International Council on Systems Engineering
A better world through a systems approach

ChatGPT Dilemma: The Effect of Generative AI on Higher Education in Systems Engineering

Emin Simsek, Gerrit Muller, Kristin Falk-
University of Southeastern Norway

INCOSE International Symposium 2025 | Ottawa, Canada

Abstract

This study examines the effects of ChatGPT on higher education in systems engineering. It focuses on how it influences student learning and academic honesty before and after the introduction of ChatGPT 3,5. Comprehensive research is limited in the literature. The research uses surveys, experiments, and case studies to understand the role of AI in education from a systems-thinking point of view. The results show that most students use AI tools in their homework, and it helps to improve grades. It also brings issues like plagiarism, critical thinking, and less effort during exams. The study points out the need for clear guidelines to provide responsible use of AI and sustain competency development. It may assist students and teachers in their learning and teaching processes and help them make ethical use of AI and find efficient ways in their professional lives.

Problem Statement

Primary Concern: Concerns regarding university academic integrity regarding using generative AI for home exams.

Academic Integrity: Teachers still determine if students' competence development remains the same when AI tools are used for home exams.

Competence Development: Some teachers are concerned that AI use may weaken critical thinking and problem-solving skills.

Problem Verification

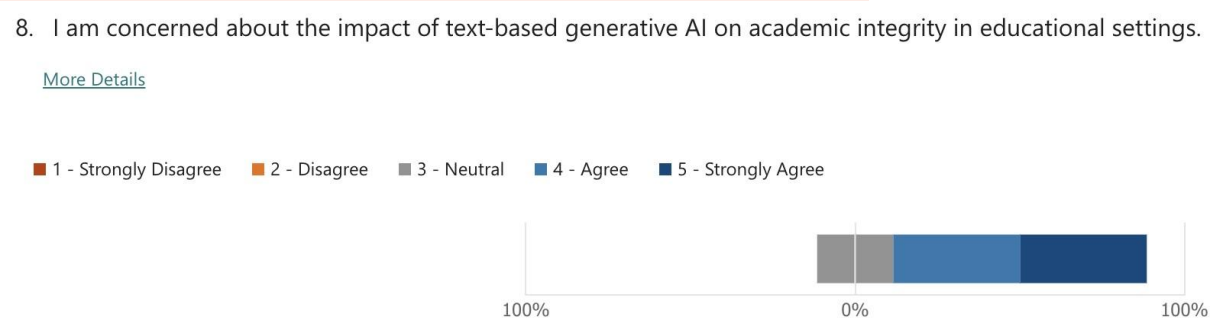
According to the teachers, Universities should not restrict the use of generative AI for courses and exams; controlling and supporting it can be necessary. 46% believe that controlling should be.

Also, 38% do not believe the university has a clear statement for using generative AI.

84% believe additional support is needed during the integration phase of the generative AI.

76%believes that universities should be able to verify the usage of GenAI for the exams.

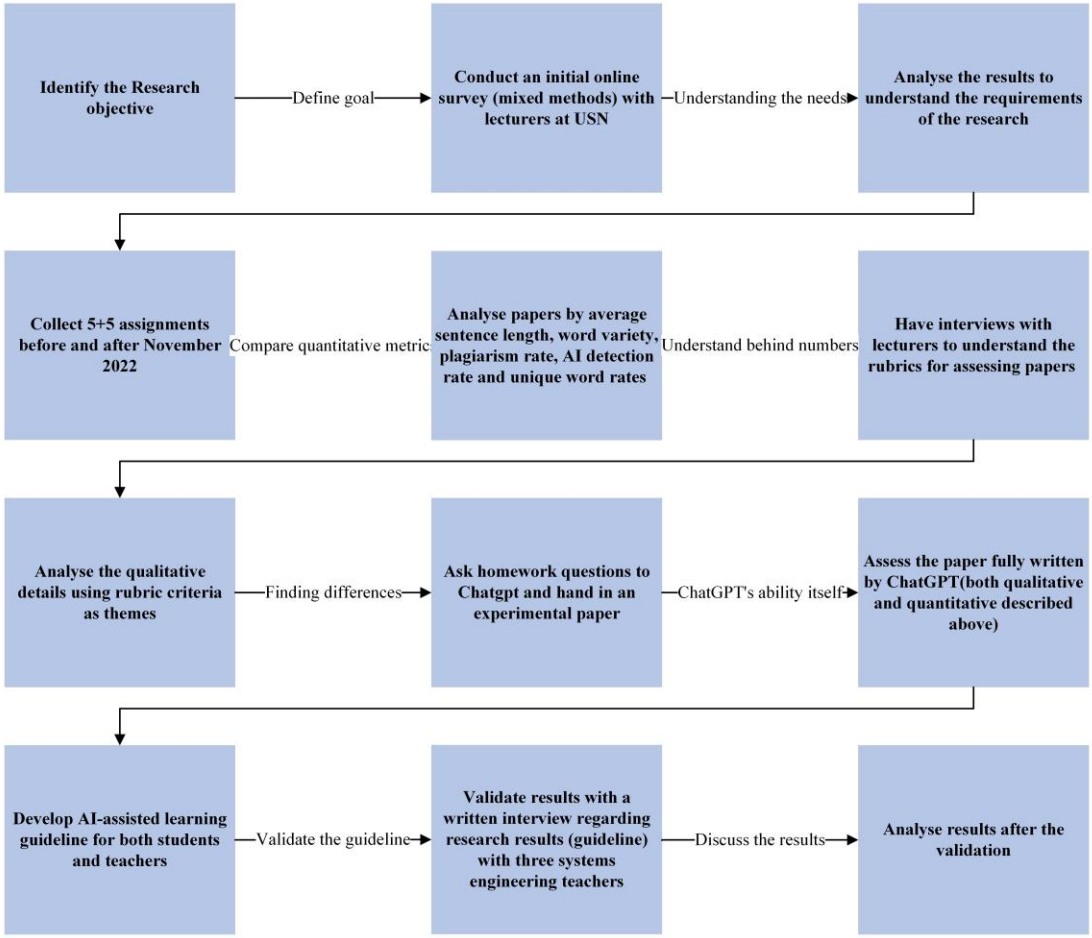
.



