Workshop AIforRSE

Exploring how AI tools can be integrated in the research software lifecycle

Date: 4 June 2024

Organized by: Research Data and Software Team (RDS)
TU Delft University | Library

TU Delft

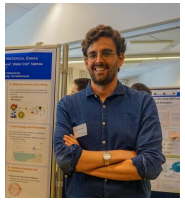
Image artificially generated using openart.ai with prompt "illustrate a workshop on AI tools for research software in Delft University"

TU Delft Library & Research Software Engineering (RSE)

The TU Delft Library is helping shape the university's strategy towards AI and RSE through:

- **Helping** define AI essentials & good practices
- **Teaching** core research software engineering skills
- **Working** to develop intermediate / advanced RSE training materials

Your Presenters:



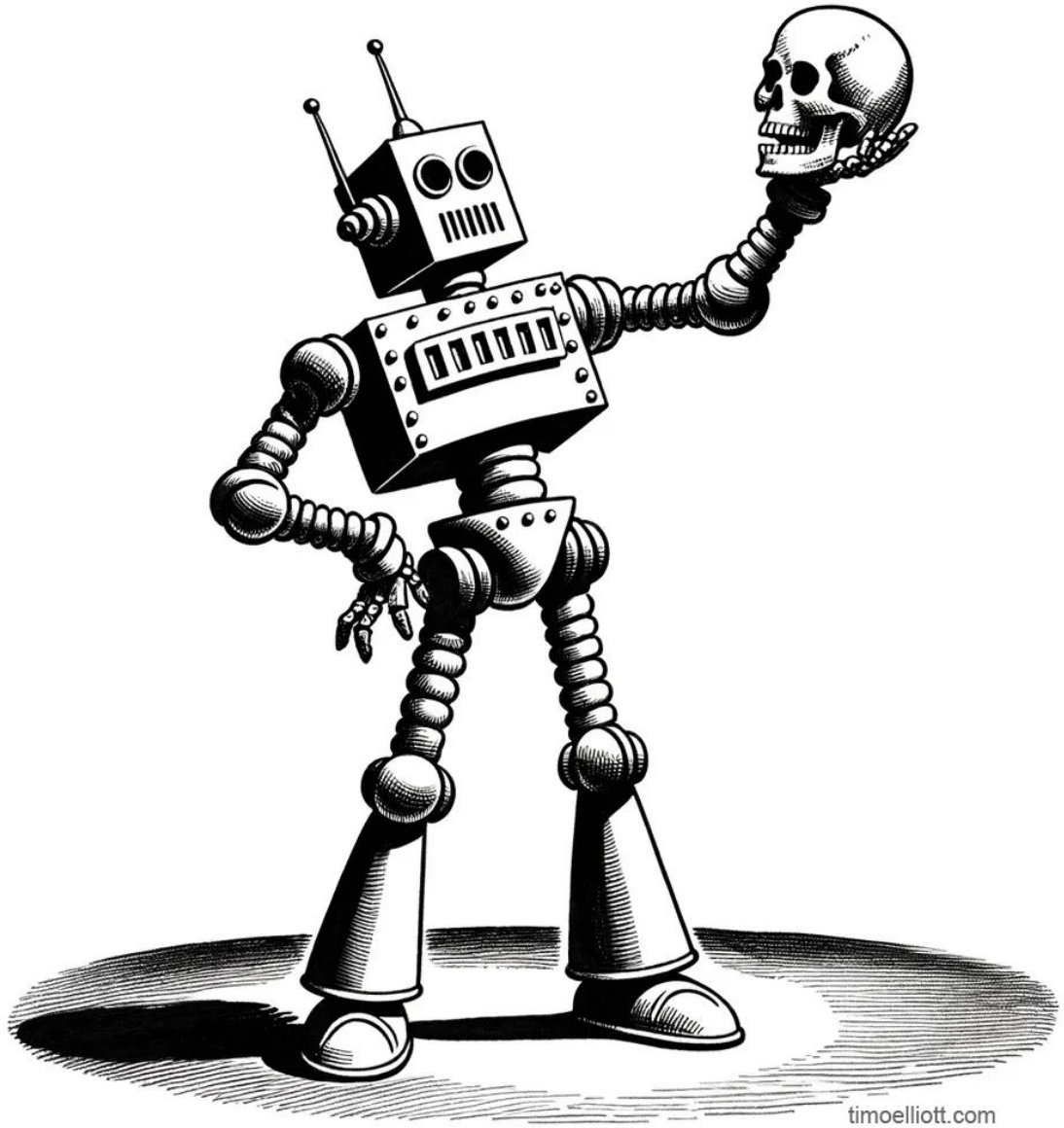
Carlos Utrilla Guerrero
RDS Trainer
https://carlosug.github.io/

Halford Dace
RDS Trainer

Fardad Maghsoudi Moud
Data Manager

TU Delft

To use or *not* to use AI for *better* research software, that is the question



"To be, or not to be? That is a syntax error..."

https://i0.wp.com/timoelliott.com/blog/wp-content/uploads/2023/11/to-be-or-not-to-be-that-is-a-syntax-error.jpeg?ssl=1

TU Delft

Agenda

1 | What is AI and GenAI?

