)	To measure (EX2)	To measure (EX2)
Languages and quality in Arabic	Arabic voices offered (vendor); Derja not tested	Multilingual (vendor); Arabic and Derja quality not tested (Part 5, gap 18)	100+ languages with lip-synced translation (vendor); Derja not tested

Criterion	D-ID	Tavus	HeyGen
Interactive agent and knowledge base	Yes: agents with a knowledge base, tested; answered every question, never "I don't know" (3B)	Yes: real-time CVI with persona and knowledge base; "Socratic tutor" only by prompt (3B)	Interactive avatar API exists; positioned for content production, not tutoring (3B)
Pricing model	Credits; free trial of about 12 credits for 14 days (3B)	From \$59/month (175 min in 3B, 100 min in 5 — to confirm); 3, 10 or 15 concurrent sessions	Freemium; credits or minutes per plan (vendor)

Table 4.23 FR5 benchmark grid (sources: Part 3B tests, Part 5, vendor pages)

4.11.5 Dated schedule (proposal)

Dates assume a start on Monday 12 October 2026 and must be shifted to match the official presentation date.

Week	Activities	Milestone
W1 · 12–18 Oct	Validate the Guiding Questions (Part 1 + Table 4.4); collect resources and cards; write the survey and interview guides; SG1–SG3	GQ list frozen; IS1 form ready
W2 · 19–25 Oct	Distribute survey IS1; experiments EX1–EX4 on the mini-prototypes (PP1); FR1–FR2 measurements	≥ 100 survey answers started; first EX results
W3 · 26 Oct–1 Nov	Interviews IS2–IS5; experiments EX5–EX7; FR3–FR5 (lecture observation, course collection, benchmark)	A/B test (EX7) done; FR5 grid complete
W4 · 2–8 Nov	PP2 latency budget, PP3 GPU sizing, PP4 usability test; analysis of all results; hand-over to Part 6	Results table to Part 6; final synthesis

Table 4.24 Dated schedule (proposal)

05

RESPONSIBLE
Youssef Jouini
3 October 2026

Legal, Ethical & Market Study

State of the art + market + constraints

Scope and method

Our product is a learning platform where GenAI-powered 3D avatars of professors, thinkers and experts answer students' questions by voice, anywhere and in their own language. This study checks what law, ethics and the market say about that idea before we build it.

JURISDICTIONS STUDIED

- **Tunisia** — where the team, ESPRIT and our first users are.
- **European Union** — the reference model Tunisia is aligning with, and a likely market (the AI Act applies to anyone placing AI on the EU market).
- **United States** — our potential partner, Indiana University, is there.
- **International soft law** — UNESCO guidance, which most education ministries follow.

METHOD

We favoured official and institutional sources (European Commission, UNESCO, national authorities, peer-reviewed studies), then company pages for product facts. Every figure carries its year. What we could not prove is listed under *Research gaps*, not guessed.

This is a student research document, not legal advice.

Partner input: Indiana University's BEAI Studio document

Indiana University sent us a reference design, **BEAI™ Studio – Agent & Digital-Twin Creation Platform** ("Building Ethical AI", ref. FBM V1A26). It confirms our technical route — a real-time talking avatar answering from a knowledge base (RAG) — and it builds in two ethical safeguards. It does not yet cover several obligations specific to education, which this study adds.

What BEAI proposes	Why it matters for this study	Section
4-layer architecture: web, mobile and 3D real-time SDK; API gateway; video generation (via the Tavus API), interactive agents, avatar management and voice cloning; speech (ASR/TTS), lip-sync, LLM + RAG; GPU cluster, Kubernetes, CDN	Avatars would run on third-party APIs and foreign clouds, so student data leaves Tunisia	5.1
Two paths from step one: "Build an AI Agent" (synthetic persona) vs "Clone Myself" (digital twin), because they carry different consent and identity risks	Matches our split between reconstructed thinkers and real, living professors	5.1, 5.2

What BEAI proposes	Why it matters for this study	Section
Consent checkbox ("I own the rights to this likeness") and a liveness check before any generation	Good practice; face and voice templates can be biometric data	5.1, 5.2
Knowledge base with retrieval settings, an "I'm not sure" fallback and a test panel showing the source of each answer	Directly addresses hallucinations	5.2
Credit/minute meter shown during generation; free voice library, voice cloning as a gated premium feature	Business model of the ecosystem	5.4
Publishing as MP4, embeddable widget or API, with private, team or public sharing	Public sharing brings platform-liability duties	5.1
Target KPIs: under 500 ms conversational latency, lip-sync within 30 ms, 200+ FPS, 3.5 GB GPU memory per 4 sessions	Useful targets, but no sources are given (citation markers like "[16†20]" have no reference list)	Research gaps

Existing solutions the document cites. D-ID is the model for the whole studio and its wireframes; Tavus is named as the video-generation engine. Both were missing from our first version and are now in 5.3. The document also mentions "Athena" for the conversational agent without explaining it.

WHAT THE DESIGN DOES NOT YET COVER FOR EDUCATION

No "this is an AI" label or watermark (required in the EU since Aug 2026); no age gate or parental consent; no Socratic mode (the agent answers directly); no hand-off to the real teacher; no low-bandwidth mode; no process for deceased thinkers' likeness; and the knowledge-base example is an e-commerce return policy, not a course.

5.1 Legal constraints

The binding constraints today are AI transparency (EU, since August 2026), data protection (Tunisia and EU), minors' consent, transfers to foreign avatar providers, and explicit consent for anyone's voice or face, which can be biometric data. Strict "high-risk" rules apply only if our avatars grade, place or admit students, and only from December 2027.

Area	Rule (jurisdiction)	What it requires	What it means for our avatars	Status (as of Oct 2026)
AI and education	AI Act, Reg. 2024/1689, Annex III (EU)	AI used in education that can decide access to education or a person's course (e.g. exam scoring) is high-risk: risk management, quality datasets, logging, documentation, human oversight, robustness	A tutor that only explains is outside Annex III. If the avatar grades, places or admits students, the full high-risk regime applies	Applies from 2 Dec 2027, moved from Aug 2026 by the AI Omnibus, in force 27 Jul 2026 (Commission)
AI transparency	AI Act, Art. 50 (EU)	Users must know they are talking to a machine; generated content must be identifiable; deepfakes must be visibly labelled	Every avatar must say it is an AI; synthetic voices and faces of real people must be labelled and marked. The BEAI design has no such label yet	In application since 2 Aug 2026; watermarking grace period to 2 Dec 2026 (Winss , citing Commission)

Area	Rule (jurisdiction)	What it requires	What it means for our avatars	Status (as of Oct 2026)
Prohibited practices	AI Act, Art. 5 (EU)	Emotion recognition in education institutions is banned	No webcam-based emotion detection to adapt lessons for schools or universities. BEAI's "sentiment control" must stay on text, not on the student's face	Banned since Feb 2025 (Commission)
Data protection	Organic Law 2004-63 (Tunisia) , supervised by INPDP	Prior declaration to INPDP; consent; authorisation for sensitive data and for transfers abroad	Student questions and voice recordings are personal data. Sending them to a foreign cloud or LLM API needs INPDP authorisation	In force. Replacement bill PLO 095/2025 (DPO, breach notification, rights against automated decisions) is in committee at the ARP (Global Law Experts , Aug 2026)
Foreign processors	Same Tunisian transfer rules; GDPR if EU users	A contract with each processor; authorisation for transfers outside Tunisia	Using D-ID (based in Tel Aviv, per its privacy policy), Tavus or an LLM API means a cross-border transfer of student and professor data	In force; no INPDP decision on these providers found
Biometric data (face, voice, liveness)	GDPR Art. 9 (EU); sensitive data under Law 2004-63 (Tunisia)	Biometric data processed to uniquely identify a person is a special category: prohibited unless an exception applies, usually explicit consent (GDPRhub)	Voice cloning and BEAI's liveness check create face and voice templates of professors: explicit, separate consent and an impact assessment	In force since 2018
Data protection of minors	GDPR Art. 8 (EU)	For online services offered directly to children on the basis of consent: parental consent under 16; states may lower to 13	Age gate plus parental consent flow if we open to EU school pupils	In force since 2018 (GDPR text)
Data protection of minors	COPPA Rule, amended 2025 (US)	Verifiable parental consent before collecting data from children under 13; written data-retention policy	Relevant only if we serve US children (e.g. via a US partner)	Compliance required since 22 Apr 2026 (Loeb & Loeb)
Platform liability	Digital Services Act (EU)	Rules for notifying platforms of illegal content and acting on it	If users can share or publish avatar content (BEAI's "public link"), we need a notice-and-action process	In force; cited by the EU IP Helpdesk as the takedown route for deepfakes
Copyright (training data)	AI Act GPAI rules on copyright and training-data summaries (EU)	Providers of general-purpose models must respect copyright and publish a summary of training content	Course material and books uploaded to the knowledge base need a licence or a lawful exception	GPAI rules in application since Aug 2025 (Commission)
Voice and likeness of living people	Danish draft amending the Copyright Act	Consent needed to publish realistic digital imitations of a person's appearance or voice	A real professor's avatar needs written consent and a licence; providers already require it (Tavus asks for on-camera consent, see 5.3)	Notified to the Commission; expected in force 1 Jul 2026, adoption not confirmed (EU IP Helpdesk , Jun 2026)
Likeness of deceased people	No specific rule found in Tunisia; industry practice	For a video of five deceased people, D-ID obtained consent from their families and co-wrote the scripts with them (Partnership on AI case study)	Avatars of recent thinkers: ask heirs or estates; ancient thinkers: label clearly as reconstructions	Practice, not law — see Research gaps
Academic use	UNESCO Guidance on GenAI in education (international)	Institutions should validate AI systems before student use; age limit of 13 for classroom use; teacher training	ESPRIT or any school would need to approve the tool before it is used in courses	Guidance, not binding (Sep 2023)