Stakeholders

Stakeholder	Role in Research
University of Southeastern Norway Management	Policy-making, AI guidelines development
Teachers (Systems Engineering & Innovation)	Teaching, assessment, and implementation of AI guidelines
ITM4200 Course Teacher	Course design, AI integration, evaluation of student work
Students (2022 & 2023 Systems Engineering)	Primary users of AI tools, source of data for the research
Generative AI tools (e.g., OpenAI, GPTZero)	Provide AI tools, ensure privacy and ethical safeguards

Methodology



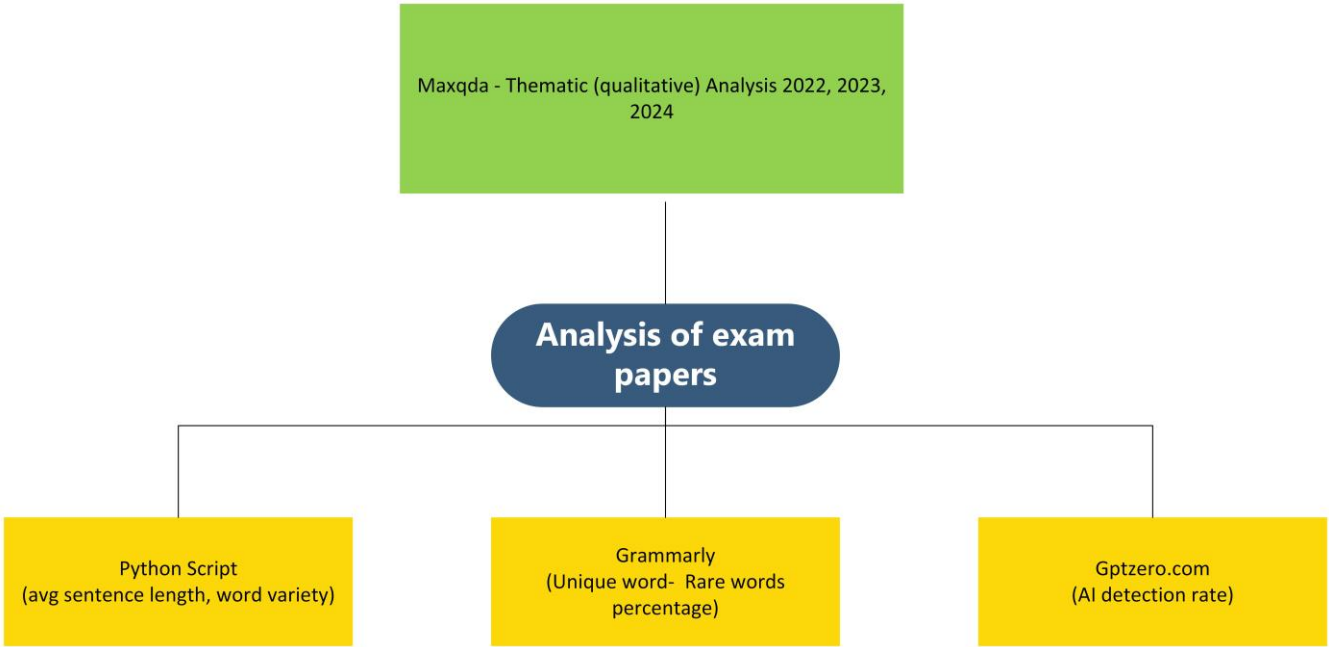
Surveys: Conducted with academic staff from the Systems Engineering program.

Case Study: Collected final exam papers from 2022 (before ChatGPT 3.5) and 2023 (post-ChatGPT 3.5).

Experiment: An experimental exam paper was created using ChatGPT-4 for comparison.

Thematic Analysis: Maxqda codes themes to compare changes in student exam papers over time.

Research Analysis



Research Risks

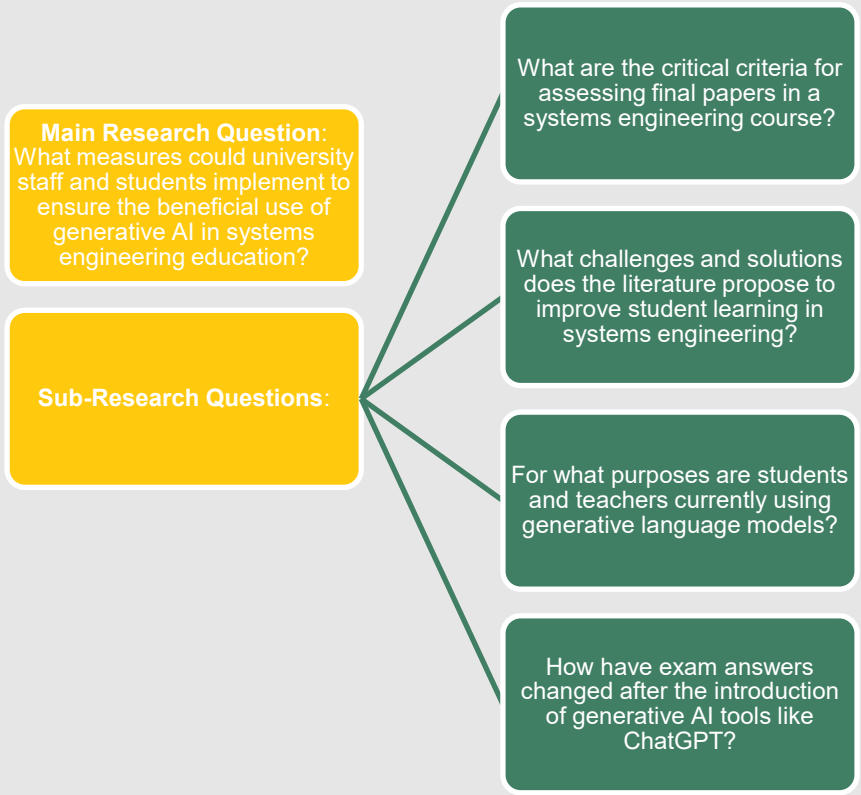
1. **Accuracy of AI Detection Tools:** AI plagiarism detectors still need to be fully reliable.

3. **Rapid Evolution of AI:** The rapid development of AI tools risks the validity of this research.

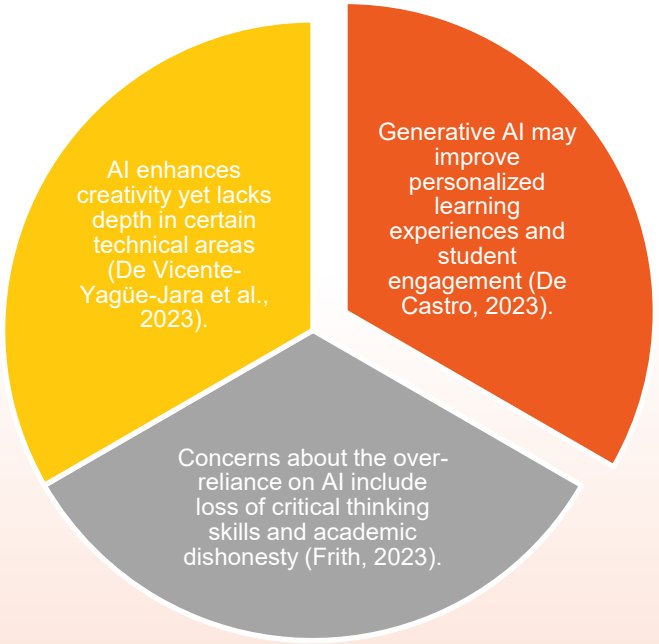
4. **Narrow Research Context:** The limited academic context (University of Southeastern Norway, Systems Engineering).

5. **Confidentiality Concerns:** Although confidentiality wasn't a primary concern for students or teachers in the study, there is an underlying risk.

