14TH INTERNATIONAL SCIENTIFIC CONFERENCE “COMPUTER SCIENCE” · SOZOPOL, 13–15 SEPTEMBER 2026

From Stakeholder Needs to Institutional Deployment

Designing a governed RAG assistant for entrepreneurship education

Angel Marinov
Technical University of Varna · a.marinov@tu-varna.bg

with H. Nenov, O. Gungor, R. Montoliu, F. Alniak, S. Franch-Ventura, M. Ivanovici, C. Bungeag, L. Staršelskis, D. Vasilev, E. Mihaylov

IN ONE LINE

A four-country needs analysis (n = 400) → explicit requirements → a retrieval-augmented assistant running on university infrastructure.

Erasmus+ SEEN · BG01-KA220-HED-000255491

Co-funded by the European Union

Mentoring does not scale. Unguided AI already does.

86%

of surveyed students already use AI tools such as ChatGPT

26%

have received any formal or informal AI training

THE INSTITUTIONAL QUESTION

What does the system know — and who decided?

Can its knowledge be inspected and corrected? What happens to user data? Who pays, and what does it cost at scale?

1 Individualised feedback is the scarcest resource

Entrepreneurial learning needs timely, iterative feedback on the learner's own idea. Human mentoring does not scale with enrolment.

2 Public chatbots answer from an opaque knowledge base

No control over sources, pedagogy, data protection or cost. The institution cannot audit, align or defend the tool.

3 The gap is institutional, not algorithmic

The question is not whether LLMs can help, but whether a university can offer AI support it governs.

Co-funded by the European Union

Evidence first: no build decision preceded the survey

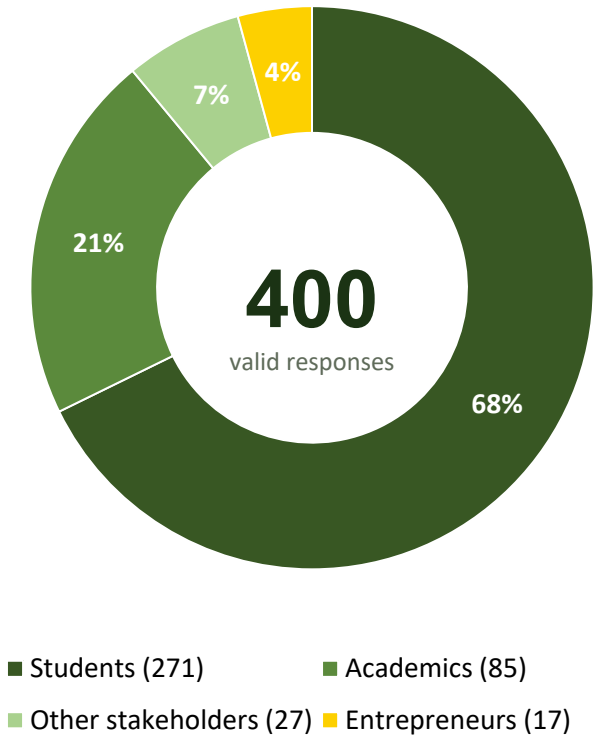


Each stage produced the input of the next.

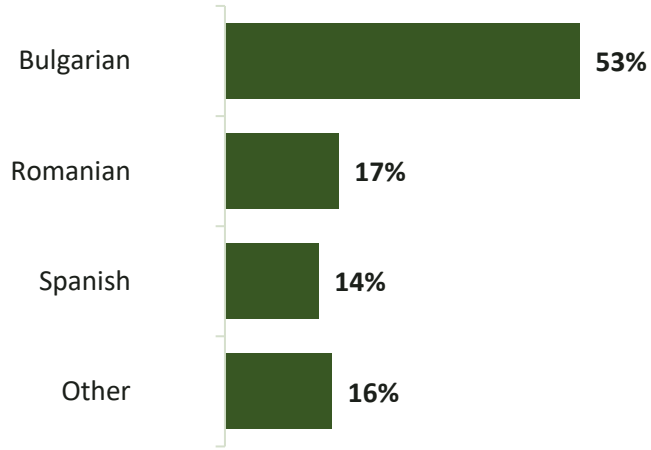
The needs analysis fixed the functional requirements; the architecture comparison fixed the implementation; the functional specification then became the acceptance reference. This ordering distinguishes the project from technology-first pilots that procure a tool and look for use cases afterwards.