(Halford)

2 | What is RSE and the challenges?

(Fardad)

3 | Why use GenAI for RSE?

(Carlos)

4 | Demo Github Copilot

5 | Conclusion

TU Delft

Image artificially generated using openart.ai with prompt "illustrate a workshop on AI tools for research software in Delft University"

Collaborative notes:
https://edu.nl/qx9qj

edu.nl/qx9qj

menti.com | use code **5384 8806**
https://www.menti.com/alxy1r19nrye

Part 1
What is AI and GenAI

Collaborative notes:
https://edu.nl/qx9qj

edu.nl/qx9qj

https://edu.nl/qx9qj

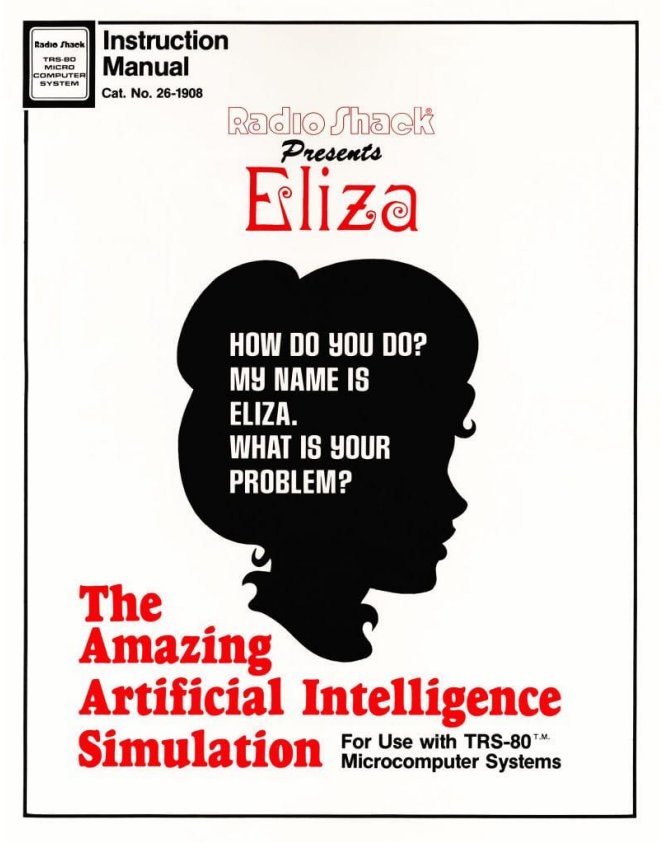
edu.nl/qx9qj

The Term “AI” can refer to:

- **Machine Learning**
  - Statistical models used for analysing / classifying / predicting based on complex data sets
  - (Mature technology: Routinely and extensively used across academic science and industry)
- **Artificial General Intelligence**
  - The ability of a machine to reason abstractly and broadly in a way comparable to a human
  - (Does not (yet?) exist. Primarily explored through science fiction. Think HAL 9000)
- **Generative AI**
  - Statistical **generative** models based on probability distribution, that can generate complex media (text, images, audio, video) in response to a prompt
  - (New, emerging technology: Lots of media attention, excitement and caution. Valid use cases emerging)

TU Delft

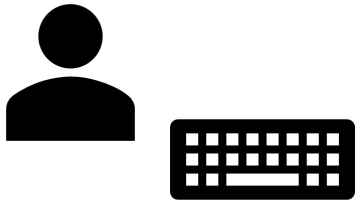
1966 computer program that has natural language interaction with humans



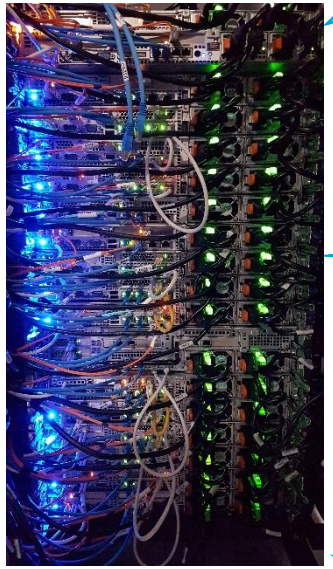
endpoint: https://web.njit.edu/~ronkowit/eliza.html

TU Delft

Generative AI



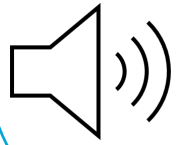
“DUCK”



Massimo Boturi via Unsplash

A duck is one of several species of waterfowl in the family Apatae

Timothy Dykes via Unsplash

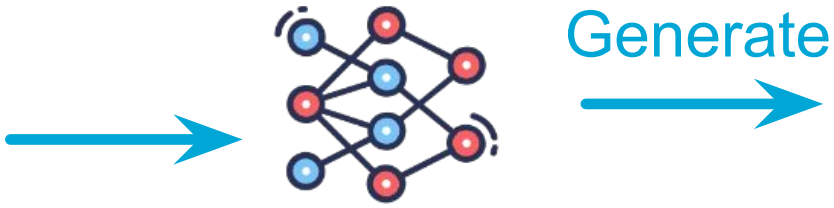


Quack

TU Delft

Text2Image Diffusion models

“Workshop in Delft”



