

Working Group 4

Transformed by Transformers: Navigating the AI Coding Revolution for CS Education

James Prather, Paul Denny,
Juho Leinonen, Brett Becker, Ibrahim
Albluwi, Michelle Craig, Hieke Keuning,
Natalie Kiesler, Tobias Kohn,
Andrew Luxton-Reilly, Stephen MacNeil,
Andrew Petersen, Raymond Pettit,
Brent N. Reeves, Jaromir Savelka.

<https://iticse23-generative-ai.github.io/>



Literature review

We aim to establish the *status quaestionis* of this new research field and provide recommendations based on the current scholarly discourse.

- Reference papers
 - 10 “seed” papers, keyword search
- Snowballing phase
 - 513 papers reviewed
- Total: 48 papers

Venue	Count
ACM	15
ArXiv	13
IEEE	3
Other Publishers	4
Grey Literature	3

LLM	Count	Language	Count
Codex	13	Python	21
Copilot	10	Java	4
GPT-3/3.5	8	C/C++	4
Other	2	Javascript	2
GPT-4	1		

Author	Title	Venue	Year
Ahmed et al.	SYNSHINE: improved fixing of Syntax Errors	IEEE Trans. Softw. Eng.	2021
Alves and Cipriano	The centaur programmer - How Kasparov's Advanced Chess spans over to the software development of the future	arXiv	2023
Austin et al.	Program Synthesis with Large Language Models	arXiv	2021
Barke et al.	Grounded Copilot: How Programmers Interact with Code-Generating Models	OOPSLA	2023
Becker et al.	Programming Is Hard - Or at Least It Used to Be: Educational Opportunities and Challenges of AI Code Generation	SIGSE	2023
Belletini et al.	Davinci Goes to Bebras: A Study on the Problem Solving Ability of GPT-3	CSEEDU	2023
Brennan and Lesage	Exploring the Implications of OpenAI Codex on Education for Industry 4.0	SEHRGMA	2022
Brusilovsky et al.	The Future of Computing Education Materials	(in draft)	2023
Bull and Kharrufa	Generative AI Assistants in Software Development Education	arXiv	2023
Dakhl et al.	GitHub Copilot AI Pair Programmer: Asset or Liability?	J. Syst. Softw.	2022
Denny et al.	Robosourcing Educational Resources - Leveraging Large Language Models for Learnersourcing	arXiv	2022
Denny et al.	Conversing with Copilot: Exploring Prompt Engineering for Solving CS1 Problems Using Natural Language	SIGSE	2023
Ernst and Ravota	AI-Driven Development Is Here: Should You Worry?	IEEE Software	2022
Finnie-Ansley et al.	The Robots Are Coming: Exploring the Implications of OpenAI Codex on Introductory Programming	ACE	2022
Finnie-Ansley et al.	My AI Wants to Know if This Will Be on the Exam	ACE	2023
Idaho et al.	Wisdom(ist): Human or AI?		2023
Jalil et al.	ChatGPT and Software Testing Education: Promises & Perils	IEEE ICSTW	2023
Kazemitabaar et al.	Studying the effect of AI Code Generators on Supporting Novice Learners in Introductory Programming.	CHI	2023
Leinonen et al.	Comparing Code Explanations Created by Students and Large Language Models	arXiv	2023
Leinonen et al.	Using Large Language Models to Enhance Programming Error Messages	SIGSE	2023
Li et al.	Competition-level code generation with AlphaCode	Science	2022
MacNeil et al.	Experiences from Using Code Explanations Generated by Large Language Models in a Web Software Development E-Book	SIGSE	2023
Gherciu	Net Impact of Large Language Models Trained on Code	Student conf.	2022
Phung et al.	Generating High-Precision Feedback for Programming Syntax Errors using Large Language Models	arXiv	2023
Piccolo et al.	Many Bioinformatics programming tasks can be automated with ChatGPT	arXiv	2023
Poldrack et al.	AI-assisted coding: Experiments with GPT-4	arXiv	2023
Prather et al.	"It's Weird That It Knows What I Want": Usability and Interactions with Copilot for Novice Programmers	TOCHI	2023
Puryear and Sprint	GitHub copilot in the classroom: learning to code with AI assistance	J. Comput. Sci. Coll.	2022
Raman and Kumar	Programming Pedagogy and Assessment in the Era of AI/ML: A Position Paper	COMPUTE	2022
Ross et al.	A Case Study in Engineering a Conversational Programming Assistant's Persona	ACM III	2023
Sandoval et al.	Lost at C: A User Study on the Security Implications of Large Language Model Code Assistants	arXiv	2022
Sarsa et al.	Automatic Generation of Programming Exercises and Code Explanations Using Large Language Models	ICER	2022
Saveika et al.	Can Generative Pre-trained Transformers (GPT) Pass Assessments in Higher Education Programming Courses?	arXiv	2023
Saveika et al.	Large Language Models (GPT) Struggle to Answer Multiple-Choice Questions about Code	arXiv	2023
Vaithilingam et al.	Expectations vs. Experience: Evaluating the Usability of Code Generation Tools Powered by Large Language Models	CHI EA	2022
Wermelinger	Using GitHub Copilot to Solve Simple Programming Problems	SIGSE	2023
Yan et al.	Practical and Ethical Challenges of Large Language Models in Education: A Systematic Literature Review	arXiv	2023
Zhang et al.	Repairing Bugs in Python Assignments Using Large Language Models	arXiv	2022

