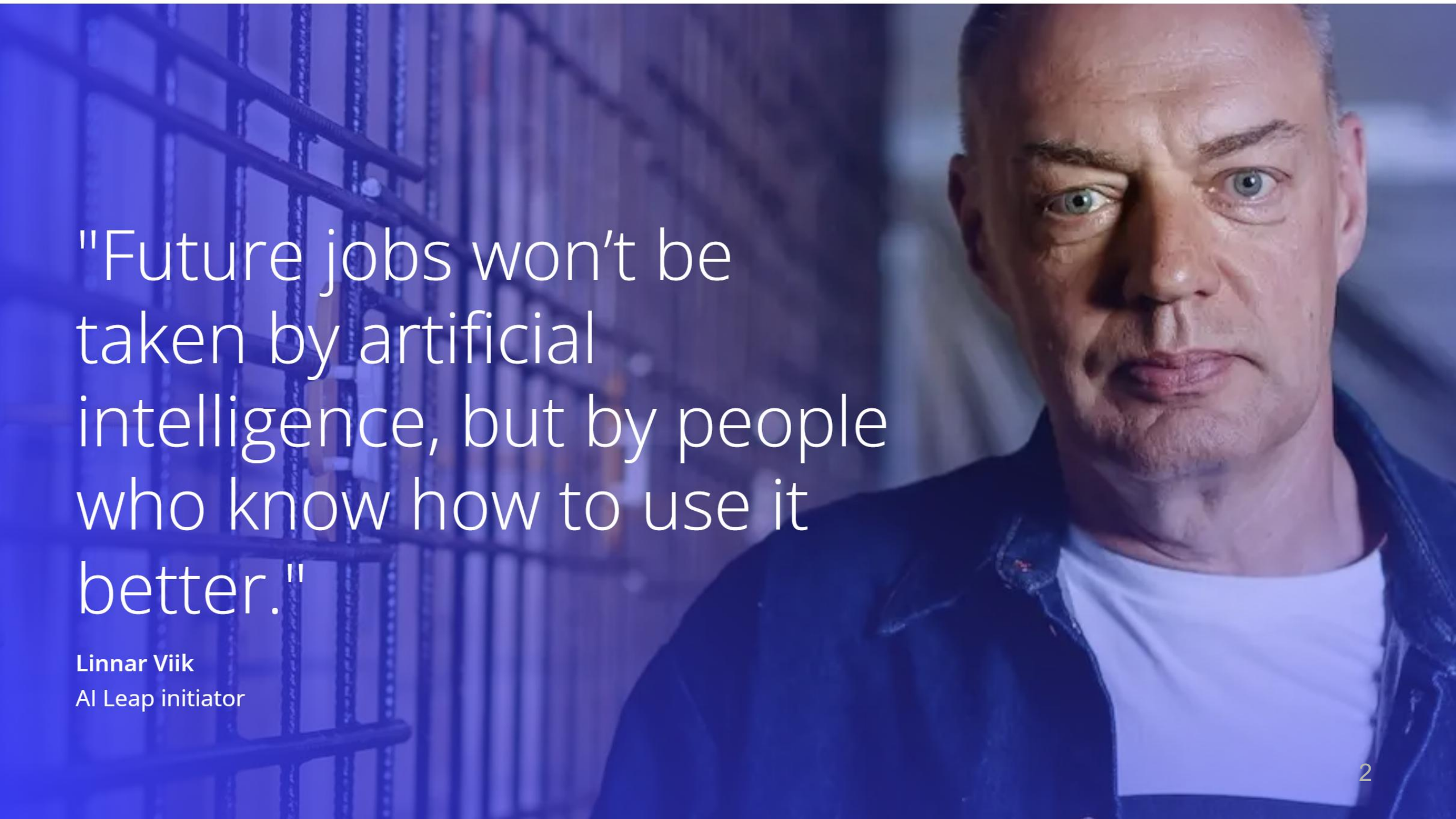
The use of AI tools in teaching written production skills in English among Estonian secondary school EFL teachers

Anneli Sigus, MA
junior lecturer of English and didactics
Liina Tammekänd, PhD
lecturer of English and linguistics
Microsoft Copilot


"Future jobs won't be taken by artificial intelligence, but by people who know how to use it better."
Linnar Viik
AI Leap initiator

Have you managed to avoid AI in the classroom?