Table 5.1 Legal constraints by area and jurisdiction

LEARNER-EXPERT INTERACTIONS

No text we found says who is liable when an avatar answers in a real professor's name. The safe reading is a contract with each expert (scope, review rights, withdrawal), a clear "AI, not the person" label, and a human escalation path. This is listed as a research gap.

5.2 Ethical questions

The biggest ethical risk is not a wrong fact but a convincing face saying it: an avatar with a famous professor's voice makes errors, bias and over-reliance more persuasive. Evidence shows design choices decide whether GenAI helps or harms learning.

Question	Evidence (year)	Risk for our avatars	Mitigation we commit to
Reliability and verification of knowledge	Field experiment with ~1,000 students: plain GPT-4 raised practice grades 48% but lowered exam grades 17% once removed; a tutor with guardrails avoided the loss (Bastani et al., PNAS 2025)	Students copy answers instead of learning	Socratic mode by default: hints and questions before answers; answers grounded in a validated course corpus, with the source shown
Misinformation and hallucinations	The same study stresses GenAI is fallible and users must check outputs (PNAS 2025); only 46% of people trust AI (KPMG/University of Melbourne 2025, from our Engage deck)	An invented quote in Socrates' or a real professor's voice looks authoritative	Retrieval from approved texts only; "I don't know" allowed; quotes linked to their source; error reporting button
Bias	LLMs studied by UNESCO (GPT-2, GPT-3.5, Llama 2) showed clear bias against women and cultural stereotypes (UNESCO/UCL, Mar 2024)	Avatars reproduce stereotypes; a catalogue of mostly Western, male thinkers narrows whose knowledge counts	Bias tests before release; diverse catalogue (including Arab and African scholars, women thinkers); audit samples each term
Confidentiality	Student voice and questions are personal data under Tunisian Law 2004-63 and the GDPR (see 5.1)	Questions reveal weaknesses, health or personal issues; voice can identify a person	Data minimisation, no training on student data without opt-in, deletion on request, local or authorised hosting
Learner safety	Character.AI removed open-ended chat for under-18s after lawsuits alleging harm to minors; the FTC is probing 7 chatbot companies (Fortune, Oct 2025)	Emotional attachment to a human-like avatar; harmful replies to a distressed student	Avatars stay on the learning topic; no companion role-play; distress detection with a route to human help; age gate
Equity of access	2.2 billion people offline in 2025; only 36% of Africa online; 23% in low-income countries (ITU Facts and Figures 2025 , reported by Guardian Nigeria)	Real-time 3D and voice need bandwidth and recent devices; the students who need help most are excluded	Low-bandwidth mode (audio or text only), offline packs, Arabic and French support, free tier for students
Human responsibility and oversight	UNESCO asks institutions to validate AI tools before student use (UN News, Sep 2023); the AI Act requires human oversight for high-risk uses (Commission)	No one checks what the avatar teaches	Each avatar has a named human owner (the professor or a curator) who reviews content and answers escalated questions
AI and human expertise	Guardrailed AI tutors help, unguarded ones can replace thinking (PNAS 2025); voices can be cloned from seconds of audio (McAfee, from our Engage deck)	Avatars used to cut teaching staff, or real experts cloned without consent	Avatars extend teachers, never replace them; only consenting experts, with revenue share and right to withdraw; historical figures clearly marked as reconstructions
Consent and likeness (added after the BEAI document)	Tavus requires on-camera consent from the person cloned and bans replicas of minors (The Rundown); D-ID obtained families' consent for avatars of deceased people (Partnership on AI); BEAI puts a consent checkbox and a liveness check before generation	A professor's face and voice used beyond what they agreed to; a deceased thinker made to say things they never said	Keep BEAI's consent-first flow; written licence per professor with scope, review and withdrawal; heirs consulted for recent thinkers; every historical avatar labelled as a reconstruction with sources

Table 5.2 Ethical questions, evidence, risks and mitigations

5.3 Study of existing solutions — state of the art

Existing solutions fall into six families: Socratic AI tutors, real-time avatar agents, expert voice clones, historical-figure chatbots, avatar video generators, and human tutoring. The real-time avatar platforms named in Indiana's document (D-ID, Tavus) already combine a talking face, live voice Q&A and a knowledge base — but for business use, not for teaching. Rows are grouped by family, not ranked.

Hands-on screenshots of Synthesia, D-ID, Tavus and HeyGen are reproduced in Part 3, section 3.1 (working notes "To be used for Part 5").

Solution	Family	Function	Target users	Access model	Main features	Documented limits
Khanmigo (Khan Academy)	Socratic AI tutor	Text tutor on Khan Academy content; teacher assistant	K-12 learners, parents, teachers, US districts	Free for teachers; paid plan for learners and families (about \$4/month, Toolchase); district pricing on request	Guides with hints instead of answers; parent moderation alerts; lesson planning	A 2-year University of Toronto study found students largely stopped using it when it asked questions instead of answering; Sal Khan called it a "non-event" for many students (Apr 2026) (KQED MindShift)
ChatGPT Study mode / Gemini Guided Learning	General LLM learning modes	Step-by-step, Socratic help on any topic	Students, lifelong learners	Included in free tiers (launched 2025) (Maginative)	Socratic questions, quizzes, visuals (Gemini, built on LearnLM)	Students can switch back to direct answers at any time (Maginative); no named teacher, no avatar; unguarded GPT-4 use lowered exam grades 17% (PNAS 2025)
D-ID (cited by Indiana as the model for BEAI)	Real-time avatar agent + avatar video	Talking-avatar videos from a photo or video, and live visual agents that answer by voice	Businesses in marketing, customer service, training and education; educators are among its most active users on G2 (G2)	Free trial; paid plans from about \$4.70–5.90/month; Lite is personal use only; agents billed 0.5 credit per 15 seconds of video (D-ID , Captterra)	RAG knowledge base from uploaded documents; 120+ languages; voice cloning; personal digital twin; embeddable in websites and LMS	Avatars can still show unnatural eye or facial motion and timing (The Rundown); credit pricing gets expensive at volume (Filmora); generic agent, no pedagogy layer