Student & instructor perspectives: we need your help!



Generative AI Working Group at iTiCSE 2023

We are exploring how Generative AI tools like ChatGPT and Github Copilot may impact computing education practice, curricula, and policy. One important aspect of this work is to understand the perspectives of students & instructors.

Here are the three ways that you can help:

1. Fill out this instructor survey (10-15 minutes)

<https://forms.gle/GmYbfoqmDMptW3N8>

2. Share this recruitment information with your students

We are a group of international researchers that want to learn more about your experiences and perspectives about using Generative AI tools like ChatGPT and Github Copilot. We are interested in hearing from a diverse range of student perspectives. If you are a university student who has taken a programming course and have 10-15 minutes and want to have a role in shaping the future of education, please share your valuable perspective using the following form.

<https://forms.gle/hhZuzug1pTKm48bx6>

If you have any questions about this study, please contact Stephen MacNeil, at stephen.macneil@temple.edu. If you have questions about your rights as a participant, please contact the Temple University's Office of Research Ethics (irb@temple.edu or (215) 707-3390)

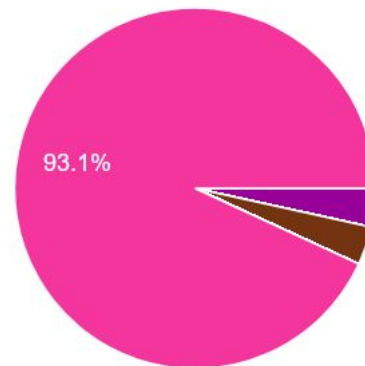
3. Share this request with your colleagues

<https://bit.ly/iticse-ai-working-group>

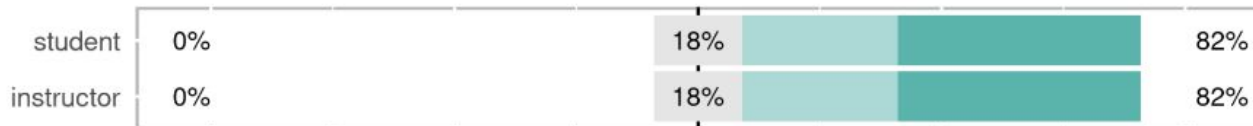
Early results (aka pilot survey)

Country

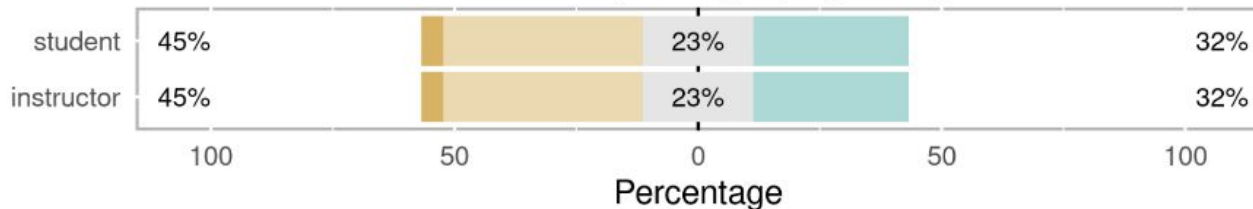
29 responses



Students must be taught how to use GenAI tools well for their future careers



Using GenAI tools frequently to generate code is harmful for my learning of programming



The Robots Are Coming: Exploring the Implications of OpenAI Codex on Introductory Programming

James Finnie-Ansley
The University of Auckland
Auckland, New Zealand
james.finnie-ansley@auckland.ac.nz

Paul Denny
The University of Auckland
Auckland, New Zealand
paul@cs.auckland.ac.nz