Aim and participants of the study

- Overview of the application of AI tools by secondary school English teachers in teaching written production skills

- Two parts:
  - i) an online survey (May 2024)
  - ii) 7 semi-structured interviews (2024/2025 autumn and spring)

- 476 secondary school English teachers (haridussilm.ee)
- 540 invitations to the study
- 79 responses from schools mainly in larger towns (17%)

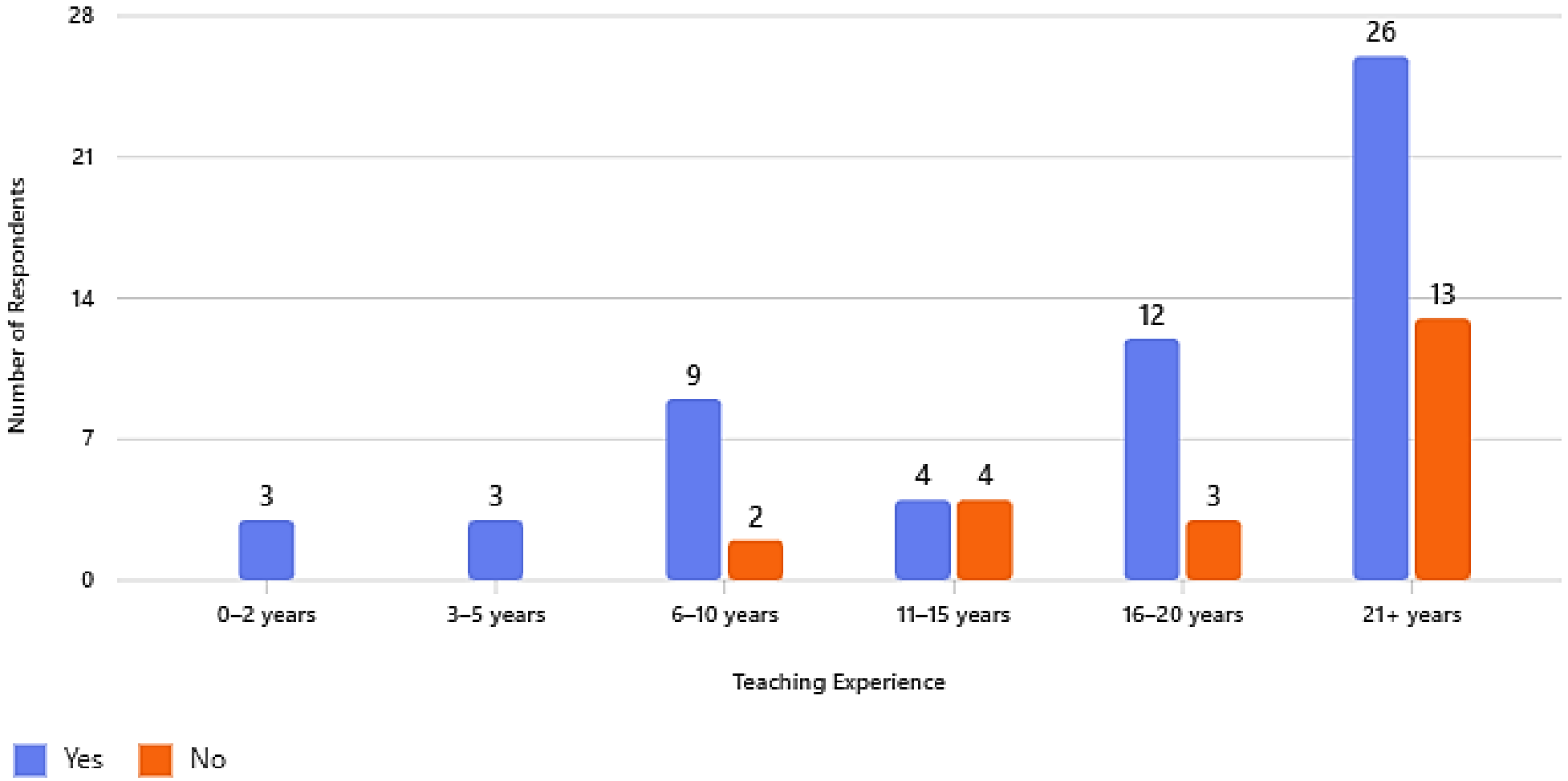