Generative model of illustrations

Language-guided artwork creation

https://chainbreakers.kath.io @RiversHaveWings

TU Delft

Generative AI Models **ARE**:

- Trained on very large datasets
- Capable of generating various media in response to prompts:
  - Text (Abstracts, documentation, code, grant proposals, essays, exam questions, advertising copy)
  - Images (Visual abstracts, figures, graphs, “original” art in various styles)
  - Audio and Video (Jingles, symphonies, animations, deepfake propaganda)
- Highly Proprietary
- Powerful tools – but how will we use them?

Generative AI Models **ARE NOT**:

- Conscious
- Factual Repositories
- Aware of the concepts of truth or falsehood
- Unbiased
- “Aware” of anything after their training date

TU Delft

Some Constraints Of These Interesting Tool

- Accuracy: AIs are statistical models of text/pixels. They don't “know” things. So output can be of variable “truth” quality
- Authorship: AIs effectively recombine other people’s prior work to produce their output. Who is the author? What are the ethical and legal issues around their original training sets?
- Bias: Outputs reflect the biases of inputs. Expect AI output to be step backwards in terms of diversity and inclusion
- Privacy Hygiene: Your data/code is uploaded to servers (often outside Europe) to generate responses. This might not meet regulatory or contract requirements for your project, so check!

TU Delft

Examples

Bard

ChatGPT

DALL·E

Potential dangers: Integrate into the classroom

How Much Research Is Being Written by Large Language Models?

Source: https://stanford.io/3wJxo7A

By the way, ChatGPT misspelled your name.

Source: https://upfront.scholastic.com/content/dam/classroom-magazines/upfront/issues/2023-24/051324/p24-cr-cartoons/PO1-UPF051324-CR.jpg

TU Delft

Potential dangers: Text2Image - Gender bias?

User input:

“create a image to illustrate a workshop on exploring the AI tools for research software in a Spring Symposium on AI education at Delft University”

Ideogram 1.0 model

https://ideogram.ai/

TU Delft

Part 2
What is RSE and challenges

Collaborative notes:
https://edu.nl/qx9qj

edu.nl/qx9qj

Research needs Software

92% use research software
67% say would not be able to do research without it
56% develop software as part of their research

Source: Software Sustainability Institute - https://bit.ly/2zZPhSa

https://www.software.ac.uk/sites/default/files/2024-01/BetterSoftwareBetterResearch.png

TU Delft

What is Research Software?

"Any software created during the research process or for a research purpose" [Gruenpeter, M. et al. Defining Research Software, 2022]

Research Software

created during the research process for a research purpose

vs.

Software in research

used for research but not created during or with research intent

Python script developed to analyse and visualise data ✔

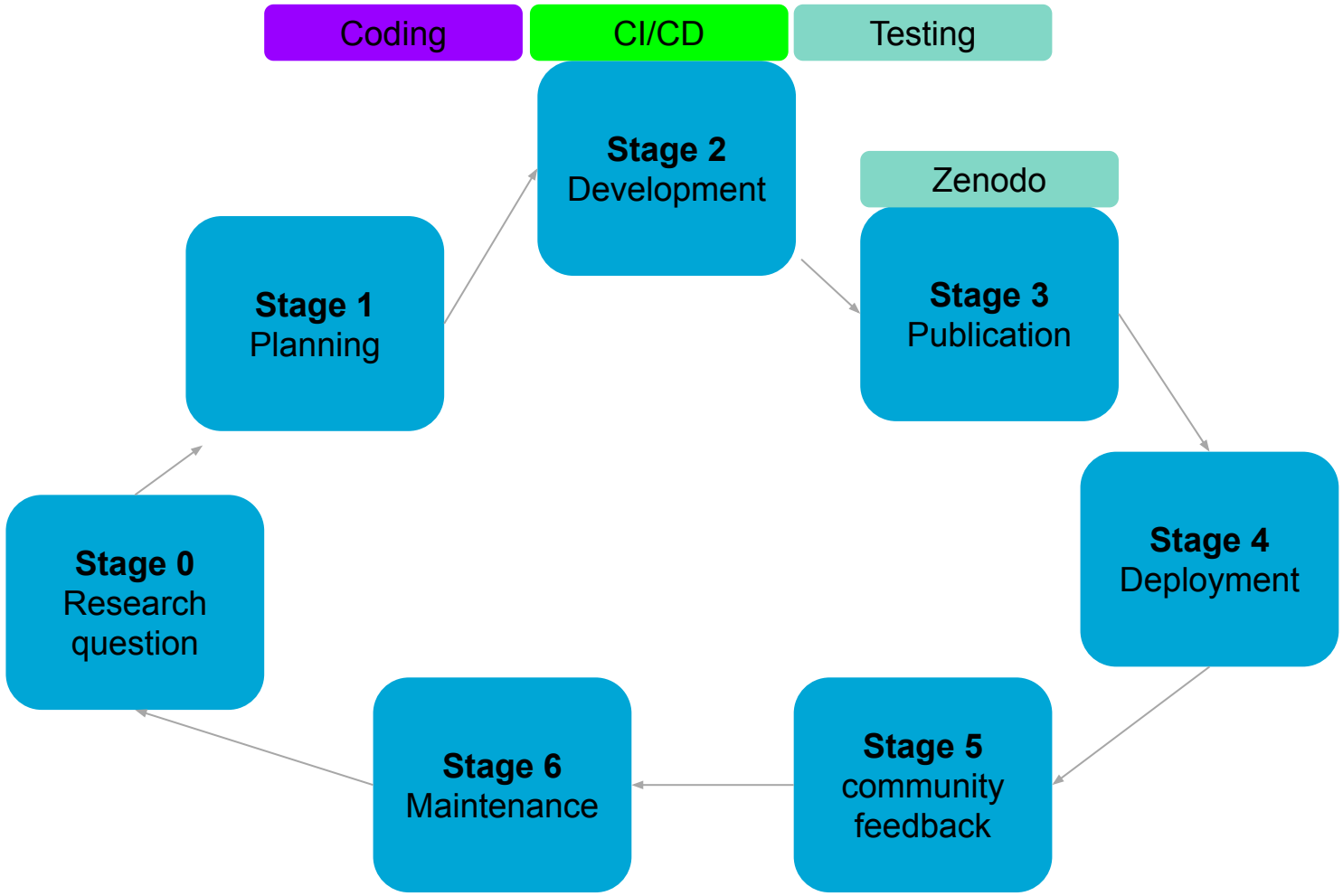
- **Excel** used to analyse and visualise data ✖