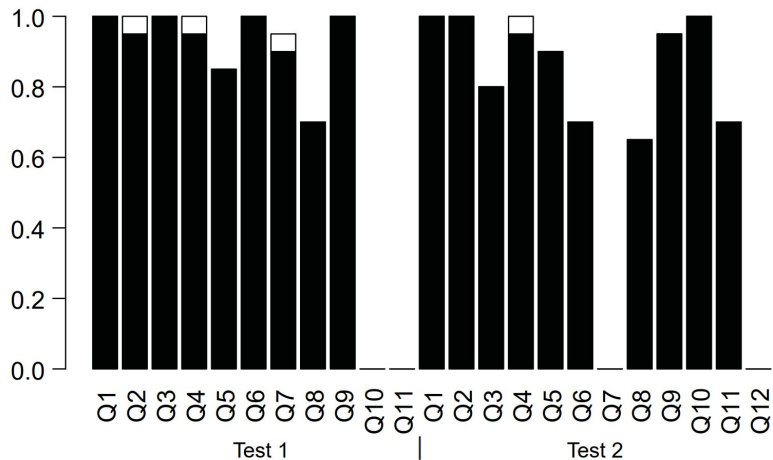
Brett A. Becker
University College Dublin
Dublin, Ireland
brett.becker@ucd.ie

Andrew Luxton-Reilly
The University of Auckland
Auckland, New Zealand
a.luxton-reilly@auckland.ac.nz

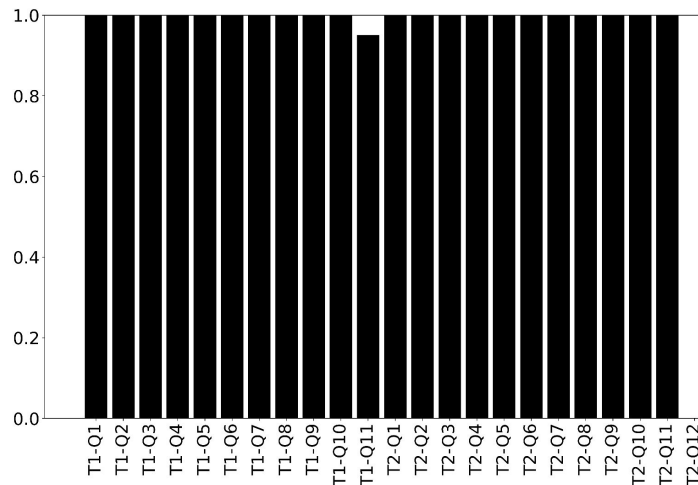
James Prather
Abilene Christian University
Abilene, Texas, USA
james.prather@acu.edu

The Robots are Here!

Then (2021, Codex)



Now (2023, GPT-4)



Novel analyses

Datasets: FalconCode, Automated Programming Progress Standard (APPS), HumanEval, EvalPlus, etc.

Tagging problems: **many** introductory/CS1 problems, very **few** CS2/OO/data structures

Running GPT-4:

FalconCode: **45.4%** (skill, lab, project) [74.1% on complete instructions]

APPS [sample]: **49.6%** (intro 72%, interview 53%, competition 29%)

bit.ly/WG4



Curriculum and assessment

LLMs are potentially changing:

- What we teach (**learning outcomes**)
- How we teach it (**course activities**)
- How we prepare to teach (**creating educational resources**)
- How we **assess** student learning

We are interviewing **educators** who have made or planned **concrete changes** to their teaching

Sign up so we can interview you!

Ethics

Mapped relationship between Academic Integrity Policies at various universities about generative AI and the general principles espoused in the ACM Code of Ethics

Looked at ethical values that are encoded in the literature on LLMs

Authorship and breaches of academic integrity in the context of generative AI:

- Users (students) are authors of content produced by generative AI.
- Incumbent on teaching staff to educate students about generative AI, and explicitly discuss restrictions of generative AI for assessed work.
- Academic misconduct arises due to use of unauthorised resources.

Ready resource to educate students on ethics of using generative AI (check site)

Working Group 4

Transformed by Transformers: Navigating the AI Coding Revolution for CS Education

James Prather, Paul Denny,
Juho Leinonen, Brett Becker, Ibrahim
Alblawi, Michelle Craig, Hieke Keuning,
Natalie Kiesler, Tobias Kohn,
Andrew Luxton-Reilly, Stephen MacNeil,
Andrew Petersen, Raymond Pettit,
Brent N. Reeves, Jaromir Savelka.

<https://iticse23-generative-ai.github.io/>