Part I – Questionnaire

Teachers' Workload and Tasks

- 22 contact lessons per week (range 3-34 lessons)
- 16 students per language group (range 9-25 students)

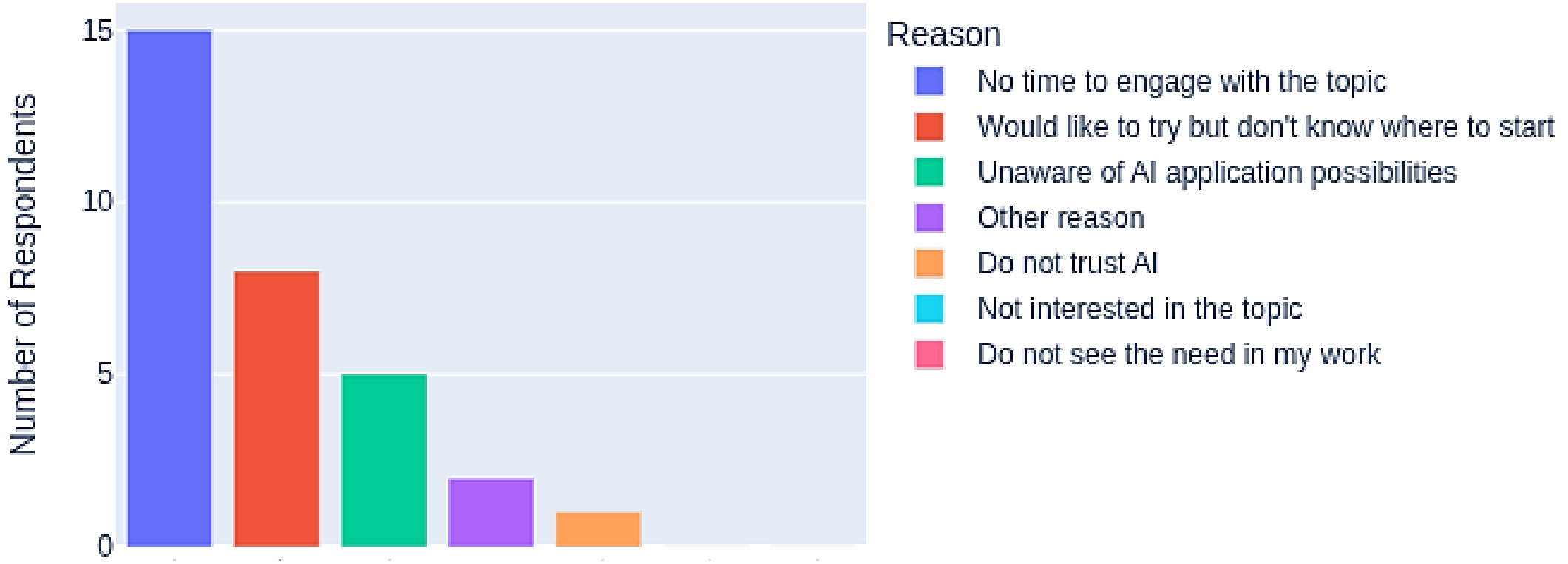
- Additional tasks:
  mentoring, supervising research projects, serving as homeroom teachers, overseeing graduating classes, curriculum development, management and coordination tasks, supporting students with special needs, and more