Research Questions



Literature Review



Literature Review

Reference	Challenges	Solutions
De Castro (2023)	Concerns over privacy and academic integrity and data manipulations.	AI used must be clearly declared. However, any use of AI is not breaching the Cambridge plagiarism policy.
Frith (2023)	Potential underdevelopment of student skills, overusing, accountability, and failure of current plagiarism detection programs.	It is suggested that students need to have more time to understand the algorithm and use it effectively.
Jungherr (2023)	Misinformation and limited temporal coverage.	Asking questions from general to specific, use theory rather than a literature review.
Walczak & Cellary (2023)	AI systems disrupt traditional education. Also, they have Originality issues and "hallucinations"	Students should develop new skills like AI- Human communication, critical thinking, etc. Because they will replace the jobs that they are preparing for.
Shaw, Morfeld & Erren (2023)	Ethical considerations, misuse and insufficiency of the ChatGPT and competitive advantage of using it.	ChatGPT should be able to detect, or assessment style may change rather than giving an essay exam. However, common policy is critical for equity.
Alqahtani et al. (2023)	Being not comprehensive in literature review, lack of complexity and ethical considerations.	Using AI for assisting, accelerating and facilitating for the literature, text generation and data analysis.
Adar & Kandemir (2008)	Concern about academic integrity and ethics	Enhancement of soft skills in the education process
Su, Scarinci & Cicirello (2023b)	Challenges in integrating LLMs into systems engineering	Use in automatic generation of systems engineering diagrams from the written texts for the requirements
Bower et al. (2024)	Unconscious effect of generative AI on education: teaching motivations and adaptions	Encouraging authentic tasks and real-world application in teaching
Eager & Brunton (2023)	Integration of AI tools to develop future skills and how to command effective?	Emphasis on industry-specific skills and academic skills in assignments and developed a command process for AI.
Joshi et al., (2023)	ChatGPT does not consistently provide accurate explanations and answers. (Sufficiency of AI)	Giving the right prompt may increase the accuracy. Also, students must use as an assistant.
AlAfnan et al., (2023)	Lack of details. Too general answers. Informality of the learning from ChatGPT and academic integrity.	Take home exam style and teacher´s assessment rubrics may change.
Szefer & Deshpande (2023b)	Students are using AI chatbots but there can be some limitations.	ChatGPT may be used as a supplementary tool, teachers may apply Laboratory-based or figure-based questions if they are willing to avoid them.
Ngoc et al., (2023)	Integration of AI assisting learning	The diagram introduced for the use cases and purposes.
De Vicente-Yagüe-Jara, M. et al., (2023).	LLM´s sufficiency and qualitative impacts	Analyzed the originality creativity and flexibility by using creative imagination test.

Analyzing ChatGPT's Aptitude in an Introductory Computer Engineering Course (Szefer & Deshpande 2023)

Assigned Work	Student Solutions without ChatGPT			Student ChatGPT-assisted Solutions			OpenAI API-generated Solutions		
	Avr. Score	*Max. Points*	*Percentage*	*Avr. Score*	*Max. Points*	*Percentage*	*Avr. Score*	*Max. Points*	*Percentage*
Quiz, Week 2	2.7	3	89%	3.0	3	100%	1.8	3	60%
Quiz, Week 3	2.8	3	94%	2.9	3	97%	2.5	3	83%
Quiz, Week 4	2.6	3	86%	2.6	3	87%	1.8	3	60%
Quiz, Week 5	2.5	3	84%	3.0	3	100%	1.3	3	43%
Quiz, Week 6	2.7	3	89%	2.7	3	90%	2.5	3	83%
Quiz, Week 7	2.9	3	95%	3.0	3	100%	2.2	3	73%
Assignment 1	22.2	25	89%	6.92	25	27%	9.2	25	37%
Assignment 2	22.9	25	91%	1.67	25	8%	1.9	25	8%
Assignment 3	21.7	25	87%	3.00	25	13%	2.0	25	8%
Mid-Term Exam	—	50	—	—	50	—	—	50	—
Lab 1 Questions	5	5	100%	2.90	5	64%	2.2	5	44%
Lab 2 Questions	5	5	100%	2.70	5	60%	2.8	5	56%
Lab 3 Questions	5	5	100%	3.00	5	70%	2.8	5	56%
Lab 4 Questions	5	5	100%	2.88	5	60%	2.2	5	44%
Lab Practical Exam	—	30	—	—	30	—	—	30	—

TABLE III

APPROXIMATE TIME TO COMPLETE THE WORK.†

Assigned Work	Approx. Time to Generate Solutions		
	Student no ChatGPT (min.)	*Student with ChatGPT (min.)*	*Automated Code (min.)*
Quiz, Week 2	5	—	0.3
Quiz, Week 3	5	—	0.4
Quiz, Week 4	5	—	0.2
Quiz, Week 5	5	—	0.4
Quiz, Week 6	5	—	0.3
Quiz, Week 7	5	—	0.3
Assignment 1	380	—	3.0
Assignment 2	560	—	2.3
Assignment 3	800	—	1.7
Mid-Term Exam	75	—	1.0
Lab 1 Questions	68	—	0.3
Lab 2 Questions	72	—	0.5
Lab 3 Questions	79	—	0.4
Lab 4 Questions	73	—	0.4
Lab Practical Exam	86	—	0.4

TABLE IV

EFFICIENCY OF GENERATING THE SOLUTIONS, IN PERCENTAGE PER MINU
HIGHER VALUE IS BETTER.

Assigned Work	Solution Generation Efficiency		
	Student no ChatGPT (% / min.)	*Student with ChatGPT (% / min.)*	*Automated Code (% / min.)*
Quiz, Week 2	17.8	—	200
Quiz, Week 3	18.8	—	208
Quiz, Week 4	17.2	—	300
Quiz, Week 5	16.8	—	108
Quiz, Week 6	17.8	—	277
Quiz, Week 7	19.0	—	243
Assignment 1	0.23	—	12
Assignment 2	0.16	—	3.5
Assignment 3	0.11	—	4.7
Mid-Term Exam	—	—	—
Lab 1 Questions	1.47	—	146
Lab 2 Questions	1.39	—	112
Lab 3 Questions	1.27	—	140
Lab 4 Questions	1.37	—	110
Lab Practical Exam	—	—	—