Solution	Family	Function	Target users	Access model	Main features	Documented limits
Tavus CVI (cited by Indiana as BEAI's video engine)	Real-time avatar agent (developer API)	Live face-to-face video conversations with an AI replica, plus scripted video generation	Developers building tutors, interviewers, sales or support agents	Free tier (25 min); Starter about \$59/month with 100 min; about \$0.32–0.37 per conversation minute (Fixed Labs, Selvia)	Under 500 ms end-to-end latency; custom replica from a short training video; perception and turn-taking models	Cost scales with every minute of conversation; requires explicit on-camera consent and bans replicas of minors (The Rundown); you build the product and consent flow yourself
Soul Machines	Real-time "digital people" (enterprise)	Emotionally responsive digital humans for customer service, health and education	Large companies	Custom enterprise contracts	Autonomous animation; integration with OpenAI, Gemini and others	Entered receivership on 5 Feb 2026 after raising over US\$135 million; normal services suspended (NZ Herald)
MasterClass On Call	Licensed expert voice persona	Voice calls with AI versions of real instructors	Adults, professionals	Separate subscription, \$10/month or \$84/year at launch (Dec 2024) (Design Compass)	Experts licensed their voice and knowledge; trained on their own material; ElevenLabs voice; over 75% of users choose voice (ElevenLabs)	Few personas at launch; some voices judged robotic; coaching and role-play, not academic curricula; labelled beta
Delphi	Expert "digital mind" platform	Experts clone themselves for text, voice or video Q&A	Coaches, creators, consultants, authors	Tiered subscriptions; experts charge their own audience (Airframe)	Trained on the expert's books, talks, podcasts; 40+ languages; voice and video calls	Quality limited by the expert's content; an unsanctioned clone of a physician had to be removed after a takedown request (VentureBeat); not built for curricula
Hello History	Historical-figure chatbot	Text chat with 400+ historical figures	Middle and high school students, teachers	Freemium app; premium about \$5.99/month (Curaited)	Wide catalogue; read-aloud audio; classroom use	Warns users to fact-check; not always historically accurate. A similar app, Historical Figures Chat, mixed up biographies and made Nazi leaders sound apologetic (Rolling Stone)

Solution	Family	Function	Target users	Access model	Main features	Documented limits
Synthesia (similar: HeyGen)	Avatar video generator	Turns a script into a video of a talking avatar	Companies, corporate trainers, course creators	Free plan (3 min/month) and Creator plan \$89/month (Toolsforhumans); Starter about \$18/month (Costbench); enterprise	140+ languages; custom avatar of a real person with consent; quizzes in video	One-way, pre-scripted video: no live questions; reviewers note challenges in realism and usability (Gartner Peer Insights)
Character.AI	Persona chat (entertainment)	Open chat with user-made characters	General public	Freemium	Very human-like, memory of past chats	Removed open-ended chat for under-18s after lawsuits over harm to minors (Oct 2025) (Fortune)
Preply	Human tutoring marketplace	Live video lessons with independent tutors	Language learners mainly; companies	Pay per lesson at the tutor's rate; platform takes 33% down to 18% commission and 100% of trial lessons (Sacra)	Real humans; tutor choice; flexible schedule	Cost per hour; tutor availability and time zones; quality varies
University office hours, email, recordings	Current baseline	Ask the teacher in person or by email; rewatch lectures	Enrolled students	Free with tuition	Real expert, accurate	Our Engage research: students don't dare ask in public, emails go unanswered, recordings don't answer questions

Table 5.3 Existing solutions, grouped by family (not ranked)

5.4 Market study

The need is best documented in three places: Tunisian families paying for private tutoring, regions short of teachers, and online learners who drop out. Demand for AI help is already there (95% of UK students use AI), while institutions lag in supporting it.

Potential users and evidence of need

User group	Evidence of need or demand (year)	What they would use us for
Secondary students and families in Tunisia	67% of families pay for private lessons; the national bill is about 1,468 million TND a year; an hour costs 15–20 TND (FTDES study, 2023) (L'Economiste Maghrébin). Some bac candidates' families spend up to 15,000 TND in a year (parents' association, May 2026) (La Presse)	Affordable help outside class hours, especially before national exams
University students (ESPRIT and others)	95% of UK undergraduates use AI and 94% use it for assessed work, but only 38% get AI tools from their institution (HEPI, survey of Dec 2025) (HEPI 2026 , Advance HE). Our own Engage research: unanswered emails, fear of asking in public	Ask "their" professor's avatar after class, in private, at any hour
Online and lifelong learners	On MIT and Harvard edX courses, completion fell from 6% (2013–14) to about 3% (2017–18) and second-year return fell from 38% to 7% (Reich & Ruipérez-Valiente, Science 2019 ; figures via Times Higher Education and Pieuvre)	A face and a voice to ask when stuck, to reduce drop-out
Learners in teacher-short regions	44 million more teachers needed by 2030; 15 million in sub-Saharan Africa (UNESCO Global Report on Teachers, 2024); Northern Africa needs 4.3 million (UNESCO factsheet, Oct 2023) (Down To Earth)	Access to expert explanations where no specialist teacher exists
Professors and experts (supply side)	Experts already license their voice and knowledge (MasterClass On Call) or sell access to their clone (Delphi) — see 5.3	Scale their office hours; earn from their knowledge

Table 5.4 Potential users and evidence of need

Institutions and organisations concerned

- **Universities and engineering schools** — ESPRIT, and Indiana University as a potential partner: buyers, content providers, and validators of the tool.
- **Ministries of Education and of Higher Education (Tunisia)** — rule-makers on school use and possible public funders of free tutoring. The parents' association is calling for public, regulated tutoring structures ([La Presse](#), May 2025).
- **INPDP** — approves our data processing and transfers (see 5.1).
- **Parents' and consumer associations** — the Tunisian parents' association and OTIC publicly track tutoring costs.
- **International bodies** — UNESCO and the Teacher Task Force (SDG 4), possible grant or pilot partners.
- **Companies** — corporate training buyers already pay for avatar video ([Synthesia](#) reports 50,000+ companies, per [Toolsforhumans](#)).

Contexts where the need is most documented

1. **Exam years in Tunisia** (bac, brevet): costs peak, and success is increasingly tied to family income ([La Presse](#), Aug 2026).
2. **Sub-Saharan and Northern Africa**: the largest teacher gaps ([UNESCO](#), above), but also the lowest connectivity (36% of Africa online, [ITU 2025](#)), so delivery must be light.

3. **Self-paced online courses:** very low completion, documented over six years of edX data.
4. **Higher education after class:** students use AI anyway, and fewer than half feel supported by teaching staff (HEPI 2026).

Business and access models already used

Model	Example	How it works
Low-cost subscription	Khanmigo, Hello History, MasterClass On Call	About \$4 to \$10 a month per learner (see 5.3)
Free for teachers, sponsored	Khanmigo	Teacher tools paid for by a sponsor (Microsoft) (Toolchase)
Institutional licence	Khanmigo District, ChatGPT Edu	The school or district pays per seat or on quote
Free inside a general AI	ChatGPT Study mode, Gemini Guided Learning	Bundled in free tiers; free student plans in some countries
Expert revenue share	Delphi, MasterClass	The expert licenses their knowledge and voice and is paid from subscriptions
Marketplace commission	Preply	Platform keeps 18–33% of each human lesson
Usage-based	Synthesia	Pay by video minutes produced
Credits or minutes of avatar time (API)	D-ID, Tavus — the model BEAI copies with its minute meter	D-ID charges 0.5 credit per 15 seconds of agent video; Tavus about \$0.32–0.37 per conversation minute, roughly \$21 an hour (Selvia)
Bespoke enterprise contracts	Soul Machines	Custom digital humans per client; the company entered receivership in Feb 2026 (NZ Herald)

Table 5.5 Business and access models in the ecosystem

WHAT THIS MEANS FOR PRICING

Live avatar time is billed by the minute. On Tavus's published plans, one hour of live conversation costs about \$21, while a private tutoring hour in Tunisia costs 15–20 TND (FTDES, 2023). A per-minute model passed straight to Tunisian students would be hard to sustain; institutional licences or a low-bandwidth voice mode would have to carry most usage.

MARKET SIZE

Grand View Research values AI in education at \$8.3 billion in 2025, rising to \$57.2 billion by 2033 (25.9% a year); higher education is the largest segment. Other firms give higher 2033 figures, e.g. \$75.1 billion ([IMARC](#)), so treat any single number with care.

5.5 Differentiation / gaps

The technology for our idea already exists and is sold by the minute: D-ID and Tavus offer real-time talking avatars that answer from a knowledge base. What does not exist is the educational layer on top — a learner's own professors and verified thinkers, with consent, pedagogy, sources and a human behind them.