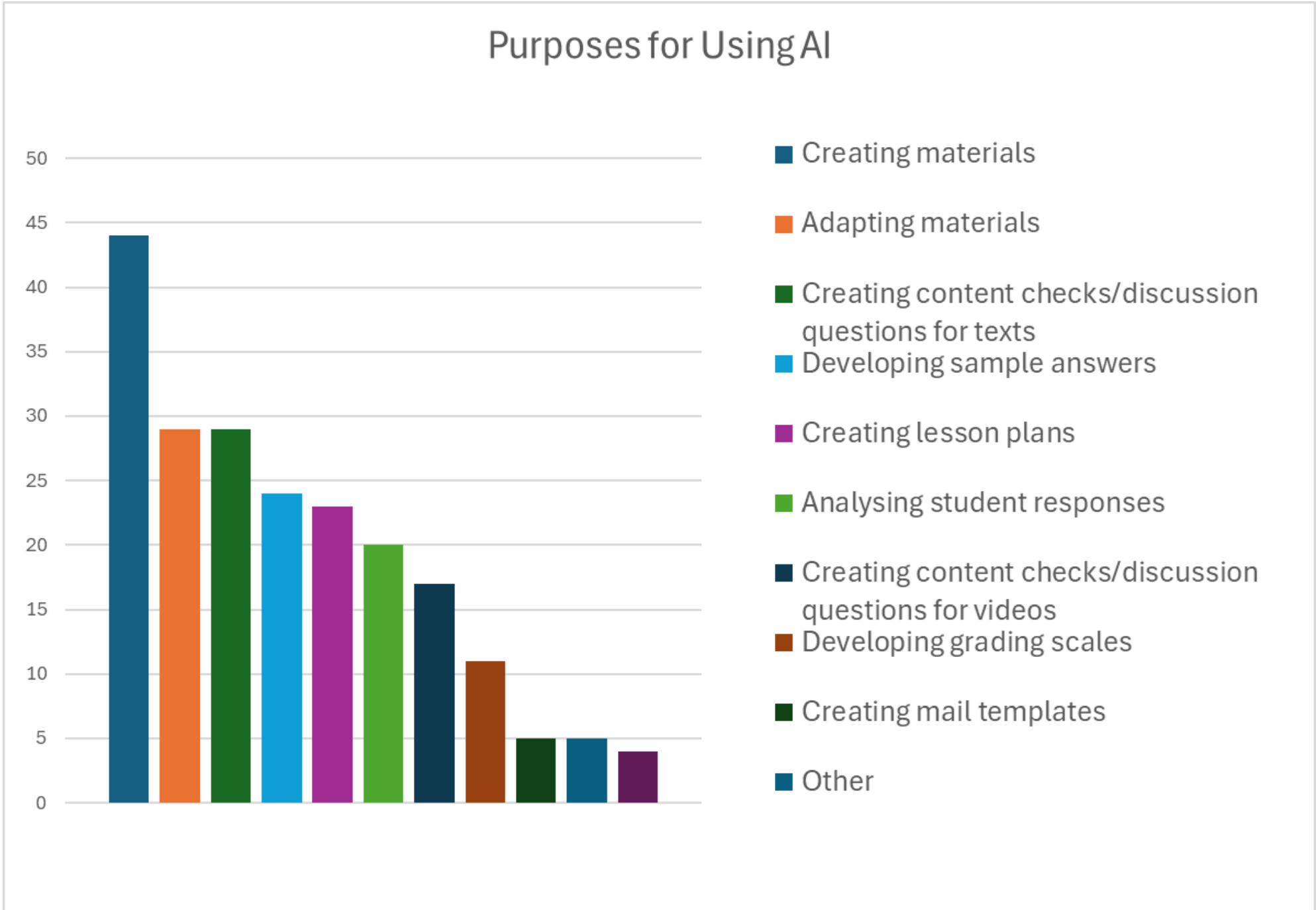
Do you use AI in teaching English? (based on teaching experience)



Reasons for avoiding AI

28% of the respondents do not use AI tools in their teaching.