TU Delft

Warning: Differentiation may vary between disciplines

Research software lifecycle and its activities

How GenAI aid to automate the RSE lifecycle?



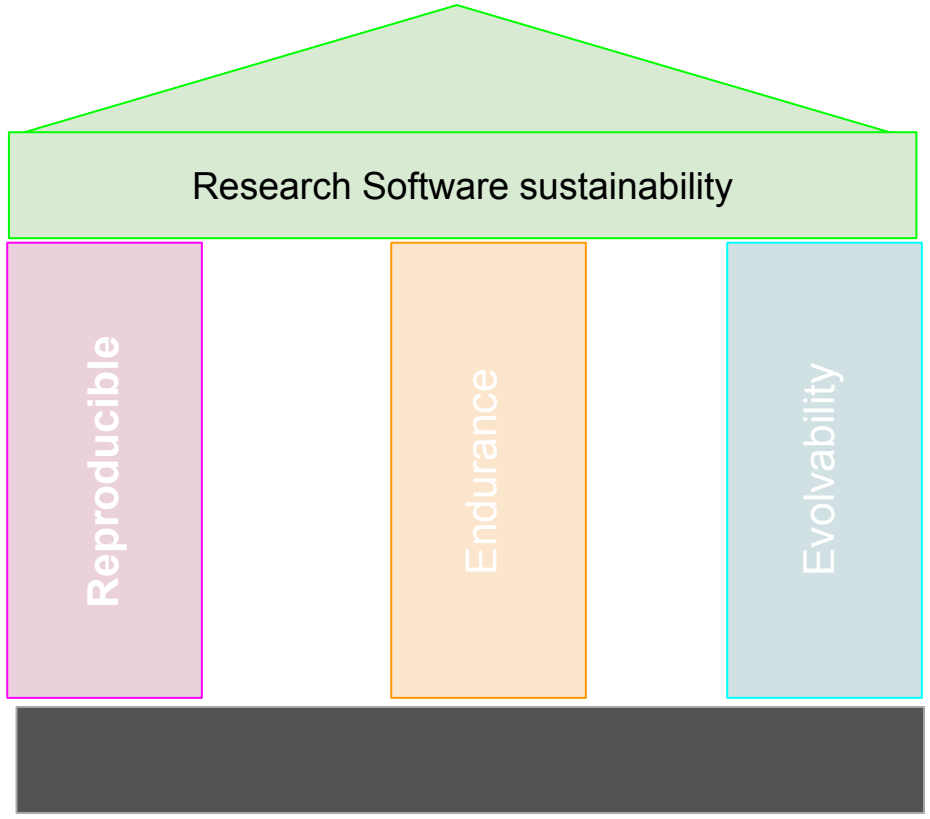
TU Delft

A research software vision: Make software first Class output

- All research software that can be is open
- All research software is reproducible
- All research software is high-quality and robust
- All research software is findable, accessible, and usable

TU Delft

RSE VISION: Making research software more sustainable, reproducible and reusable



> *"the software will continue to be available in the future, on new platforms, meeting new needs"*

https://danielskatzblog.wordpress.com/2016/09/13/defining-software-sustainability/