Capability	Who covers it today	Coverage
Real-time talking avatar answering from uploaded documents (RAG)	D-ID Agents, Tavus CVI; the BEAI design reproduces it	EXISTS for business use
Digital twin of a real person with consent capture	D-ID, Tavus (on-camera consent), BEAI (checkbox + liveness check)	EXISTS
Socratic AI tutoring by text	Khanmigo, ChatGPT Study mode, Gemini Guided Learning	EXISTS
Voice conversation with a licensed real expert	MasterClass On Call, Delphi	EXISTS but for coaching and creators, not courses
Talking avatar video in 120+ languages	Synthesia, HeyGen, D-ID	EXISTS mostly one-way and pre-scripted
Conversations with historical thinkers	Hello History, Historical Figures Chat	PARTIAL documented factual errors, disclaimers instead of sources
An avatar of the student's own professor, trained on that course	Technically possible with D-ID or Tavus; no education product found	PARTIAL tools exist, no packaged academic service
Pedagogy inside a live avatar (Socratic mode, reaching out to students who don't ask)	None found in avatar platforms; Khanmigo is redesigning to reach out first	NOT SUFFICIENTLY COVERED
Hand-off from the avatar to the real teacher	Only human-only channels (office hours, Preply)	NOT COVERED in the AI tools reviewed
"This is an AI" label, age gate and parental consent in the avatar flow	Not shown in the BEAI design; providers leave it to the customer	NOT SUFFICIENTLY COVERED
Consent process for deceased thinkers	One documented case (D-ID with victims' families)	PARTIAL practice, no standard
Affordable, low-bandwidth, Arabic/French use for North Africa	Live avatars are billed per minute and need bandwidth; language support not verified for education	RESEARCH GAP to verify, not to assume

Table 5.6 What exists, what is partial, what is missing

Differences between approaches

Existing tools differ on three axes. **Who speaks:** a generic AI (ChatGPT, Khanmigo), a licensed living expert or clone (MasterClass, Delphi, D-ID and Tavus twins), or a reconstructed historical figure (Hello History). **How:** text, voice, recorded avatar video, or live video. **From what knowledge:** the open web, the expert's own corpus, a company knowledge base, or a school curriculum. Tutors are strong on pedagogy but faceless; avatar platforms have the face, voice and RAG but no pedagogy; expert clones serve coaching; historical chatbots have the catalogue but not the accuracy. A business model gap sits underneath: avatar platforms bill per minute, tutors per month, and the most expensive bespoke model (Soul Machines) collapsed.

CRITICAL PARAGRAPH ON THE GAPS

Indiana's document makes one thing clear: building a talking 3D avatar that answers from documents is no longer an innovation — D-ID and Tavus sell it, and BEAI copies their pipeline. Our differentiation cannot be the technology; it has to be what education needs and these platforms leave to their customers. The evidence points to four gaps. **First, engagement:** Khanmigo, the most visible AI tutor, was largely ignored by students who did not want to be questioned, and plain chat assistants raise practice scores while lowering real learning — no avatar platform reviewed adds pedagogy. **Second, trust:** historical-figure apps show the appeal of "talking to great minds" and the cost of doing it without sources, while consent for real and deceased people is handled case by case. **Third, compliance:** the BEAI flow has consent and grounding but no AI label, age gate or teacher hand-off, which EU law and UNESCO guidance expect for learners. **Fourth, cost and access:** per-minute avatar pricing (about \$21 an hour on Tavus) and bandwidth needs do not fit Tunisian students or 36% connectivity in Africa. Nothing reviewed lets a student privately ask their own professor's avatar at 2 a.m., in their language, with a sourced answer and the real professor able to step in. That is our position, but it is defensible only if we build that educational layer. We still have no evidence that a 3D avatar improves learning over voice alone; this must be tested in the Act phase, not claimed.

Research gaps

We could not find solid evidence for the points below. They are open questions, not conclusions.

Open question

- 1 Status of Tunisia's data protection bill (PLO 095/2025). Sources conflict: one summary site says it entered into force in July 2026, while a legal review dated Aug 2026 says it is still in committee. Check the [ARP dossier](#) and the INPDP.
- 2 Tunisian rules on voice and image rights, including for deceased people (e.g. a Tunisian scholar's avatar). Nothing found.
- 3 Who is liable for an avatar's answer given in a real expert's name. No specific rule found in any jurisdiction studied.
- 4 ESPRIT's own policy on AI and academic integrity. Not checked; ask the administration.
- 5 US student-record rules (FERPA) if we pilot with Indiana University students. Not researched.
- 6 Whether Denmark's deepfake amendment was finally adopted (expected 1 Jul 2026).
- 7 Copyright term and text-and-data-mining rules in Tunisia for training on thinkers' works and course material. Not verified.
- 8 Tunisian higher-education data on students' unmet support needs. No Tunisian equivalent of the HEPI survey found.
- 9 Evidence that a 3D avatar improves learning compared with voice or text alone. None found.
- 10 Arabic, French and low-bandwidth support in the tools reviewed. Not verified product by product.
- 11 Prices of Khanmigo, Synthesia and Hello History come partly from secondary review sites; confirm on official pages.
- 12 Trust (46%, KPMG/University of Melbourne 2025) and voice cloning (3 seconds, McAfee) come from our Engage deck; the original reports were not reopened for this study.

Added after Indiana University's BEAI document

- 13 BEAI's performance targets (under 500 ms latency, 30 ms lip-sync, 200+ FPS, 3.5 GB GPU per 4 sessions) have no sources in the document. Ask Indiana for the references behind the "[16†20]" markers. Tavus does advertise under 500 ms latency (Fixed Labs), which is consistent.
- 14 "Athena" is named as the conversational engine but never defined. Ask Indiana what it refers to.
- 15 Whether D-ID and Tavus are used in education with measured learning results. We found marketing claims and user profiles, no study.
- 16 Tavus pricing differs between review sites (\$39 vs \$59 for Starter); confirm on the official pricing page.

Open question

17 Whether INPDP would authorise transfers of student and professor data to D-ID (Israel) or Tavus. No decision found.

18 Arabic and Tunisian dialect quality of D-ID and Tavus voices and lip-sync. Not tested.

Sources

All pages accessed on 3 October 2026. Year = date of the document or data.

Law and regulation

Source	Type	Year
European Commission — AI Act page	Official (EU)	Updated Aug 2026
Winss — EU AI Act education deadline deferred	Secondary, citing Commission	Aug 2026
Organic Law 2004-63 (Tunisia)	Official text	2004
INPDP	Official (Tunisia)	—
Global Law Experts — Tunisia data protection bill	Legal commentary	Aug 2026
ARP — bill PLO 095/2025	Official (Tunisia)	2025
GDPR Article 8	Legal text	2016
Loeb & Loeb — amended COPPA Rule	Legal commentary	2025
EU IP Helpdesk — Denmark's deepfake copyright proposal	Official (EU)	Jun 2026
UN News — UNESCO guidance on GenAI in education	Institutional	Sep 2023
GDPRhub — Article 9 GDPR (biometric data)	Legal text and commentary	2016 text, page 2026
Indiana University — BEAI™ Studio, Agent & Digital-Twin Creation Platform (FBM V1A26)	Partner document	2026

Ethics and learning research

Source	Type	Year
Bastani et al., "Generative AI without guardrails can harm learning", PNAS	Peer-reviewed	2025
UCL / UNESCO — bias in large language models	Academic / institutional	2024
Fortune — Character.AI ends open chat for under-18s	Press	Oct 2025
ITU — Facts and Figures 2025	Institutional	2025
Guardian Nigeria — ITU figures for Africa	Press	2025
KQED MindShift — students and Khanmigo	Press, reporting a study	2026

Existing solutions

Source	Type	Year
Khanmigo pricing	Company	2026
Toolchase — Khanmigo	Review site	2026
Maginative — ChatGPT Study mode	Tech press	2025
Maginative — Gemini Guided Learning	Tech press	2025
Design Compass — MasterClass On Call	Press	Dec 2024
ElevenLabs — MasterClass case study	Company	2025
VentureBeat — Delphi	Tech press	2024
Airframe — Delphi analysis	Analyst site	2026
Hello History — for research	Company	—
Curaited — Hello History	Review site	2025
Rolling Stone — Historical Figures Chat	Press	2023
Toolsforhumans — Synthesia	Review site	2026
Costbench — Synthesia pricing	Review site	2026
Gartner Peer Insights — Synthesia	User reviews	2026
Sacra — Preply business model	Analyst site	—
D-ID — AI Agents page	Company	2026
D-ID — Privacy policy (Tel Aviv address)	Company	2026
G2 — D-ID pricing	Review site	2026
Captterra — D-ID pricing	Review site	2026
The Rundown — D-ID review	Review site	2026
Filmora — D-ID review	Review site	2026
Partnership on AI — D-ID consent case study	Non-profit / industry body	2025
The Rundown — Tavus review (consent rules)	Review site	2026
Fixed Labs — Tavus pricing and latency	Review site	2026
Selvia — Tavus cost per minute	Competitor analysis (biased)	2026