Purposes for Using AI
50
45
40
35
30
25
20
15
10
5
0
Creating materials
Adapting materials
Creating content checks/discussion questions for texts
Developing sample answers
Creating lesson plans
Analysing student responses
Creating content checks/discussion questions for videos
Developing grading scales
Creating mail templates
Other

AI Use regarding different variables

No discernible correlation between AI use and:

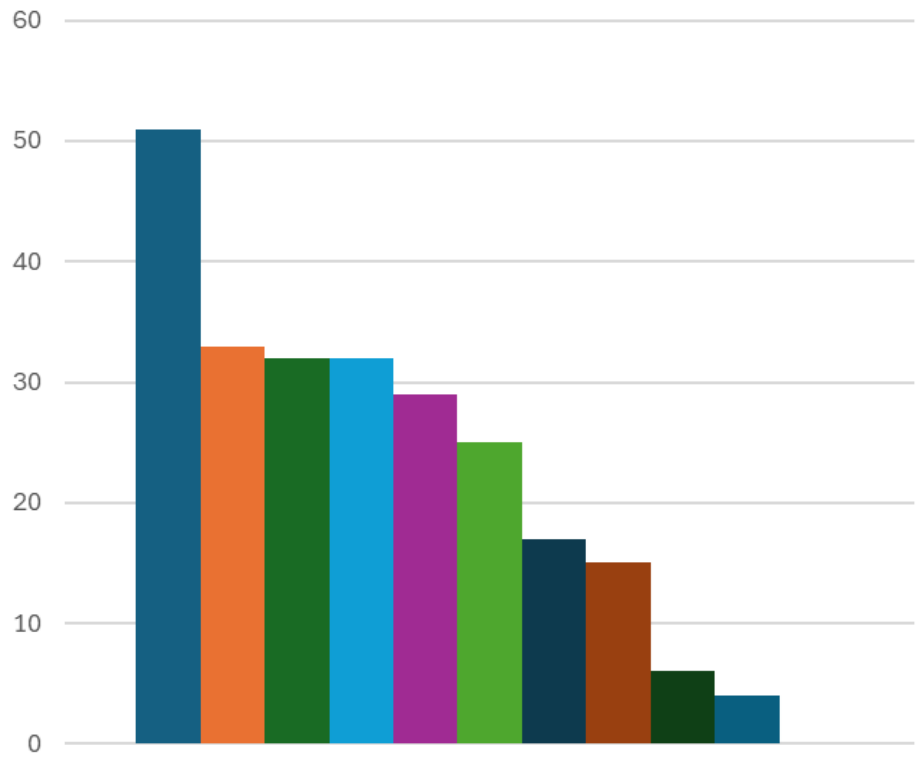
- Location of the school
- Teacher's workload
- Additional tasks
- Group size
- Years of experience (more experienced teachers are *slightly* less likely to report using AI and *slightly* less likely to select restrictive/caution items)

Which tools are used most?

ChatGPT (free version)
Microsoft Copilot

MagicSchool
Twee
Perplexity
Brisk
Canva
Google Gemini

Changes in teaching related to AI



- I teach students to critically evaluate AI output
- I explain to students about the skills they develop by completing tasks
- I use different types of assignments
- I have students write assignments by hand in class
- I explain to students that using AI is academic dishonesty
- I encourage my students to use AI
- I use AI detection programs
- I have changed assessment criteria
- Other
- I collect a language sample from students
- AI has not influenced my teaching in any way

How do these results compare to your experience?

Part II - interviews

Qualitative part of the study

- 7 semi-structured interviews (avg. 1 hour 9 mins)
- We wanted to know:
  a) How have teachers used AI tools in EFL classes?
  b) What are teachers' perceptions of the usefulness of AI in teaching written production?
- Interview questions composed based on the questionnaire results
- Inductive content analysis to determine key attitudes and experiences with AI tools and to identify overlapping points

What leads teachers to use AI tools?