Co-funded by the European Union

Four languages, four countries, 400 stakeholders



Nationality of respondents



Convenience sampling through partner channels; the coordinating Bulgarian university is over-represented. Results are descriptive, not nationally representative.

WHAT WAS ASKED

Students

Career intentions · support received vs desired · importance of 4 entrepreneurship and 6 AI topics · AI use, usefulness, training · interest in institutional AI tools

Academics

Same topic and AI blocks with educator wording · willingness and incentives to support student ventures

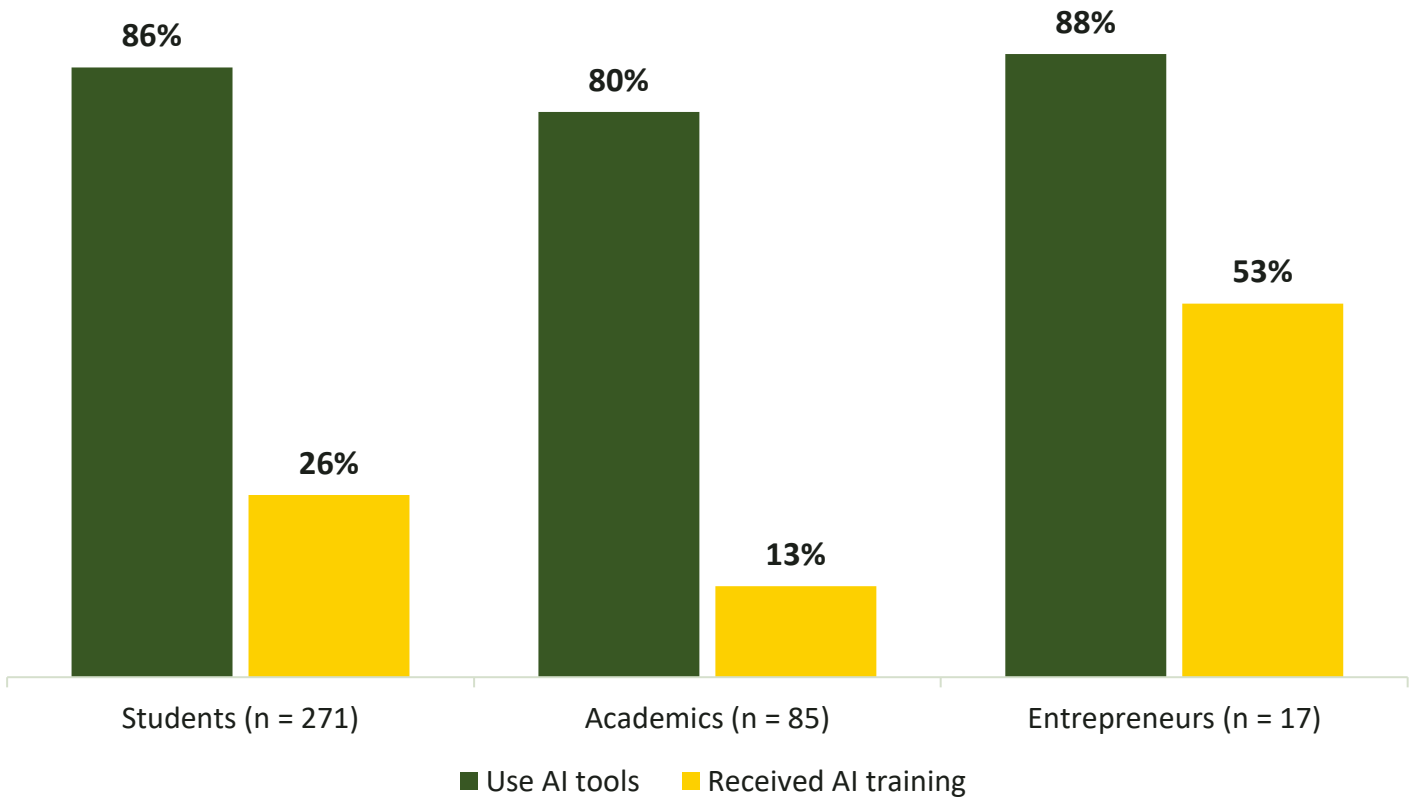
Entrepreneurs & stakeholders

AI in business operations · importance of AI tools and student innovation in higher education

Languages: BG · EN · RO · ES. Instrument public: github.com/astmarinov/SEEN-Surveys

Co-funded by the European Union

Adoption runs far ahead of preparation in every group



Perceived usefulness among student users

3.77 / 5

Academic use counts any of teaching, research or development. A further 33% of academics work in computer science and were treated as domain-prepared.