ADORE .SOFTWARE

<ReSA> Research Software Alliance

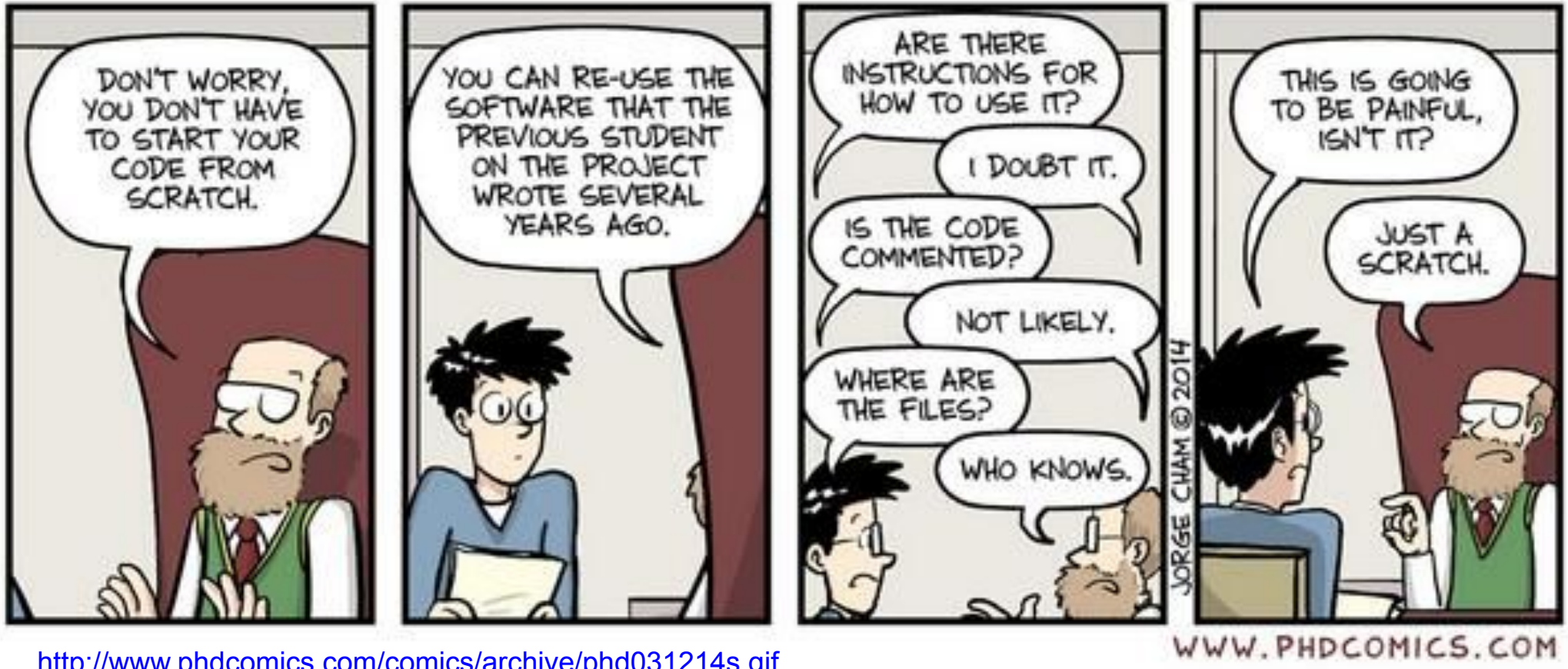
RSE

SOFTWARE SUSTAINABILITY INSTITUTE

Findable Accessible Interoperable Reusable

TU Delft

Can you relate to this situation?



http://www.phdcomics.com/comics/archive/phd031214s.gif

TU Delft

Making Good Software Is Hard

CHALLENGES

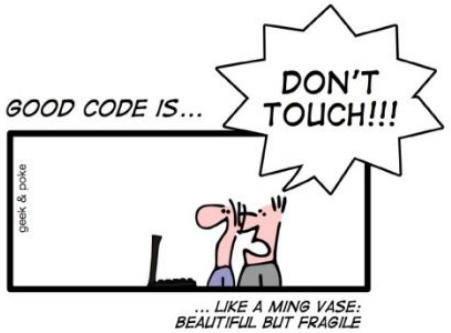
- Code Quality (Does it do what we intend it to?)
- Usability ("This software is not for stupid people" – Anonymous bioinformatics developer)
- Continuity (Documentation and maintenance after the first PhD has left!)
- Design Stamina Hypothesis (https://martinfowler.com/bliki/DesignStaminaHypothesis.html)

Source: ProgrammerHumor

TU Delft

Discussion: How do you teach students to develop good RSE practices?