- Personal interest in ICT tools

  Int4: *“Using AI tools depends only on the teacher’s personality - not every teacher is ready to try it.”*

- Students’ use of AI

  Int5: *“Students are afraid to admit they use AI, but I want there to be open discussions about it.”*

  → ⅔ of high school students in Estonia use AI (Norstat 2024)

- Perceived need to teach students to use AI effectively

  Int4: *“School is a safe place to experiment with AI.”*

General use of AI tools (our analysis)

- Mainly positive experience - planning and brainstorming
  - Lesson plans and course syllabi
  - Structuring the teaching process
  - Comparable, but different tasks for students

- Mainly negative experience - creating study materials (not according to Copilot!)
  - A lot of time and effort needed to get usable materials

- Potential benefits:
  - AI mistakes as learning opportunities
  - inclusive education and differentiated teaching

Using AI for teaching written production

- **Student must be the creator of the original text**

  Int3: *“It is a part of learning. Creating [the text] yourself is thinking.”*

- **Teaching “critical evaluation” of AI generated content highlighted - mostly done by comparison and discussion**
  - student’s own text and AI-generated sample
  - student’s own text and the same text improved by AI

- **Better experiences with specific AI tools rather than generative AI**
  - Monitoring the writing process with AI tools (e.g. Brisk Teaching)

AI for formative feedback on writing

- **AI feedback is lengthy, but not specific enough to be useful**

  Int1: "*It's so generic, it's pointless. It doesn't help the student develop.*"

- **Lack of students' engagement or skill in reading the feedback**

  Int3: *"Before we can let students use AI for feedback, we need to teach them how to interpret it."*

- **AI tends to give directive feedback or corrects the text itself**

- Potential: AI feedback on language issues;
  the teacher can focus on higher-order skills.
- Problem: The teacher needs to know their students' progress

  Int1: *"If I don't read their papers myself, I am letting my students down."*

AI for assessment of written production

- **Assessment is not transparent**

  Int2: *“some essays got a B and some an A, but based on what - I don’t know.”*

- **AI can include (positive) aspects that the teacher might not notice**
- **Without special training, AI gives average scores**, even with uploaded marking scales
- AI was **more generous** in marking than the teacher.
- Ethical issues and data privacy.
- The question of the **teacher’s authority** and respect.

  Int7: *„The students asked right away, why I could use it and they weren’t allowed to.“*

- **Potential: AI for creating marking scales**

Analysis done with AI (Copilot)

🗺️ Thematic Map Diagram

- 🧰 Workload Changes
- ✍️ Writing Process Integrity
- 👩‍🏫 Teacher Oversight
- 📚 Exam Preparation
- 🔐 Ethical Use & Literacy
- 🚀 Future Skills & Training

✅ Areas with Mainly Positive Experiences (Copilot)

Lesson Planning & Material Creation - Brainstorming lesson ideas, generating warm-ups, comprehension questions, and assessment rubrics.

Grammar & Vocabulary Support – Tools like Brisk and RoadtoGrammar help students identify errors and improve lexical range.

Formative Feedback - AI provides instant feedback, motivating students and reducing teacher workload for surface-level corrections.

Idea Generation for Writing - Students use AI to brainstorm essay topics or arguments, especially for advanced learners.

❌ Areas with Mostly Negative Experiences (Copilot)

AI-Generated Assessment Items - Multiple-choice questions from AI were often poor quality (implausible distractors, too text-based).

AI as a Speaking Partner - Attempts to use AI for oral exam practice often failed because bots dominated the conversation or misunderstood prompts.

Automated Grading - AI grading with broad rubrics produced generic feedback and averaged scores, missing task-specific nuances.

Lower-Level Learners - For beginners, AI suggestions were too complex or confusing, requiring heavy teacher mediation.

Hallucinations & Inaccuracies - Teachers encountered factual errors and odd translations, which required extra checking.

Overall remarks

- Teachers feel they need more time to experiment with AI
- AI practice activities should not be assessed (*“learning for yourself”*)
- AI for writing assessment is difficult to achiece (needs knowledge, input and training supervision)
- Lower-level students seen as benefitting less from AI in writing skills
- Open discussion valued for both learning and relations with students
- Final exams used as a motivator for personal need *(„AI won’t be there to help you at the exam*“)

Discussion

What to do with this data?

What should we do with AI so that teaching English and teaching writing skills would benefit from it?