WHY IT MATTERS FOR DESIGN

A governed assistant responds to an existing, largely unguided practice — it does not create a new one. Guidance must therefore be applied rather than introductory, and verifiable rather than oracular.

Co-funded by the European Union

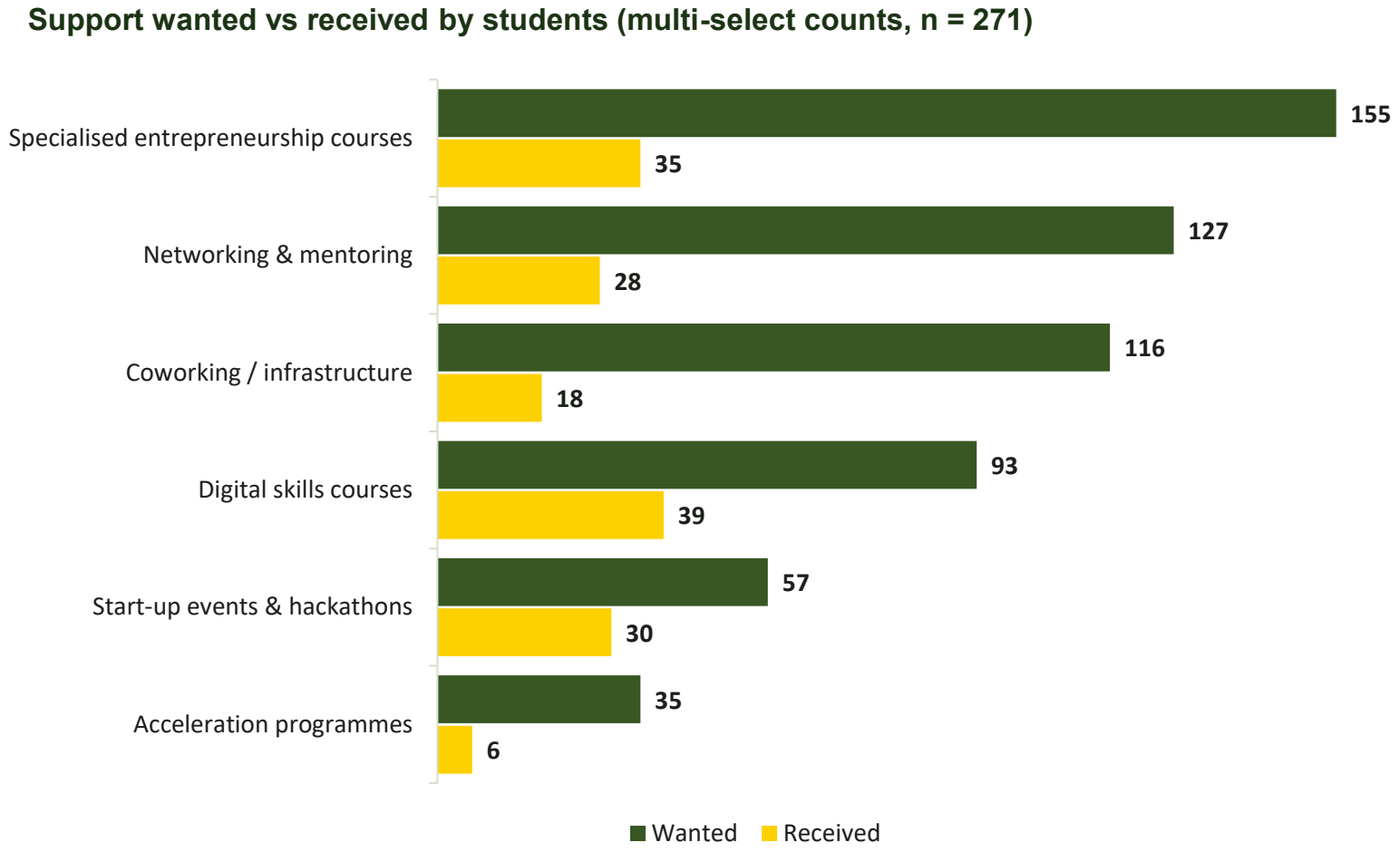
Two-thirds of students get no entrepreneurship support

182 / 271

students reported no support at all relevant to starting a business

THE GAP IS WIDEST WHERE HUMAN ATTENTION IS REQUIRED

Courses, mentoring and coworking are the most wanted and the most attention-intensive forms of support. A tool giving individualised, on-demand guidance at negligible marginal cost addresses exactly this gap.