Reproducible Environments

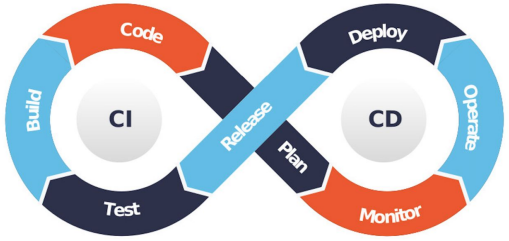


Code quality

Code testing

Code review

Documentation



Continuous Integration

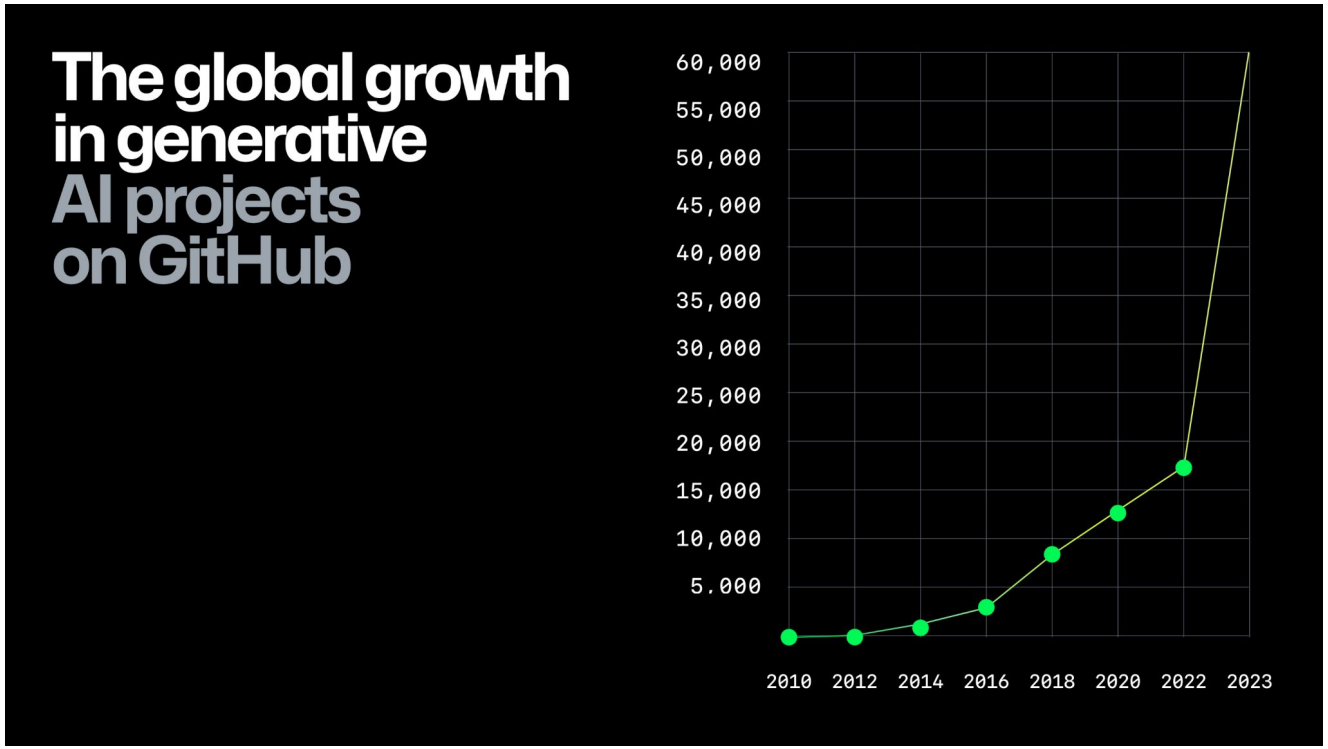
TU Delft

Part 3
AI in RSE

Collaborative notes:
https://edu.nl/qx9qj

edu.nl/qx9qj

Significant effort to install, interpret and make sense of AI projects for new purposes



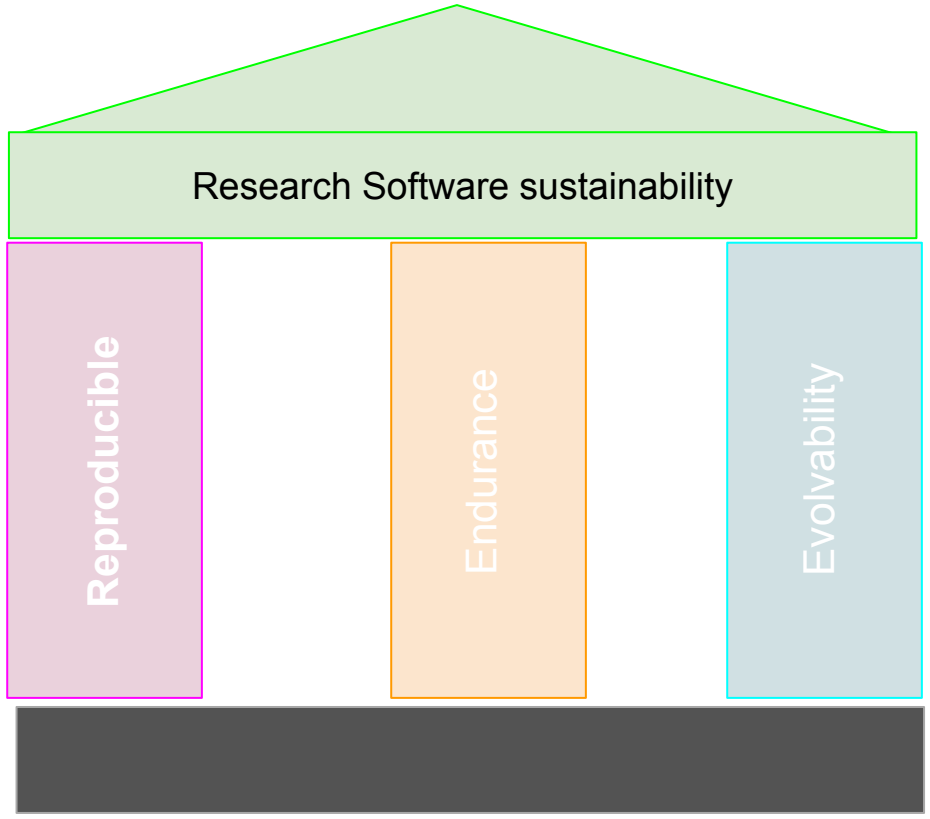
https://github.blog/2023-11-08-the-state-of-open-source-and-ai/#the-explosive-growth-of-generative-ai-in-2023