Recommendations from Copilot (1)

1. **Curate & integrate**

   Publish a short, **subject-specific toolset** (e.g., planning: Perplexity/Magic School; language: RoadtoGrammar; feedback: Brisk) and, where possible, **integrate** launchers in your LMS (Teams/eKool).

2. **“In-class for marks; at home for practice” policy**

   Keep assessed writing **in class**; allow home use of AI for **idea generation and language polishing**, with **disclosure**.

3. **Critical AI literacy sequence (3 mini-lessons)**

   Prompting 101: Role–Audience–Criteria–Constraints–Tone; iterate.

   Verify: Compare AI claims against **cited sources** (Perplexity links, Wikipedia with references).

   Ethics & privacy: What’s acceptable help vs authorship; minors’ use; don’t upload peers’ work into public AIs.

Recommendations from Copilot (2)

4. Adopt a two-tier marking workflow

Student phase: require **self-revision** using approved tools (e.g., Brisk, RoadtoGrammar) + short **revision log**.

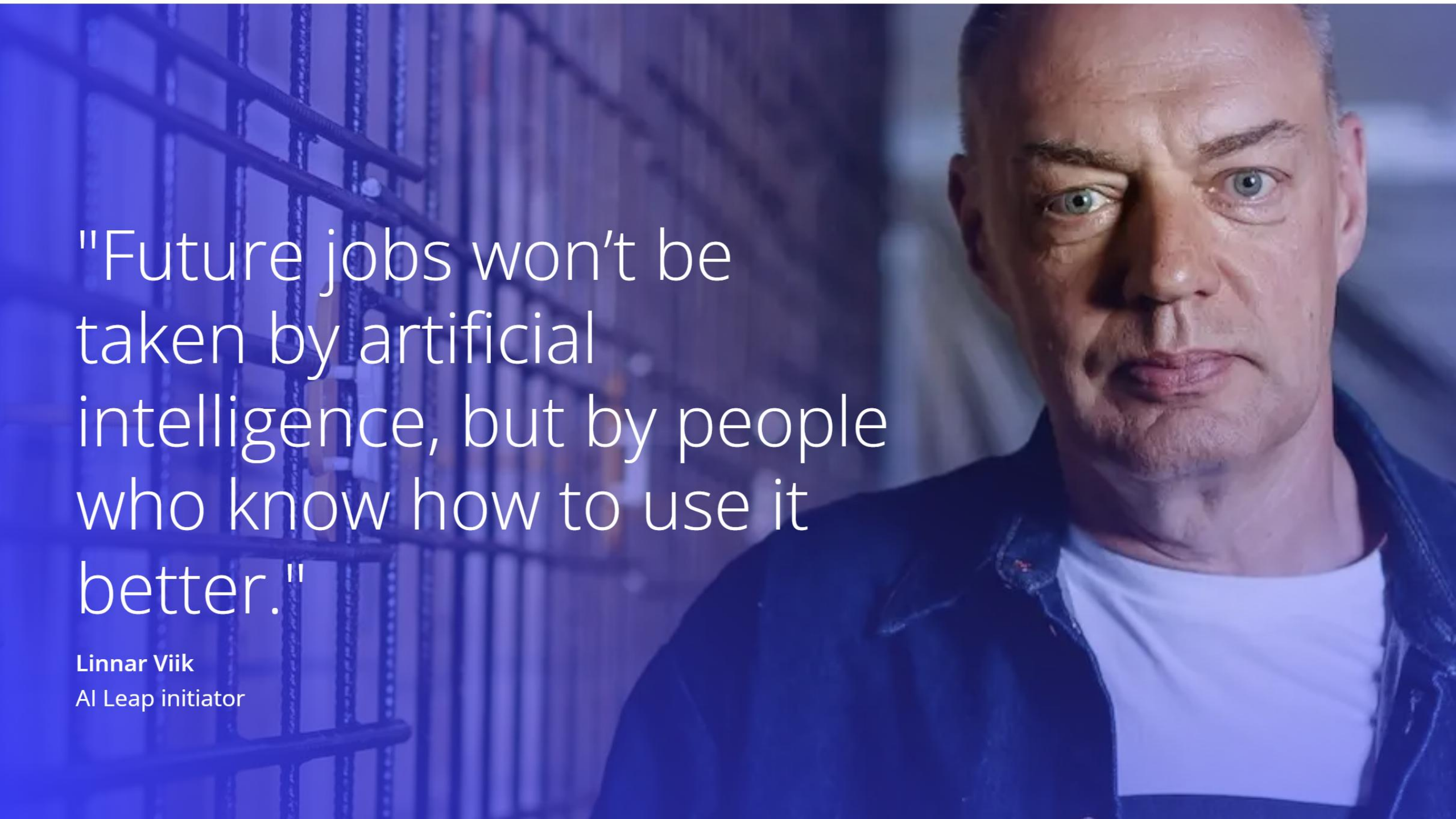
Teacher phase: focus time on **content/structure**; use **task-specific rubrics**; optionally let AI pre-fill rubric notes **after** seeding with **exemplars**—then you finalize.