1. Impact of AI on Education Quality

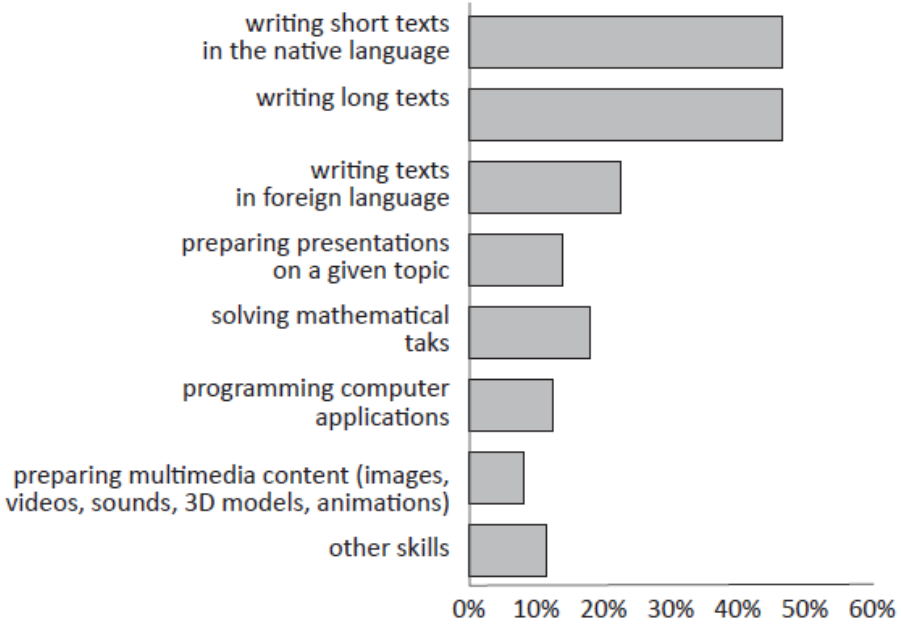
De Castro (2023) found that while AI like ChatGPT can significantly boost student engagement and provide personalized learning, it raises **serious concerns** about academic integrity and privacy. In some cases, universities reported **an alarming increase in plagiarism** and over-reliance on AI-generated content.

.

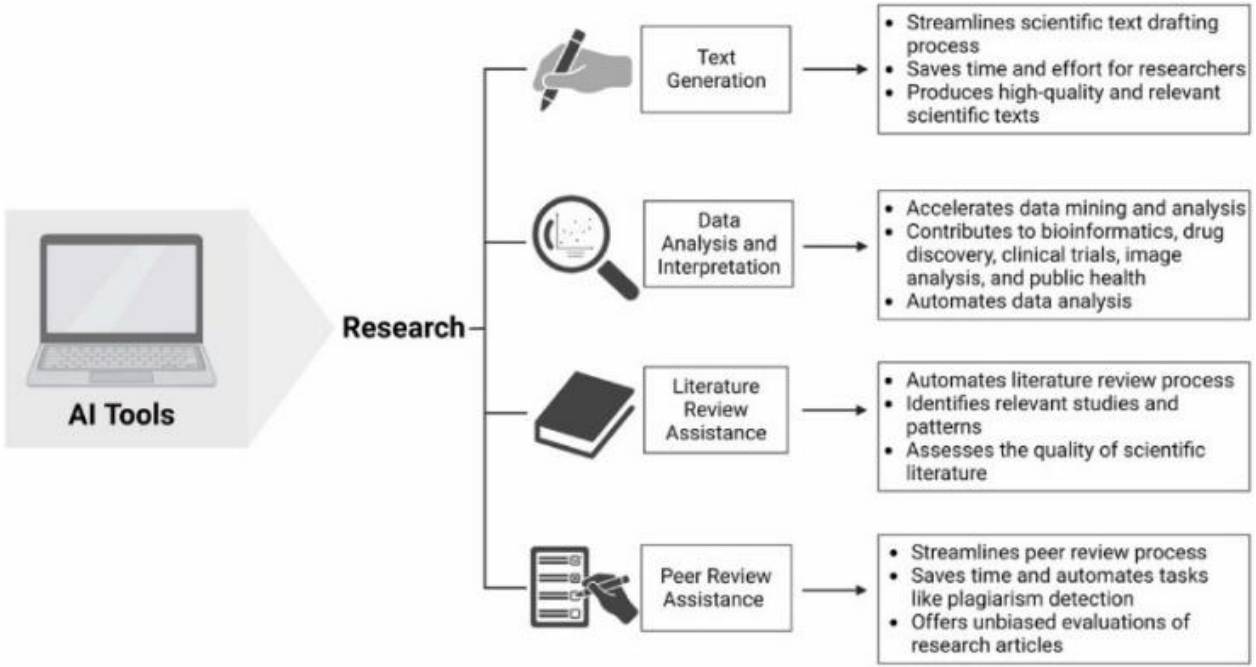
2. Decline in Critical Thinking

- **Frith (2023)** called ChatGPT a **"disruptive educational technology"** and pointed out that the rise of AI could cause **catastrophic reductions in critical thinking**. Rather than deeply engaging with the material, students often rely on AI to produce **surface-level answers**, leading to a worrying decline in analytical skills.

How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence? Outcomes of the ChatGPT teacher survey (Bower et al. 2024)



The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. (Alqahtani et al. 2023)



Literature Review

Creativity vs. AI Limitations

De **Vicente-Yagüe-Jara et al. (2023)** discovered that AI could enhance fluency and adaptability in writing but noted a **disturbing trend**: students using AI displayed **far less originality**. AI-generated work often lacked creative insights, leading some educators to fear that the next generation might rely heavily on AI and lose the ability to innovate.

The Rise of "CheatGPT"

Shaw et al. (2023) highlighted the rising use of AI for unethical educational practices, coining the term **"CheatGPT."** The authors detailed cases where students used ChatGPT to create entire assignments and bypass plagiarism detectors, leading some institutions to resort to **pen-and-paper exams** to restore academic integrity

Unequal Learning Environments

Cassidy (2023) described a dramatic divide between universities that allow AI use and those that don't, dubbing this phenomenon **"AI privilege."** Institutions that ban AI face an **uphill battle** with students finding ways to cheat, while those that embrace AI without proper guidelines risk graduating students with **shallow, AI-generated knowledge**.

Literature Review

ChatGPT as an Educational Tool:

Opportunities, Challenges, and Recommendations for Communication, Business Writing, and Composition Courses (AlAfnan et al. 2023)

han can be managed effectively, and a perception that overload in his sense creates a degree of stress for which his or her coping strategies are ineffective' (p. 34). Table I provides the similarity

Table I. Similarity index and grading outcome for the theoretical communication question

Test number	Turnitin similarity index (%)	Grade given	Additional notes
Test 1	50	93/100	This answer entails misconduct as it is heavily plagiarized.
Test 2	13	90/100	Comprehensive answer but conversational.
Test 3	0	95/100	Well written and well structured.
Test 4	20	89/100	A comprehensive answer, but conversational. The student may lose more grades on originality.
Test 5	19	89/100	This answer is an obvious paraphrase of test 4. Even though Turnitin did not detect the similarity, the similarity is obvious. The student may lose more grades on originality.

Table II. Similarity index and grading outcome for the case-based communication question

Test number	Turnitin similarity index (%)	Grade given	Additional notes
Test 6	3	83/100	A comprehensive analysis that synthesizes the situation but lacks summary and support (incorporating additional resources)
Test 7	0	81/100	A comprehensive analysis of the situation. However, the suggestions are almost the same as in test 6. ChatGPT paraphrased the answer of test 6.
Test 8	9	82/100	A comprehensive analysis of the case. The paraphrase is also obvious.
Test 9	0	81/100	A comprehensive analysis of the case. The paraphrase is also obvious.
Test 10	0	83/100	A comprehensive analysis of the case. The paraphrase is also obvious.