Challenges:

- Structured representation
- Search is time consuming
- Compare tools is time-sink
- Diverse practices in RSE

TU Delft

RSE VISION: Making research software more sustainable, reproducible and reusable



> *"the software will continue to be available in the future, on new platforms, meeting new needs"*

https://danielskatzblog.wordpress.com/2016/09/13/defining-software-sustainability/

ADORE .SOFTWARE

<ReSA> Research Software Alliance

RSE

SOFTWARE SUSTAINABILITY INSTITUTE

Findable Accessible Interoperable Reusable

Can GenAI tools help achieve RSE VISION?

TU Delft

How should AI course evolve? Teaching students how to develop reusable, sustainable, and robust research software

Ten simple rules for teaching sustainable software engineering

Kit Gallagher[1], Richard Creswell[2], Ben Lambert[3], Martin Robinson[2], Chon Lok Lei[4], Gary R. Mirams[5], David J. Gavaghan[1*]

1 Doctoral Training Centre, University of Oxford, UK
2 Department of Computer Science, University of Oxford, UK
3 Department of Statistics, University of Oxford, UK
4 Faculty of Health Sciences, University of Macau, Macau, China
5 Centre for Mathematical Medicine & Biology, School of Mathematical Sciences, University of Nottingham, UK

* david.gavaghan@dtc.ox.ac.uk

[K. Gallagher et al., 2024] https://arxiv.org/pdf/2402.04722

FAIR4RS DCC TUDELFT

RSE

RESEARCH SOFTWARE TRAINING NL

CODE REFINERY

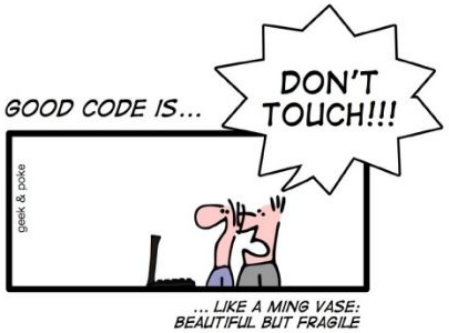
software carpentry

netherlands eScience center

TU Delft

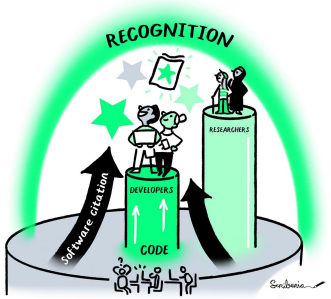
Students skills necessary to apply practices for reproducible research, and possibility to automate them with GenAI