Co-funded by the European Union

NEED ANALYSIS → REQUIREMENTS

Every design decision traces back to a survey finding

EVIDENCE	REQUIREMENT	SYSTEM RESPONSE
High AI use, little formal training	**Applied guidance with verifiable sources**	**Grounded responses with source provenance**
Two-thirds of students without institutional support	**On-demand, scalable individualised support**	**Five structured entrepreneurship workflows**
Strong interest in institution-provided tools	**University governance and identity**	**Authenticated deployment on institutional infrastructure**
Demand for practical and critical AI literacy	**Structured interaction with visible limitations**	**Stage-specific workflows, constrained prompting, provenance**
Erasmus+ funding ends in 2026	**Inspectable, updatable knowledge base**	**Curated corpus, administrator upload, incremental indexing**

Co-funded by the European Union

Three approaches, five requirement-derived criteria

	APPROACH 1 API only	APPROACH 2 · SELECTED API + RAG over a curated corpus	APPROACH 3 Custom / self-hosted open model
Domain specificity	Low — generic model knowledge	**High — grounded in curated corpus**	High — if trained on domain data
Content governance & provenance	None	**Full — inspectable corpus, cited passages**	Full
Data protection	Provider-dependent	Provider-dependent; no special-category data	Maximal — on-premises
Implementation & maintenance effort	Minimal	Moderate — corpus pipeline, vector index	High — training, MLOps, hardware
Indicative monthly cost	≈ USD 5–60	≈ USD 15–110	≈ USD 200–500 flat

Assessment at time of selection (2025); costs indicative for 300–1000 requests/month. A design decision against project needs, not a controlled benchmark.

Co-funded by the European Union

RAG is an institutional choice, not a technical fashion

PUBLIC CHATBOT	GOVERNED RAG DEPLOYMENT
Opaque, shifting knowledge base	**Curated corpus the institution can inspect and update**
Cannot cite institutional materials	**Responses carry provenance links to vetted passages**
No role-based control	**Authenticated access, RBAC, admin panel**
Data logged under external terms	**Minimal, GDPR-aware logging bound to local policy**
Can change or disappear without notice	**Service evolves only when the institution decides**
Needs ML specialists to adapt	**Maintained by ordinary academic staff: upload a document**

None of this requires the underlying model to be better than the public one. The value lies in the wrapper of governance around it.

Co-funded by
the European Union

DEPLOYED SYSTEM · ARCHITECTURE

Four cooperating layers, wrapped in operational safeguards

Interaction layer
Authenticated web UI · five structured workflows · role-based admin panel (users, corpus) · secure PDF/TXT upload

Orchestration layer
Multi-step agentic flow with dedicated agent roles · input validation · retrieval · grounded prompt assembly around task templates (problem–solution fit, market hypotheses, value proposition, risks, next steps)

Knowledge layer (RAG)
Curated corpus (PDF + TXT) · transformer embeddings · managed vector index (Pinecone) with provenance metadata

Generation layer
Model-agnostic API access under institutional credentials (currently OpenAI GPT-4 family) · scope, tone and grounding constraints · admin-configurable agent behaviour

Operational safeguards
Role-based access control · GDPR-aware minimal logging · secret management via environment variables · cost guardrails at the API boundary · response provenance for audit

The workflows mirror canonical early-venture stages — and they also define the five functional areas of the evaluation.