Market

Source	Type	Year
UNESCO — Global Report on Teachers	Institutional	2024
Down To Earth — UNESCO teacher factsheet	Press, citing UNESCO	Oct 2023
HEPI — Student Generative AI Survey 2026	Research institute	Mar 2026
Advance HE — HEPI 2026 summary	Institutional	2026
MIT Teaching Systems Lab — The MOOC Pivot (Science)	Peer-reviewed	2019
Times Higher Education — MOOC data	Press	2019
Pieuvre — MOOC completion figures	Press	2019
L'Economiste Maghrébin — FTDES study on school costs	Press, citing FTDES	2023
La Presse — tutoring up to 15,000 TND	Press	May 2026
La Presse — tutoring costs before exams	Press	May 2025
La Presse — back-to-school costs	Press	Aug 2026
Grand View Research — AI in education market	Market research	2026
IMARC — AI in education market	Market research	2025

Source	Type	Year
NZ Herald — Soul Machines in receivership	Press	Feb 2026

06

RESPONSIBLE
Yasmine Nouisser
6 October 2026

Synthesis, Objective Tree & Global Paragraph

Assembled from the findings of Parts 1 to 5 · Part 1 shared by all members

The investigation shows that the technology to give every learner a face-to-face mentor already exists; what is missing is the educational layer on top of it — pedagogy, sources, a learner model, consent, transparency, a human behind the avatar, and an access model that fits Tunisian students.

STATUS OF THIS SYNTHESIS

All parts are now in the report. Part 1, shared by all members, consolidates the evidence of Parts 2–5 and is to be reviewed by the team; Part 2 figures were checked (Table 2.5). The questions still marked **PENDING** in 6.6 need experimental results (EX1–EX7, IS1), not more desk research.

6.1 Collect the findings

Findings are grouped by theme, not by member. Each theme cites the parts that support it.

Theme	Key evidence collected	Parts	Evidence quality
Availability of teachers and experts	44 million additional teachers needed by 2030 (UNESCO 2024); video conferencing removes distance, not the time constraint; experts are costly and limited	3, 5	INSTITUTIONAL
Demand and cost of help in Tunisia	67% of families pay for private lessons; about 1,468 million TND a year; 15–20 TND per hour (FTDES 2023); up to 15,000 TND for some bac candidates (2026)	5	PRESS, CITING A STUDY
Passive online learning	edX completion fell from 6% to about 3% (2013–2018) and second-year return from 38% to 7% (Science 2019)	5	PEER-REVIEWED
Existing tools	26 tools compared on documented evidence and 8 tested hands-on: no tool covers every required capability; educational tools lack persona, voice and avatar, avatar tools lack learner model, pedagogy and identity control	3, 5	HANDS-ON TESTING
Pedagogy and reliability	Unguarded GPT-4: +48% practice, –17% exam; guardrailed tutor avoids the loss (PNAS 2025); AI tutor beat active learning at Harvard (2025); RAG reduces but does not remove hallucinations; ASR can invent words; Khanmigo largely abandoned when it only asked questions	3, 5	PEER-REVIEWED
Language	Vendors claim 100–140+ languages, with no documented quality for dialects; low-resource languages are weaker (NLLB 2022); TounsiBench and ArabCulture-Dialogue rank Tunisian Arabic among the weakest varieties; Latin-script, code-switched Derja is unbenchmarked; no educational Tunisian speech corpus (Part 2)	2, 3, 4, 5	LITERATURE + VENDOR CLAIMS

Theme	Key evidence collected	Parts	Evidence quality
Law and ethics	AI Act Art. 50 (AI label, since Aug 2026); Art. 5 bans emotion recognition in education (since Feb 2025); face and voice templates are biometric data; INPDP authorisation for transfers abroad; minors' consent; no rule on deceased likeness or on liability for an expert's avatar	3, 4, 5	OFFICIAL TEXTS
Cost, scale and access	Tavus about \$21 per live hour; concurrency caps of 3–15 sessions; 42 GPUs (24 GB) for 1,000 students, as a lower bound; 2.2 billion people offline and 36% of Africa online (ITU 2025)	3, 4, 5	MIXED: VENDOR + ESTIMATE
Data and context	No database of consented professor recordings; no educational Tunisian Arabic speech corpus; talking-face datasets (LRS3, MEAD...) are research-only and cannot be used commercially; local course content not available under licence; students' real needs not measured; household data more than 5 years old	2, 4, 5	CHECKED (TABLE 2.5)
Reliability at every layer	RAG legal tools hallucinate 17–33%; Whisper invents text in about 1% of clips, 38% of them harmful; Khanmigo arithmetic errors; watermarks removable and C2PA stripped by all 8 platforms tested	3	PEER-REVIEWED + PRESS
Human in the loop	Tutor CoPilot: +4 pp mastery (+9 pp for weaker tutors) at about \$20 per tutor per year; LearnLM supervised by tutors: 66.2% vs 60.7% transfer	3	RCTS
Trust in synthetic teachers	62% of computing students would not trust AI-presented videos as much; 4 of 6 voice-cloning tools without consent barrier; Arup \$25.6M deepfake fraud; only 46% of people trust AI	3, 5	SURVEYS + PRESS

Table 6.1 Findings grouped by theme

- Every finding above has a source in Parts 2, 3, 4 or 5. Part 3A now provides 42 numbered references with links; Part 2 figures still need to be checked against their sources (see Annex).
- Two figures need a check before the presentation: Tavus pricing (\$59 for 175 min in Part 3, \$59 for 100 min in Part 5, \$39 on some sites) and the BEAI latency targets, which have no sources in the BEAI document.

6.2 What our investigation revealed

Documented facts are kept apart from the team's interpretation. A finding confirmed by several parts is marked as such.

Finding	Documented fact	Interpretation	Parts
1. The barrier is availability, not distance	Video conferencing is mature since 2020 but still depends on the professor being available; 44 million teachers are missing by 2030	A 24/7 mentor addresses the real constraint: time	3, 5
2. Learners already pay for help	67% of Tunisian families pay for private lessons; 95% of UK students use AI, but only 38% get tools from their institution	The demand exists; institutions are not meeting it	5
3. The technology exists, for business	D-ID and Tavus sell real-time talking avatars with a knowledge base; BEAI reproduces their pipeline; they are positioned for marketing, sales and support	The technology is not our differentiator	3, 5
4. No tool combines pedagogy and a face	No tested tool covers every column of the coverage matrix; no avatar platform has a learner model or Socratic pedagogy	The gap is the educational layer	3, 5
5. Design decides whether AI helps	Guardrails avoid the –17% exam loss (PNAS 2025); a well-designed AI tutor beat active learning (2025); students left Khanmigo when it only asked questions	Pedagogy must guide without frustrating	3, 5
6. Language quality is unproven	100–140+ languages are vendor claims; no documented quality for Derja	Arabic and Derja must be measured, not assumed	3, 4, 5

Finding	Documented fact	Interpretation	Parts
7. Law limits what we can sense	Emotion recognition is banned in education (AI Act Art. 5); AI interactions must be labelled (Art. 50); faces and voices are biometric data	Module 5 must rely on conversation, not on the face	4, 5
8. Consent is self-declared today	An Einstein agent was created on D-ID without any authorization; BEAI uses a checkbox and a liveness check	An institutional consent process is needed	3, 5
9. Per-minute pricing does not fit	About \$21 per live hour (Tavus) vs 15–20 TND per tutoring hour; concurrency caps of 3–15 sessions	Institutional licences and a low-bandwidth mode are required	3, 4, 5
10. Teaching value is not yet proven	No evidence found that a 3D avatar improves learning over voice or text alone	The A/B test (EX7) is the key experiment	4, 5
11. The data for Tunisia does not exist yet	No consented professor recordings, no educational Derja speech corpus, research-only talking-face datasets, no measured student needs	Data must be collected with consent, professor by professor, and a student survey is a prerequisite	2, 4
12. Fluent errors are the main risk	RAG 17–33% hallucination; Whisper fabrications (38% harmful); misgrounded citations look credible	A teacher's face makes a wrong answer more convincing: sources, "not in your course" and teacher review are mandatory	3
13. AI works best anchored to a human	Tutor CoPilot +4 pp; LearnLM 66.2% vs 60.7% under tutor supervision; unguarded GPT-4 –17%	The mentor extends the real teacher rather than replacing them	3, 5

Table 6.2 Facts and interpretations

6.3 Global vision / emerging themes

THEME 1 · AVAILABILITY OVER ACCESS

Students have content and connections; what they lack is someone who answers *their* question when they need it.