Literature Review

Writing, creativity, and artificial intelligence. ChatGPT in the university context (De-Vicente-Yagüe-Jara et al. 2023b)

Table 4. Creativity of AI systems and learners

Creativity indicators	Human Intelligence		Artificial Intelligence		U	Z	p	Effect Size
	M	SD	M	SD				
G2 Fluency	10,47	5,756	29,30	16,927	228,000	-6,499	<,001	.445
G2 Flexibility	7,56	3,637	12,35	5,334	790,000	-4,363	<,001	.299
G2 Narrative Originality	11,02	8,337	37,40	20,451	246,500	-6,422	<,001	.440
G3 Fluency	9,52	5,576	10,65	7,443	1798,000	-,504	,614	.034
G3 Flexibility	6,87	3,467	6,90	4,340	1913,000	-,065	,948	.004
G3 Narrative Originality	9,35	8,509	7,80	7,818	1644,500	-1,090	,276	.075

Moreover, no statistically significant differences were found in the indicators of game 3, indicating similar levels between the students (human intelligence) and the AI. However, it should be noted that students show a higher mean on the narrative originality indicator in game 3. Also, fluency and narrative originality in game 2 have moderate effect sizes, while the other indicators in game 2 and game 3 have low effect sizes.

Literature Review

(Pham et al., 2023) Digital Transformation in Engineering

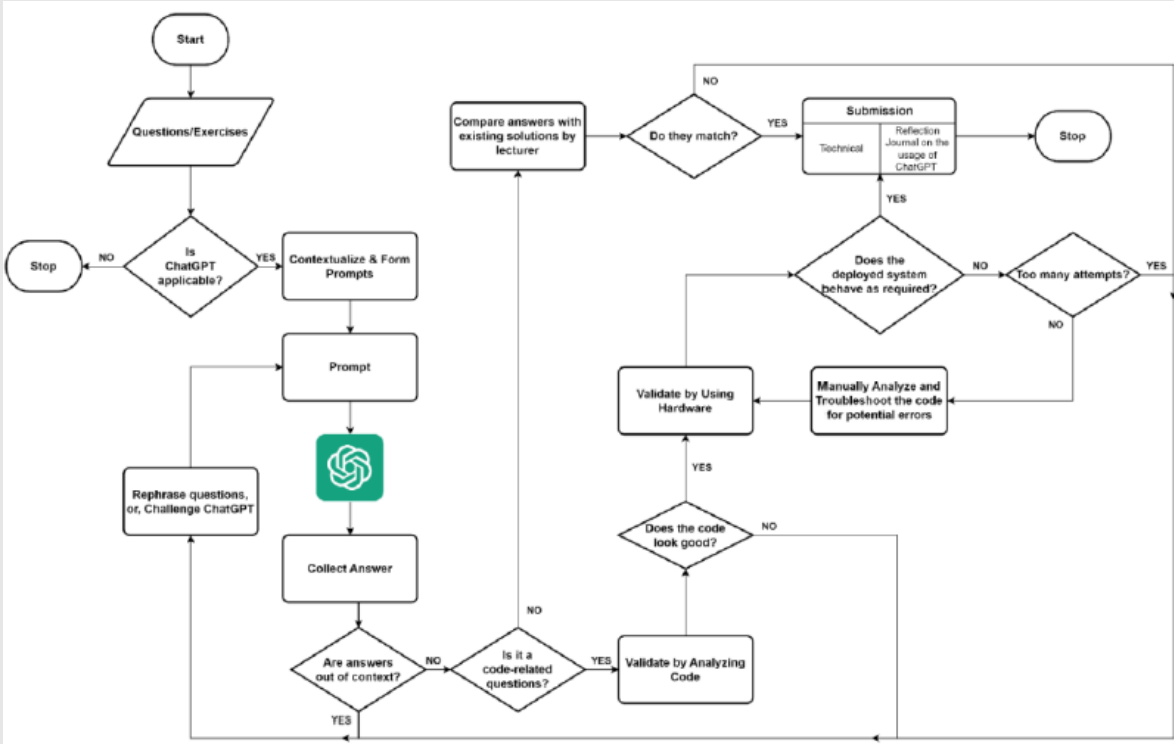


Figure 1. Proposed flow of using ChatGPT in assisting learning

Exam Questions

a) Provide an introduction for your selected case describing how it aligns with achieving the target of Sustainable Development Goal 7.b. Describe how you consider the resources available, socio, political, and geographical aspects. **4p**

b) Formulate a one sentence problem statement. Then create a CONOPS ("AS-IS" and "TO-BE" scenarios) diagram describing the problem domain. Describe the challenges in "AS-IS" situation

2

Universitetet
i Sørøst-Norge

and anticipated outcomes from "TO-BE" (Describe *What* the outcomes are, not *How* the problem is solved). **8p**

c) Refer to your CONOPS diagram, identify 4 key stakeholders of the problem and elaborate on their needs/concerns. (identify the originating stakeholder). **8p**

d) List the criteria (min. 8) for selecting/prioritizing a concept, and provide rationale considering Stakeholder needs. **8p**

e) Suggest 3 concepts that might meet the requirements, select a concept based on Pugh matrix approach by assessing the criteria from (d). Explain the key drivers in finalizing your concept. **10p**

f) Identify/draw the Dynamic Behavior happening for the selected concept. **8p**

g) List and describe the Enabling systems in realizing the concept chosen in (e). **4p**

h) Describe 4 functions of your concept/system and write corresponding system level requirements. **8p**

i) Suggest methods to verify each functional/system requirements during i) Design phase and ii) Prototype phase. Describe the verification methods in sufficient detail. **12p**

j) What trade-offs/risks have you considered in realizing the concept, provide rationale. **5p**

k) For your chosen solution, identify and explain three aspects that you would consider when applying Design for Cost (part of DfX model). **8p**

l) As part of SEMP, you are requested to identify at least 4 decision gate reviews across the system lifecycle. Describe the intention of review and which personnel roles should be involved. Provide supporting rationale. **7p**

m) Write your reflection on application of Systems Engineering on SDG 7.b (Max. 1 page) **10p**

2023- ITM4200 Exam questions

a) Provide an introduction for the selected case describing how it aligns with Sustainable Development Goal 9.

b) Formulate a one sentence problem statement. Then create a ConOps/context diagram describing the problem domain. Briefly describe the items you place on your diagrams.

c) Identify the stakeholders of the problem and their needs. Write four stakeholder requirements (identify the originating stakeholder).

d) Suggest 3 concepts that might meet the requirements, select the best concept based on Pugh matrix approach and explain the limitation of using this decision-making method.

e) Write 4 system level functional requirements for the selected concept and suggest methods to verify the requirements. Describe the verification methods in sufficient detail. Provide traceability between stakeholder and system requirements.

f) Prepare a risk matrix that summarizes the pros and cons of your proposed solution and

2

Universitetet
i Sørøst-Norge

provides suggested mitigations for each of these risks.

Q2. Total value 30 points

Identify an application of good engineering practices in any product that you are reasonably familiar with, covering as many phases in the **life cycle** as possible. You could use your own experiences researching your bachelor's thesis. Document the case in no more than **six pages**. Show what has been done and hypothesize how the systems approach was applied. Indicate what was not done, or what could have been done in a different way.