5. Accessibility by default

Normalize **text-to-speech**, **auto-captions**,and **screen-reader-friendly** handouts; consider AI-assisted **description** for visuals.

6. Avoid auto-generated MCQs for summative use

Use AI to **draft** items but **hand-craft** distractors, or focus AI support on **open/constructed** tasks where its strengths (idea surfacing, rephrasing) help learning without undermining validity.


"Future jobs won't be taken by artificial intelligence, but by people who know how to use it better."
Linnar Viik
AI Leap initiator

The Big Picture (1)

Goal? What do we want them to learn about writing?

- „writing“ is not just words connected into sentences, but a way to express one’s thoughts and ideas in a structured manner.
- skills are developed only through repeated practice

1. Focus on the writing **process and message**, not the end result.

The Big Picture (2)

Kids are feeling isolated and have social anxiety ...

And we want them to talk MORE with software algorithms?

2. Keep in mind that the point of language classes is to teach students to communicate well in English.

- person-to-person interaction
- **communicative purpose** of a task (also with AI)

→ safe space to communicate

→ discuss difficult topics

→ practise respect and consideration

Look into the scales of the Common European Framework of Reference (CEFR) Companion Volume for diffrerent purposes of writing and competences connected to it.

Language activities (general purpose of the communication):
- written production (pp 66-70)
- written interaction (pp 81-89)
- written mediation (pp 90-121)

Linguistic competence (language used in the communication):
Chapter 5 (pp 129-142)

Thank you!

References

Laak, K. J., & Aru, J. (2024). Generative AI in K-12: Opportunities for Learning and Utility for Teachers. In *International Conference on Artificial Intelligence in Education* (pp. 502-509). Cham: Springer Nature Switzerland.

Li, Yingxia, Wanyi Du, Bike Zhu. Cognitive Engagement of Low Proficiency Learners: Evaluating the Impact of ChatGPT Feedback in EFL Writing. *Frontiers in Educational Research* (2023) Vol. 6, Issue 28: 69-73. https://doi.org/10.25236/FER.2023.062809

Marzuki, & Widiati, Utami & Rusdin, Diyenti & Darwin. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*. 10. 10.1080/2331186X.2023.2236469

Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, *2*(2), Article 100050. https://doi.org/10.1016/j.rmal.2023.100050

Norstat 2024 = Samsungi tellitud uurimus ideekonkursi Solve for Tomorrow raames. 2024. Kättesaadav: https://www.solvefortomorrowbaltics.com/et/uuring-just-nendes-oppeainetes-saavad-noored-tehisintellektist-enim-abi/

Grading the Graders: Comparing Generative AI and Human Assessment in Essay Evaluation - Elizabeth L. Wetzler, Kenneth S. Cassidy, Margaret J. Jones, Chelsea R. Frazier, Nickalous A. Korbut, Chelsea M. Sims, Shari S. Bowen, Michael Wood, 2025

Image from: The AI Leap homepage TI-Hüpe (accessed on 22 August 2025)