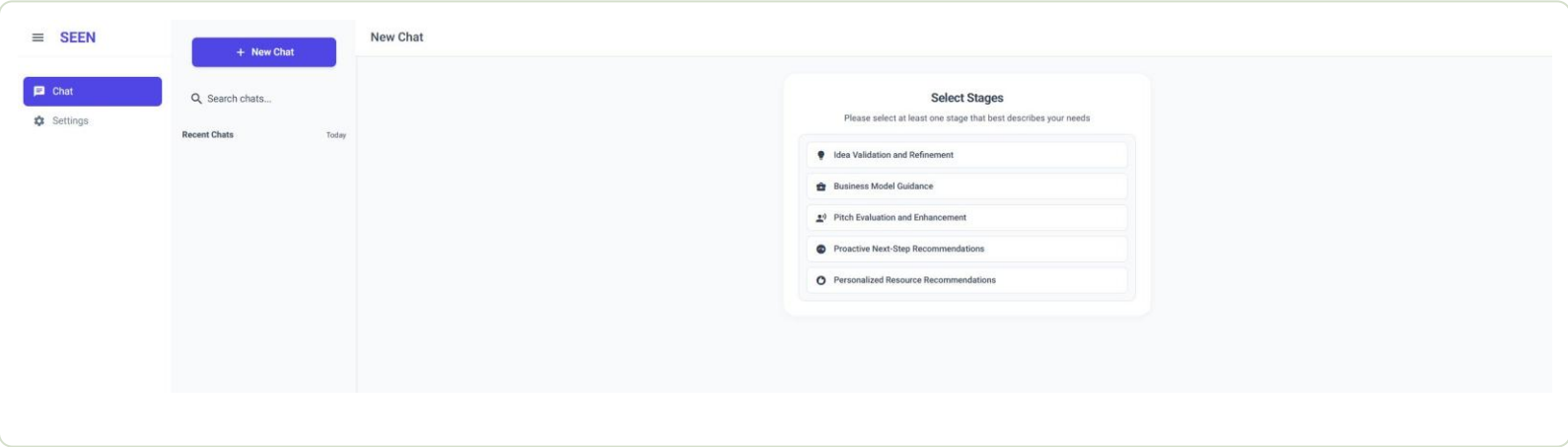
FIVE STRUCTURED WORKFLOWS

1. **Idea validation**
   problem–solution fit, validation planning
2. **Business model**
   structure and revenue logic
3. **Next steps**
   prioritised action planning
4. **Resources**
   matching to learning resources
5. **Pitch evaluation**
   structured feedback on pitch texts

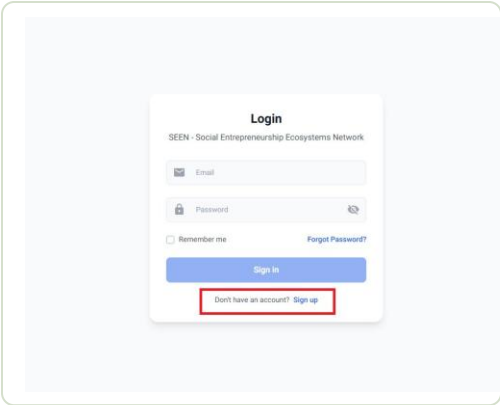
Co-funded by the European Union

DEPLOYED SYSTEM · THE PRODUCT

In production on university servers: seen.tu-varna.bg



Stage selector: the user picks the workflow before the first message — scope is part of the user experience.



DEPLOYMENT PRACTICE

- HTTPS on the institution's official subdomain, behind a reverse proxy
- Role-based access control; institutional credentials at the API boundary
- Minimal, purpose-bound logging; no special-category data; removal on request
- Cost guardrails at the API boundary; secrets via environment variables
- Admin panel: users, roles, corpus upload, index refresh, agent configuration

WHAT A USER GETS

Grounded feedback

a human-readable brief plus a structured object: strengths, gaps, recommendations, referenced sources

Provenance

each response links back to the retrieved passages, so a sceptical student or instructor can check where the advice comes from

Visible limits

stage-specific workflows and a defined no-answer path when the corpus does not contain the evidence

Export for mentor handover

structured feedback can be taken to a human mentor as a first pass

Co-funded by the European Union

CONCLUSIONS

The contribution is the chain, not the model

1. **Design from evidence**
   Every architectural decision traces to a survey finding; the requirements sheet doubled as the acceptance reference.

2. **Governance is the point of RAG**
   An inspectable, updatable, citable corpus and controlled deployment answer the questions a public chatbot cannot: what does it know, who decided, where does the data go.

3. **Governance is affordable**
   Tens of dollars a month at pilot scale; after the project, maintenance is a document collection plus an API subscription, run by ordinary academic staff.

4. **A reusable template**
   Substitute corpus and workflows — legal clinics, thesis supervision, design review — and the architecture and governance model stay the same.

Not claimed: learning effectiveness (needs longitudinal designs) · scalability under high concurrent load · elimination of hallucination — retrieval mitigates it.

SEEN AI platform · seen.tu-varna.bg
Survey instruments: github.com/astmarinov/SEEN-Surveys

Angel Marinov
Technical University of Varna
a.marinov@tu-varna.bg

Co-funded by the European Union (Erasmus+, SEEN). Views are the authors' own.

Co-funded by the European Union