THEME 2 · PEDAGOGY OVER TECHNOLOGY

Avatars, voices and RAG are commodities. Learning gains depend on guardrails, Socratic guidance and a learner model.

THEME 3 · TRUST BY DESIGN

Sources, an AI label, consent, liveness, a named human owner and no facial emotion analysis are conditions for adoption, not options.

THEME 4 · FIT FOR THE LOCAL CONTEXT

Arabic and Derja quality, low bandwidth and an affordable access model decide whether the students who need help most can use it.

Connections. The teacher shortage (theme 1) explains the demand; commercial avatars answer it technically but without pedagogy (theme 2); a convincing face makes errors more persuasive, so trust must be built in (theme 3); and the most affected learners are those with the least bandwidth and the least-supported languages (theme 4).

Tensions and contradictions that need attention

Tension	What the parts say	How Part 6 handles it
Emotion-aware module vs AI Act Art. 5	Our first objective tree detected the learner's emotion from the face; Parts 4 and 5 show this is banned in education since Feb 2025	Module 5 refocused on conversational signals and learner modelling; the avatar's own expression stays (BEAI Sentiment Control)
Socratic guidance vs engagement	Guardrails protect learning (PNAS 2025), but students abandoned Khanmigo when it only asked questions	Requirement: guide with hints first, without refusing to help
Realism vs deception	Realistic avatars engage; the "Real or AI?" game (SG2) is designed to measure how easily people are fooled	Requirement: visible "This is an AI" label and watermark
Latency targets	BEAI gives < 200 ms end-to-end (Monitoring) and < 500 ms (KPIs), while the LLM first token alone is 200–500 ms	Question to clarify with the supervisor (PP2)
Tavus pricing	Part 3: \$59 for 175 min; Part 5: \$59 for 100 min, \$39 on some sites	Confirm on the official pricing page before quoting
Data locality vs foreign providers	BEAI relies on third-party APIs (Tavus, D-ID in Tel Aviv); Tunisian law requires INPDP authorisation for transfers	Requirement: authorised or local hosting
Tunisian data law status	Sources conflict on whether PLO 095/2025 is in force	Kept as a research gap (Part 5)
Research datasets vs commercial use	Part 4 plans experiments on LRS3, MEAD and similar; Part 2 shows these licences forbid commercial use	Use them for Investigate/Act experiments only; collect consented data for any product

Tension	What the parts say	How Part 6 handles it
Labels vs provenance that survives	AI Act Art. 50 requires labelling; Part 3A shows watermarks are removable and C2PA metadata is stripped by platforms	Requirement: visible label inside the experience plus provenance, not provenance alone

Table 6.3 Tensions between findings

6.4 Structural gaps

Gaps confirmed by at least two parts, or by hands-on testing plus documentation.

What exists today	What is needed	Confirmed by
Avatar agents that answer everything	A mentor that cites the course and says "this is not covered in your course"	3, 5
Pedagogy in text tutors; face and voice in avatar tools	Pedagogy inside the live avatar (hints first, Socratic, reaching out)	3, 5
Tone and language adaptation only	A learner model that adapts explanations to mastery	3
No link back to the real professor	Professor review of answers and student escalation to a human	3, 5
Self-declared consent	University-controlled authorisation, publication and deactivation of a professor's mentor	3, 5
No AI label, age gate or parental consent in the BEAI flow	Transparency and minors' protection in the user journey	5
100–140+ languages announced	Measured quality for Arabic, Derja and technical vocabulary	3, 4, 5
Per-minute pricing and high bandwidth	Affordable, low-bandwidth access (voice or text mode)	3, 4, 5
Analytics on views, credits and latency	Analytics on learning progress and weak topics	3
No study of avatars' learning effect	Measured evidence of teaching value	4, 5
Talking-face datasets licensed for research only; no Derja educational speech corpus	Consented, licensed data collected per professor and per language	2, 4
Labels and watermarks that are stripped or removed	Transparency that stays visible in the experience	3, 5

Table 6.4 Structural gaps

Remaining limitations: no local data on student needs (confirmed by Part 2; survey IS1 pending), Part 2 figures to verify, GPU cost at scale only estimated (PP3), Tunisian rules on voice and image rights not found, liability for an expert's avatar not defined in any jurisdiction studied.

6.5 Strategic conclusions

- 1. Our value is the educational layer, not the avatar.** Talking avatars with RAG are sold today; what is missing is a mentor tied to the student's own course and professor.
- 2. Pedagogy decides learning.** Guidance with guardrails, a learner model and sourced answers matter as much as realism.
- 3. Trust and compliance are design requirements.** AI label, consent and liveness, biometric protection, no facial emotion recognition, minors' protection and authorised data transfers must shape the solution from the start.
- 4. The human stays in the loop.** Each mentor needs a named human owner who reviews content and takes over escalated questions.
- 5. Local fit is a condition of equity.** Arabic and Derja quality, low bandwidth and an institutional access model determine who benefits.
- 6. Value must be measured before it is claimed.** The A/B test (EX7) and the student survey (IS1) are the priority evidence for the Act phase.

These conclusions set requirements; they do not describe the detailed solution, which belongs to the Act phase.

6.6 Guiding Questions coverage

Plan-level questions (Part 1 structure) and the module-level questions of Part 4 are mapped to the evidence available today.

Code	Guiding Question	Evidence so far	Coverage
Plan-level questions (Part 1, Table 1.6)			
Context	How do learners access teachers, experts and resources today?	Part 1 (1.1, Table 1.1); Part 3 (tools); Part 5 (baseline)	COVERED
Problem	What prevents access to quality knowledge?	Part 1 (1.2); Part 5 (teacher shortage, cost, connectivity); Part 3A	COVERED
Causes	Which technical, human, economic and language factors?	Part 1 (1.3, Table 1.3); Parts 3A and 5	COVERED
Consequences	What are the effects of unequal access?	Part 1 (1.4); Part 5 (MOOC drop-out, tutoring costs)	COVERED
Stakeholders	Who produces, transmits and regulates knowledge?	Part 1 (1.5, Table 1.5); Part 5	COVERED
Data	What data exists, and what is missing?	Part 2 (data table, gaps, legal constraints, sources); Part 4 (datasets named)	COVERED
Techniques	Why are current tools not enough?	Part 3A (evidence review, needs × techniques matrix) and 3B (testing, coverage, gaps)	COVERED
Legal & ethics	Which constraints apply?	Part 5 (5.1, 5.2); Part 4 (points of attention)	COVERED
Market	Who needs this, and what already exists?	Part 5 (5.3, 5.4, 5.5)	COVERED
Module-level questions (Part 4)			
GQ1.1	Understand a student speaking Derja, French or both, even in noise?	Part 3 (ASR limits, Koenecke 2024); EX1 and FR1 planned	PARTIAL

Code	Guiding Question	Evidence so far	Coverage
GQ1.2	How much audio to clone a faithful voice?	BEAI: 30–60 s; McAfee: seconds (to verify); EX1 planned	PARTIAL
GQ2.1	Every answer from the course, with its source?	Part 3 (RAG); Part 5 (BEAI test panel with sources)	PARTIAL
GQ2.2	What to do when the answer is not in the course?	BEAI "I'm not sure" fallback; Part 3 (robustness gap); EX3 planned	PARTIAL
GQ3.1	Best balance of realism, lip-sync and real time?	Part 4 (Wav2Lip, SadTalker, MuseTalk, Hallo-Live); EX2 planned	PENDING
GQ3.2	Are BEAI targets (30 ms, 200+ FPS) achievable?	Targets unsourced (Part 5); Tavus advertises < 500 ms	PARTIAL
GQ4.1	Cascade or single model to translate while keeping the voice?	Part 4 (Seamless, UniSS); EX4 planned	PENDING
GQ4.2	What quality for Arabic and Derja?	Part 3A (TounsiBench, ArabCulture-Dialogue, DialectalArabicMMLU); Part 2 (no Derja corpus); EX4 still to measure	PARTIAL
GQ5.1	Adapt expression and tone without analysing the face?	Part 4 refocus; Part 5 (AI Act Art. 5)	PARTIAL
GQ5.2	Do students perceive the avatar's expressions correctly?	EX5 planned	PENDING
GQ6.1	Do generated quizzes and summaries have teaching value?	PNAS 2025 (guardrails); EX6 planned	PENDING
GQP.1	Is the end-to-end latency achievable?	Part 4 (PP2: inconsistent targets)	PARTIAL
GQP.2	Collect, verify and withdraw a professor's consent?	Part 5 (biometric data, Tavus on-camera consent, BEAI liveness); Part 3 (no institutional workflow)	COVERED
GQP.3	Is the user journey clear?	PP4 planned	PENDING

Table 6.5 Guiding Questions coverage — 10 covered, 8 partial, 5 pending

Further research needed: team review of Part 1; the student survey (IS1); measurements from EX1–EX7 for the "Pending" module questions.