Include the Systems Engineering Management Plan (SEMP) for case/product/system under consideration. In SEMP provide an example of at least one verification method in each life cycle stage.

2022-ITM4200 Exam questions

Part 1 of the 2022 questions are the same as all 2023 assignments. All the papers compared answered the same questions in around 15 pages.

Case Study(pre-post) Results

Year	Paper	Grades	Average Sentence length	Unique words	Word variety	Plagiarism rate	AI Plagiarism	using ChatGPT	What was their use case?
2022	Paper1	C	18.7	23 %	578	2 %	2 %	No	N/A
2022	Paper2	A	22.7	18 %	750	4 %	1 %	No	N/A
2022	Paper3	A	20.6	20 %	766	7 %	2 %	No	N/A
2022	Paper4	C	24.3	18 %	605	5 %	1 %	No	N/A
2023	Paper1	C	18.7	14 %	1092	8 %	82 %	Yes	Better writing, concept and idea creation
2023	Paper2	A	14.5	16 %	944	1 %	11%	Yes	Asking specific questions/Chatting
2023	Paper3	A	14.7	15 %	761	4 %	82 %	Yes	Asking specific questions/Chatting, concept and idea creation
2023	Paper4	A	22.1	19 %	1118	2 %	1%	Yes/No	Asking specific questions/Chatting
2023	Paper5	C	12.6	19 %	785	3 %	16 %	Yes	Literature review, idea concept creation, Chatting
2023	Paper6	A	18.3	18 %	1089	7 %	40 %	No	Better writing
2024	ExperimentalPaper1	C (Gerrit and John)	16.5	20 %	1078	4 %	100 %	Yes (100 %)	All

Case Study(pre-post) Results

☑ **Have you used ChatGPT in your assignment?**

Number of submissions: **14**

Submissions	Count	% of submissions	Chart
Yes	10	71.4%	71.4%
No	5	35.7%	35.7%

☑ **Do you have a privacy concern about using Chatgpt?**

Number of submissions: **14**

Submissions	Count	% of submissions	Chart
Yes	5	35.7%	35.7%
No	9	64.3%	64.3%

Case Study(pre-post) Results

How do you use ChatGPT or other chatbots in general?

Number of submissions: **14**

Submissions	Count	% of submissions	Chart
Asking for better writing (spelling, grammar etc.)	6	42.9%	42.9%
Asking for idea or concept creation	3	21.4%	21.4%
Literature review	2	14.3%	14.3%
Asking spesific questions (chatting)	7	50%	50%
Decision making	1	7.1%	7.1%
N/A	2	14.3%	14.3%

Experimental Paper Analysis

Criteria	Assessor 1	Assessor 2
Overall Impression	At first glance, the exam looks very good.	The paper jumped to a solution without properly exploring the problem space and understanding stakeholders.
Steps and Methodology	It makes all the steps as prescribed.	When developing solutions using AI, it's important to balance technical detail with conceptual modeling.
Diagrams	Diagrams are very bad; they suffer from being beautiful instead of functional.	Diagrams had "some nice pictures" but the text could not be seen clearly.
Quantification	Submission seems to miss quantification nearly entirely; this is a significant minus. Some quantification appears on page 11 but isn't well-argued.	It did not provide enough concrete details, numbers, and criteria to justify the proposed solution.
Dynamic Behavior	Only covers the obvious. Relevant use cases are when there is too much or too little power or energy, short or longer duration.	Lack of originality in the images used in dynamic behavior.
Enabling Systems	Missing procurement (supply chain), transportation, and installation. Cleaning is especially relevant due to dust risk in Chad.	Any proposed solutions need to consider the local context, challenges, and resources.
Verification Phase	Verification phase stops after the prototype. Missing FAT and SAT.	There's no verification method provided as clear and usable.
Mechanistic Answers	Answers feel mechanistic (simulate and measure prototype). In real life, this would be insufficient for a 7.5 ECTS course.	Not specifically mentioned.
Trade-offs and Risks	Trade-offs and risks feel mechanistic.	Assessments should look for more critical analysis to minimize impacts and costs.
Decision Gates	Feels rather theoretical.	Not specifically mentioned.
Reflection	Reads as a sales flyer, not as a reflection. The exam barely touches on socio-economic and environmental aspects.	Not specifically mentioned.
References	It provides unexpected references. I checked Sheard and Mostahari, an existing paper but not the core. Overall, it seems superficial.	Not specifically mentioned.
Final Grade and Comments	Would grade a C and discuss with another assessor to calibrate.	Would grade a C, but it may be B.

Case Study(pre-post) Results

Criteria	Student Paper (2023)	ChatGPT Paper (2024)
Concept Creation	Detailed	Lacked depth
Stakeholder Analysis	Comprehensive	Broad and shallow
Verification Methods	Specific	Generic explanations

Table 3: Experimental Paper Analysis

The Rubric

candidate	scaled	max pts	ID
total		ind.	
Letter			
graded		0	
Q1	**30**		
a. Sustainable development Goal 9	5		
introduction		2	
Alignmment elements to G9		3	
b. Goal / Problem Statement	5		
Problem staement Sentance		1	
ConOps / Context of problem domain		2	
Descriptions		2	
c. Stakeholders	5		
Stakeholders and Needs		2	
Stakeholder Rqmnts (ID origin)		3	
d. Concepts	5		
3 concepts described v requirements		2	
Pugh Matrix		2	
Limitations		1	
e. verification methods	5		
System Requirements		2	
verification (detailed)		2	
Tracability to stakeholder		1	
e. Risk Matrix	5		
Pros		2	
Cons		2	
mitigations		1	

Thematic Analysis

Criteria	2022 Exam papers	2023 Exam papers	ChatGPT Exam Paper
a) Introduction-	Explains SDG but often lacks depth and comprehensive coverage of its components.	Comprehensive and detailed introduction on SDG, covering its components and relevance.	Clear and well-structured, strong alignment with energy goals. But lack of details and starts with solution.
a) Problem Statement	Generally clear but could be more concise and specific; often lacks depth in problem articulation.	Clear and well-articulated problem statement with specific details and context.	Problem defined too general and background information is too short.
b)Concept creation as-is, to-be	Present but often lacks detailed descriptions and comprehensive context.	Detailed with thorough context and comprehensive descriptions.	As-is/to-be is listed directly but it is comprehensive. CONOPs diagram is just a picture and is good enough.
c)Stakeholder analysis	Requirements are listed but not thoroughly explained or aligned with stakeholder needs.	Detailed and well-aligned with stakeholder needs, providing clear and specific requirements.	Broad identification but it is just listed and described in one sentence.
d)Concept creation and decision making	Described with Pugh Matrix but often lack full exploration of limitations and detailed comparisons.	Thoroughly described with detailed Pugh Matrix analysis, including comprehensive comparisons and limitations.	While deciding the solar panel concept, criteria pointed with similar grades in Pugh matrix. So, decision is not explained well.
e)Verification method	Described but need more detail and comprehensive traceability to stakeholder requirements.	Detailed and comprehensive with clear traceability to stakeholder requirements.	Logical and includes explanations. However, it is only verifying design and prototype phases.
f) Risk assessment	Risks are mentioned but not thoroughly analyzed, with limited discussion on mitigations.	Comprehensive analysis of pros, cons, and detailed mitigations with thorough discussion.	Risks are well listed and explained but all the risks are just financial, social and climate based. There is no technical risk identified.