6.7 Foundation for the solution

The future solution must meet the requirements below, drawn from the investigation, without being designed in detail before the Act phase.

Requirement	Type	Drawn from
Available 24/7, in private, for each student	MUST-HAVE	Findings 1, 2
Answers grounded in the student's own course, with the source shown, and "not in your course" when needed	MUST-HAVE	Findings 4, 5 · Parts 3, 5
Pedagogy inside the mentor: hints first, Socratic guidance, without refusing to help	MUST-HAVE	Finding 5 · Parts 3, 5
Multimodal interaction: voice, face, text and image combined	MUST-HAVE	Theme 2 · Parts 3, 4
Multilingual by design, with measured quality for Arabic and Derja	MUST-HAVE	Finding 6 · Parts 3, 4, 5
Visible "This is an AI" label inside the experience, plus watermark and C2PA provenance (which alone can be stripped)	CONSTRAINT	AI Act Art. 50 · Parts 3A, 5
No facial emotion recognition of students; adaptation from conversation only	CONSTRAINT	AI Act Art. 5 · Parts 4, 5
Explicit, separate consent and liveness check before any cloning; right to withdraw	CONSTRAINT	GDPR Art. 9, Law 2004-63 · Parts 3, 5
Age gate and parental consent where minors are concerned	CONSTRAINT	GDPR Art. 8, COPPA · Part 5
Authorised or local hosting of student and professor data	CONSTRAINT	INPDP · Part 5
A named human owner per mentor, with review and hand-off to the real teacher	CONSTRAINT	UNESCO guidance · Parts 3, 5
Low-bandwidth mode (voice or text) and an institutional access model	OPPORTUNITY	Finding 9 · Parts 3, 4, 5
Learner model that adapts explanations to mastery	OPPORTUNITY	Part 3 (knowledge tracing)
Learning analytics for teachers (progress, weak topics)	OPPORTUNITY	Part 3
Measured teaching value (A/B test) before any claim	OPPORTUNITY	Finding 10 · Parts 4, 5
Consented, licensed data: professor recordings collected one by one; research-only datasets kept out of the product	CONSTRAINT	Part 2 (2.5, 2.6)
Locally licensed course content as the knowledge base	MUST-HAVE	Part 2 (2.5) · Finding 12

Table 6.6 Requirements for the Act phase

Alignment with BEAI Studio

All six modules map onto BEAI Studio's AI/ML services; three of them extend the platform. The findings of Parts 2 to 5 add one correction (module 5) and four requirements BEAI does not yet cover.

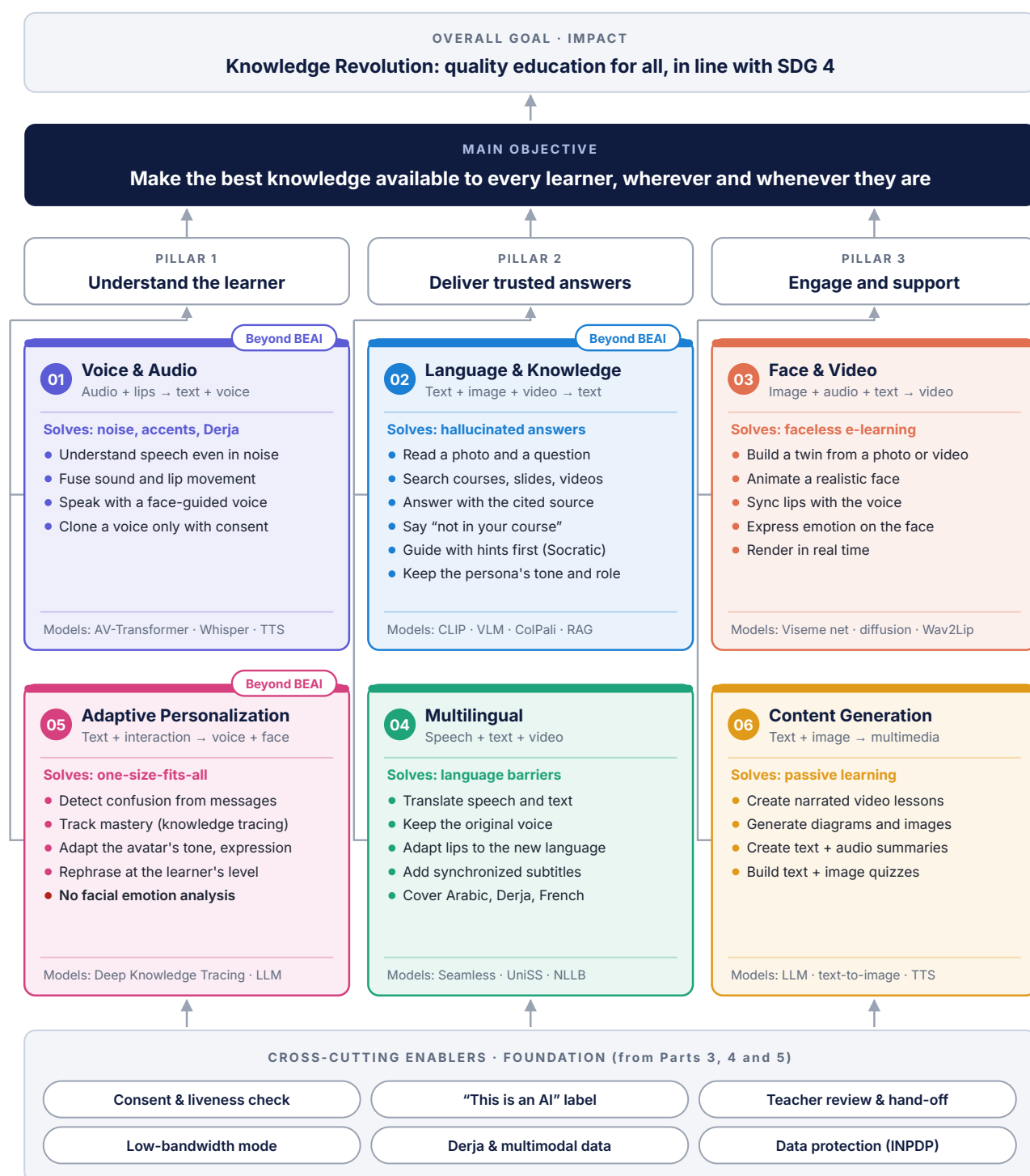
Module	BEAI Studio component	Alignment
01 Voice & Audio	Speech Processing: ASR, TTS, voice cloning	Aligned; lip reading is our extension (camera opt-in)
02 Language & Knowledge	LLM/NLP: custom LLM, RAG, knowledge base, "I'm not sure" fallback, personality settings	Aligned; photo questions and Socratic mode are our extensions
03 Face & Video	Avatar creation from a photo or short video; Vision: viseme-to-frame, diffusion, lip-sync	Fully aligned

Module	BEAI Studio component	Alignment
04 Multilingual	120+ languages, dubbing and subtitles in the Review step	Aligned; Derja quality to be measured
05 Adaptive Personalization	Memory settings, Sentiment Control (avatar expression)	Refocused: learner model from conversation, no facial emotion analysis
06 Content Generation	Video Generation Service, Generative UI (images, videos, quizzes)	Fully aligned
Foundation	Capture & Consent step with liveness check; review before publishing	Extended with AI label, age gate, teacher hand-off and low-bandwidth mode (Part 5)

Table 6.7 Modules and the BEAI Studio reference design

Objective Tree

Read bottom-up: six cross-cutting enablers support six multimodal deep learning modules (intermediate objectives), each with its specific objectives; the modules form three pillars that deliver the main objective and its impact on SDG 4. Every level is linked to the findings above.



Read bottom-up: each level is the means to reach the level above. Beyond BEAI = our extension of the reference platform.

Figure 6.1 Objective tree — 1 main objective, 3 pillars, 6 intermediate objectives, 29 specific objectives, 6 enablers

CHANGES FROM THE PREVIOUS VERSION, DRIVEN BY PARTS 3–5

- **Module 5** becomes *Adaptive Personalization*: confusion is detected from messages, pauses and repeated questions, and a learner model tracks mastery; facial emotion analysis of students is removed (AI Act Art. 5, Parts 4 and 5).
- **Module 2** adds "not in your course" and Socratic hints first (PNAS 2025, Khanmigo, BEAI fallback).
- **Module 4** adds Arabic, Derja and French as explicit targets (Parts 3, 4, 5).
- **Enablers** add the AI label, teacher hand-off, low-bandwidth mode and data protection (Part 5, 5.5).
- **Parts 2 and 3A** confirm the tree: the "Derja & multimodal data" enabler answers the data gaps of Part 2, and Part 3A's evidence (hallucination rates, removable watermarks, human-in-the-loop gains) supports modules 2 and 5 and the label and hand-off enablers.

Global paragraph

Our investigation shows that the main barrier to learning from the best is no longer distance but availability. The world will need 44 million additional teachers by 2030, Tunisian families already spend about 1,468 million dinars a year on private tutoring, and completion of open online courses fell to about 3% on edX, which shows that access to content alone does not keep learners engaged. The technology to answer that need already exists: platforms such as D-ID and Tavus sell real-time talking avatars that answer from uploaded documents in more than 100 languages, and the BEAI Studio design reproduces their pipeline. Yet none of the 26 tools we compared, and none of the eight we tested, combines a mentor's face and voice with pedagogy, a model of what the student knows, institutional control of identity, or a link back to the real professor. The research also shows that design decides whether AI helps or harms learning: unrestricted GPT-4 raised practice grades but lowered exam grades by 17%, while a guardrailed tutor avoided the loss, and students abandoned a Socratic tutor that only asked questions; the strongest gains come from AI supervised by a human, while grounded tools still hallucinate 17 to 33% of the time and speech recognition can invent sentences. Law and trust set firm limits: since 2025 the EU AI Act bans emotion recognition in education, since August 2026 users must be told they are talking to an AI, and a professor's face and voice are biometric data that require explicit consent and, in Tunisia, INPDP authorisation for transfers abroad. Data is the next obstacle: there is no educational Tunisian Arabic speech corpus, no consented professor recordings, and the talking-face datasets are licensed for research only. Cost and access complete the picture, since an hour of live avatar time costs about \$21 while a tutoring hour in Tunisia costs 15 to 20 dinars, and only 36% of Africa is online. The investigation therefore points to an educational layer rather than a new technology: a multimodal, multilingual mentor grounded in the student's own course, transparent about being an AI, built on consent, able to hand over to the real teacher and affordable on a low-bandwidth connection, whose teaching value must still be measured before it is claimed.

Sources: all figures come from Parts 2, 3, 4 and 5 and their source lists (UNESCO 2024, Magesh et al. 2025, Koenecke et al. 2024, Wang et al. 2024, FTDES 2023, Reich & Ruipérez-Valiente 2019, Bastani et al. 2025, AI Act, Organic Law 2004-63, ITU 2025, Tavus pricing).

A

Completeness check for each part

What is still missing, against the deliverables of the Investigate plan

Each item below compares a part with the deliverables listed for it in the Investigate plan. Items marked **DONE HERE** were completed in this report; **BLOCKING** items prevent Part 6 from being finalised (none left); **TO COMPLETE** items are required by the plan; **TO IMPROVE** items strengthen the report.

Part 1 — Context & Problem Structuring · All members **TEAM REVIEW**

Element	Priority
Structured text 1.1–1.6, Context → Problem → Impact schema and Guiding Questions list: consolidated from every member's findings (Parts 2–5)	DONE HERE
Whole team: review the text, schema and Guiding Questions together; each member checks the facts drawn from their own part	TO COMPLETE
Tunisia-specific data on students' unmet needs (no HEPI equivalent found) — to come from survey IS1	TO IMPROVE

Part 2 — Data Study · Maryem Ouichka **COMPLETE**

Element	Priority
All plan sections present (2.1–2.6, data table, data gaps, legal constraints, sources)	COMPLETE
Year of each source and check of the figures (2.7, Table 2.5): 27 consistent, 3 to verify, 1 correction	DONE HERE
Apply the PISA correction in Table 2.2 (Tunisia not in PISA 2022); confirm the Tunisian textbook link, Tunisia's latest TIMSS cycle and the status of bill PLO 095/2025	TO COMPLETE
Original in French — translated in this report	DONE HERE

Part 3 — Techniques, Tools & Limits · Med Rayen Ayat (3A) · Islem Tellili (3B) **COMPLETE**

Element	Priority
Comparison table Technique/Tool · Description · Strengths · Limits · Link (26 tools), "What is missing today" and 42 references with links — provided by Part 3A	COMPLETE
Part 3B reference list completed (Bloom 1984, VanLehn 2011, Corbett & Anderson 1995, Piech 2015, Lewis 2020, Ji 2023, Prajwal 2020)	DONE HERE
Verification flags listed by the author (3A.6): Perplexity error rate, four figures from secondary reporting, vendor claims; re-check the tool table before submission	TO IMPROVE
Tool 3 (Canvas) and Tool 2 notes in 3B are very short; the date of each test is not recorded	TO IMPROVE
Tavus pricing (\$59 for 175 min in 3B) differs from Part 5 (\$59 for 100 min); agree on one figure from the official page	TO IMPROVE

Element	Priority
Two parallel documents merged as 3A (evidence review) and 3B (hands-on testing); 3B numbering harmonised	DONE HERE

Part 4 — Guiding Activities & Resources · Malek Tirellil **COMPLETE**

Element	Priority
Activity · Skill · Why needed · Resource · Type · Link table for the 24 activities (Table 4.20)	DONE HERE
Priority skills list (Table 4.21) and links for the digital media (Table 4.22)	DONE HERE
FR5 benchmark grid pre-filled (Table 4.23); lip-sync ratings and creation times still to measure with trial accounts	TO COMPLETE
Dated schedule proposed (Table 4.24); align with the official presentation date	TO IMPROVE
Activities planned but no result reported yet (experiments, survey, interviews)	TO COMPLETE

Part 5 — Legal, Ethical & Market Study · Youssef Jouini

Missing element	Priority
All plan deliverables are present (solutions table, legal/ethical table, critical paragraph, sources)	COMPLETE
18 open research gaps listed by the author — priority: status of bill PLO 095/2025, INPDP position on foreign transfers, ESPRIT's AI policy, official Tavus and Khanmigo prices	TO IMPROVE
Two figures taken from the Engage deck (46% trust, 3 s voice cloning) to verify in the original reports	TO IMPROVE

Part 6 — Synthesis · Yassmine Nouisser

Missing element	Priority
All parts integrated; update after the team review of Part 1	LATER
Update the coverage table once experiments EX1–EX7 and survey IS1 have results	LATER

MESSAGE TO THE TEAM

Every part of the plan is now covered. Sections marked "Completion" were added to fill the gaps and must be reviewed by their owners: the whole team reviews Part 1 together; Maryem applies the PISA correction and confirms three figures; Malek checks the Activity → Skill table and the dated schedule. What remains is field work: survey IS1, interviews, experiments EX1–EX7 and the FR5 measurements. Module 5 follows the AI Act (no facial emotion analysis of students).

Part 1 is now the only blocking part: without the Guiding Questions list and the Context → Problem → Impact schema, the synthesis cannot be finalised. Part 2 must verify its figures and add the year of each source. Part 4 still needs its Activity → Skill table, priority skills list and resource links. Part 3 is complete. Module 5 has been refocused to comply with the AI Act (no facial emotion analysis of students): anyone preparing slides or simulations on emotions should follow the new definition. Parts 1 and 2 are blocking: without the Guiding Questions list and the data table, the synthesis cannot be finalised. Part 3 needs its reference list and Link column, Part 4 its Activity → Skill table and priority skills list. Module 5 has been refocused to comply with the AI Act (no facial emotion analysis of students): anyone preparing slides or simulations on emotions should follow the new definition.