Thematic Analysis

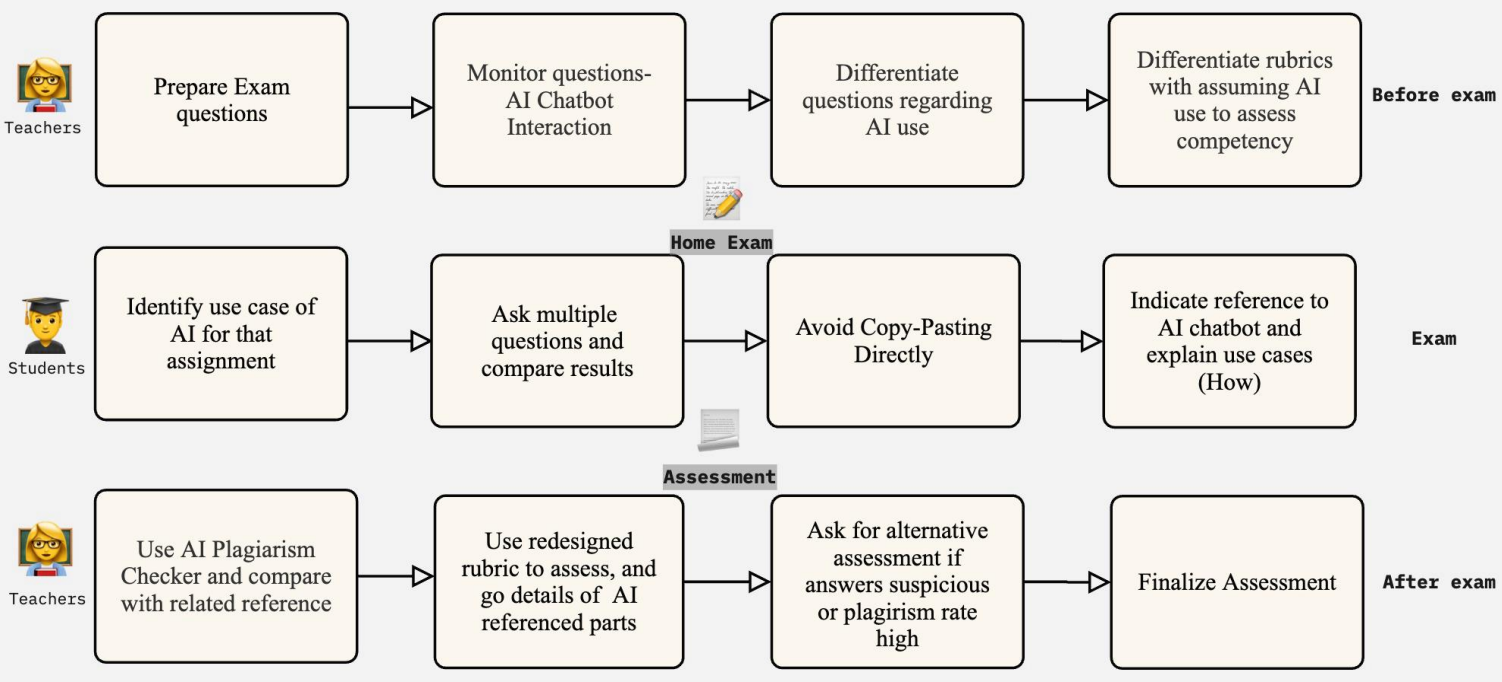
Theme	2022 Papers	2023 Papers	ChatGPT Paper (2024)
Problem Definition	Lacked depth	Detailed	Generalized
Stakeholder Analysis	Incomplete	Detailed	Minimal analysis
Concept Creation	Limited comparison	Detailed analysis	Generic decisions

Table 4: Thematic Analysis of Exam Papers

Summary of Data for Guideline Requirements

Requirement	Supporting Data
Transparency in AI Use	**84% of teachers** require clearer guidelines; **76% of teachers** want verification tools for AI use in exams.
Redesign of Exam Questions	**38% increase** in AI plagiarism detection post-ChatGPT; Originality and critical thinking most probably dropped in many 2023 exam papers.
Academic Integrity Tools	AI detection rates ranged from 1-2% to 38%, showing that tools are helpful but need to be supplemented with other methods, such as oral exams.
Differentiating AI-Assisted Work	The **shorter sentences, fewer unique words**, and responses in AI-generated work emphasize the need for rubrics focused on **creativity and depth**.
Guideline for AI Use	**83% of students** used ChatGPT for idea generation or writing help in 2023; proper guidelines are needed to promote responsible AI use.
Privacy and Security Awareness	While **67% of students** had no privacy concerns, educators need to increase awareness of data privacy risks associated with using AI tools.

AI-assisted Exam guideline



Validation

Validation from Experts:
Three experts from the Systems Engineering department provided insights into AI's impact on academic assessments and recommendations.

1. Changes in Assessment Practices

Interview 1: AI-generated answers tend to be generic. The expert introduced frequent feedback during the assessment to ensure deeper student engagement.

Interview 2: The expert emphasized adding dynamic, real-world problems to assessments to challenge AI-generated content.

Interview 3: Switched to case-based exams to ensure students' understanding of practical applications, making it harder to rely solely on AI.

Validation

2. Challenges in AI Declaration	**Interview 1 & 2**: Both experts highlighted the need for clear AI-use declarations, ensuring transparency in how students are using AI tools.	**Interview 3**: AI usage should be part of the academic honesty declaration to maintain integrity in submissions.
3. AI Plagiarism Detection	**Interview 1 & 2**: AI plagiarism tools are useful but unreliable. Additional methods, like oral exams, were recommended to confirm students' understanding.	**Interview 3**: AI detection should be part of the assessment, but the emphasis must be on assessing critical thinking and original contributions.

Validation

Interview 1: Rubrics do not require major changes but should include criteria that assess critical thinking and originality.

Interview 2 & 3: Rubrics should emphasize creativity and application-based work to evaluate students' understanding and ability to apply theoretical knowledge.

Discussion

Main Research Question:
What measures could the university staff and students implement to ensure a beneficial use of generative AI in systems engineering education?

- Guidelines are essential
- Transparency
- Assessment redesign
- Follow-up verification methods

Discussion

Sub-Question 1. According to educators, what are the critical criteria for assessing final papers in an introductory course called ‘systems engineering’?

- **Teacher Criteria**: Teachers focus on alignment with Sustainable Development Goals (SDGs), problem definition, stakeholder needs, and decision-making frameworks (e.g., Pugh matrix).
- **AI-specific Criteria**: Teachers also emphasize AI-specific aspects like “clarity and coherence,” “depth and accuracy,” “engagement and interactions,” “critical thinking,” and “originality and creativity.”

Discussion

Sub-Question 2. What challenges and solutions has the literature proposed for enhancing students' learning in fields related to systems engineering?

- Challenge: Over-reliance on AI

Solution: Adapt teaching and assessment methods

- Challenge: Privacy and data security

Solution: Guidelines and education on AI use

Discussion

Sub- Question 3. For what purposes are students and teachers in Systems Engineering currently using generative language models?

- Students' use of AI: Most students (83%) use ChatGPT for writing assistance, idea generation, and problem clarification rather than solving entire assignments.
- Primarily for writing improvement and quick Q&A support (50%), but also for literature review, concept creation, and decision-making.
- 14% of the students use it for literature review, 21% for concept creation, and 7% for decision-making during homework

Risks Challenges

Reliability Challenges

- AI Rapid Development
- Secondary Data Dependence

Scope and Sample Limitations

Risk Mitigations:

- **Expanded Scope**: To improve reliability, the author expanded the scope of analysis by comparing findings with existing literature and using different student surveys.
- **Experimental Paper**: Including an experimental paper helped validate results, providing a baseline for comparing pre- and post-AI impacts on student performance.

Conclusion

Summary of Key Findings

Enhanced Student Responses: AI contributes to more sophisticated, well-structured exam answers but may reduce critical thinking.

AI Usage Guidelines: Clear policies are needed for responsible AI use in academic settings; students should disclose and contextualize their AI use.

Adapted Assessment: Teachers should adopt rubrics to evaluate AI-related competencies, emphasizing creativity, depth, and engagement.

Recommendations

Application-based Exams

Privacy Awareness

Future Research Directions

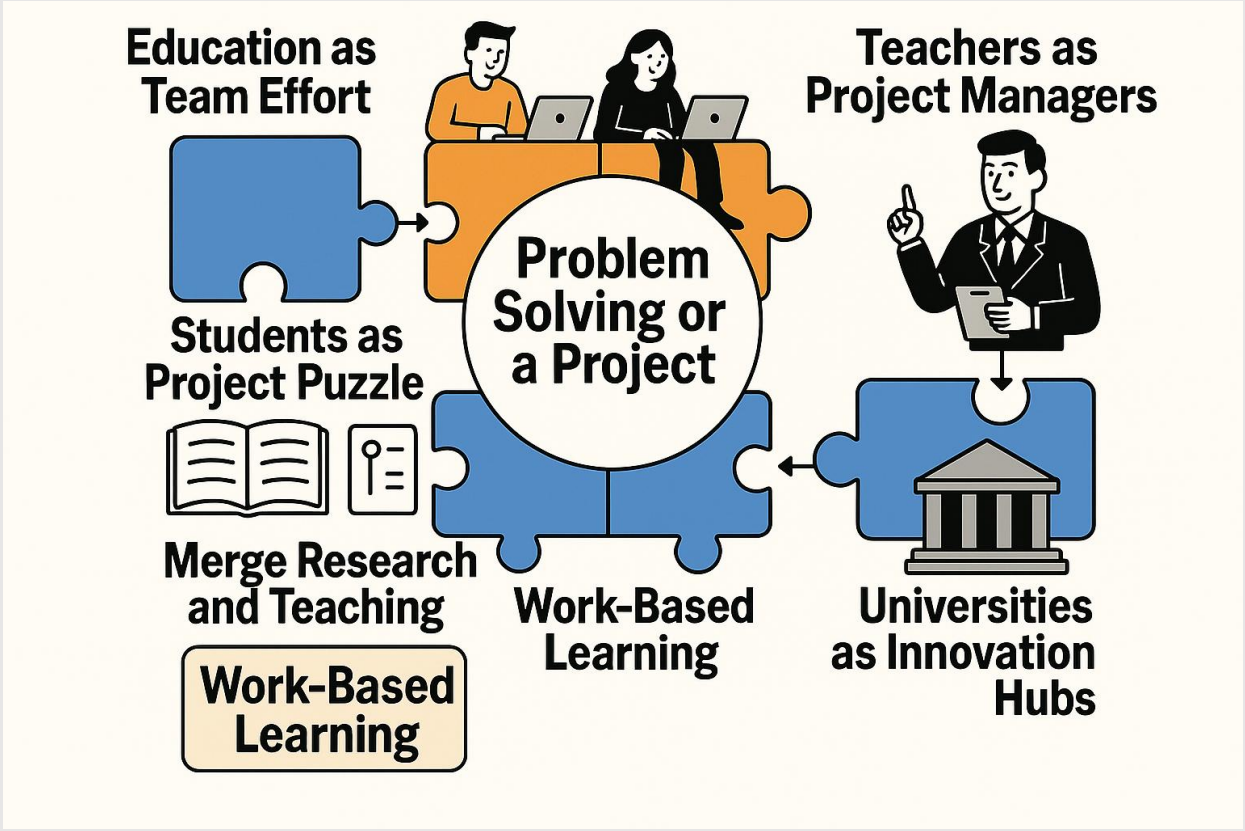
Broader Sample: Broaden research with more diverse samples to validate findings.

Competency Tracking: Explore competency changes in students before and after AI integration.

Conclusion

1. The increase in AI-generated content led to higher detection rates of plagiarism, showing that students often use ChatGPT without fully understanding ethical boundaries.
2. Findings indicate a decline in originality in post-ChatGPT exam papers, emphasizing the need for new strategies to preserve core competencies in systems engineering education.
3. The absence of clear guidelines on AI use has led to inconsistent practices among students and educators.
4. Traditional exam designs may need to be revised to evaluate AI-enhanced work. Exams emphasizing real-world problem-solving, oral evaluations, and practical applications will be required.
5. When students and ChatGPT are used together, the results are more detailed, and students achieve better grades than when ChatGPT alone is used.

How Can BE?


Education as Team Effort
Teachers as Project Managers
Problem Solving or a Project
Students as Project Puzzle
Merge Research and Teaching
Work-Based Learning
Universities as Innovation Hubs
Work-Based Learning

New Rubric

Category	Max Points	Score
a. Alignment with Objectives	**5**	
Introduction and context setting	2	
Alignment with objectives/topics	3	
b. Clarity and Coherence	**5**	
Clarity of explanations	2	
Logical flow and structure	3	
c. Depth and Accuracy	**5**	
Accuracy of facts and data	3	
Depth of content	2	
d. Engagement and Interaction	**5**	
Use of engaging language	2	
Interaction with questions	3	
e. Critical Thinking and Analysis	**5**	
Evidence of critical thinking	3	
Inclusion of perspectives	2	
f. Originality and Creativity	**5**	
Originality in approach	2	
Creativity in problem-solving	3	
Total	**30**	

35th Annual INCOSE international symposium

hybrid event

Ottawa, Canada
26 - 31 July 2025

Emin Simsek

MSc. Systems Engineering from USN

Gerrit Muller

Systems Engineering Prof. and Excellent educator at USN

Kristin Falk

Subsea and Systems Engineering Prof. at USN