Create container



Modular coding

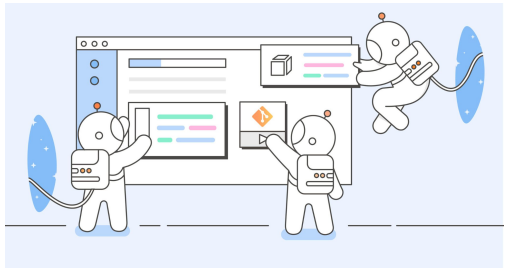
Unit Test



add .cff file

Peer-review

Create README



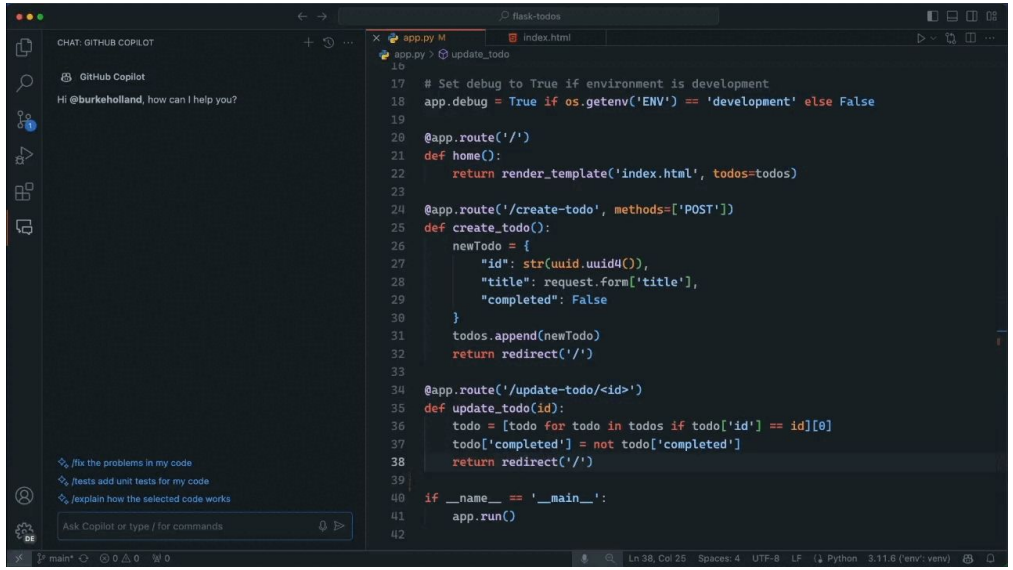
Create pull request

Publish Your Software in conda-forge

TU Delft

The potential benefits of AI code assistant: increase productivity, greater efficiency & collaboration, and speed up learning

Answer coding questions



OpenAI Codex: trained on **54 million Github repositories**

Fix Github-issues

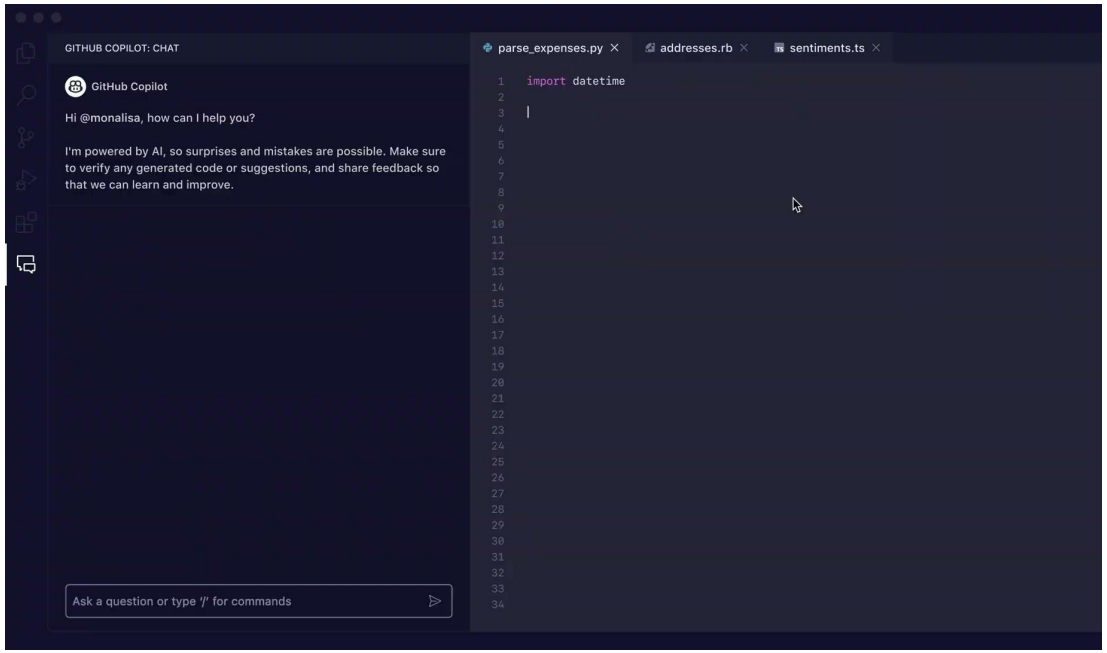
SWE Autonomous coding agent
[John Yang et al., 2024]
https://swe-agent.com/ @princeton-nlp

TU Delft

Other workloads AI code assistant can support:

- Run terminal commands
- Write pull request summaries
- Suggest code refactoring
- Chat with your codebase
- Generate commit messages

Generate unit test cases



https://github.blog/2024-01-22-10-unexpected-ways-to-use-github-copilot

TU Delft

The raise of *"free (?)"* AI-partner in code

https://www.tabnine.com/

https://codeium.com/

https://www.wispr.ai/

https://github.com/copilot

https://cursor.sh/

https://zed.dev/

https://www.jetbrains.com/ai/

https://useadrenaline.com/

https://aws/codewhisperer/

https://snyk.io/

https://alphacode.deepmind.com/

https://replit.com/learn/intro-to-ghostwriter

TU Delft

Mentimeter

Which tools?

menti.com | use code **5384 8806**

https://www.menti.com/alxy1r19nrye

TU Delft

Mentimeter

For the AI tools you use as part of your RSE development workflow, what are the MOST important benefits you are hoping to achieve? Please check all that apply.

- Increase productivity
- Speed up learning
- Greater efficiency
- Improves accuracy in coding
- Improve collaboration

menti.com | use code **5384 8806**

https://www.menti.com/alxy1r19nrye

TU Delft

Mentimeter

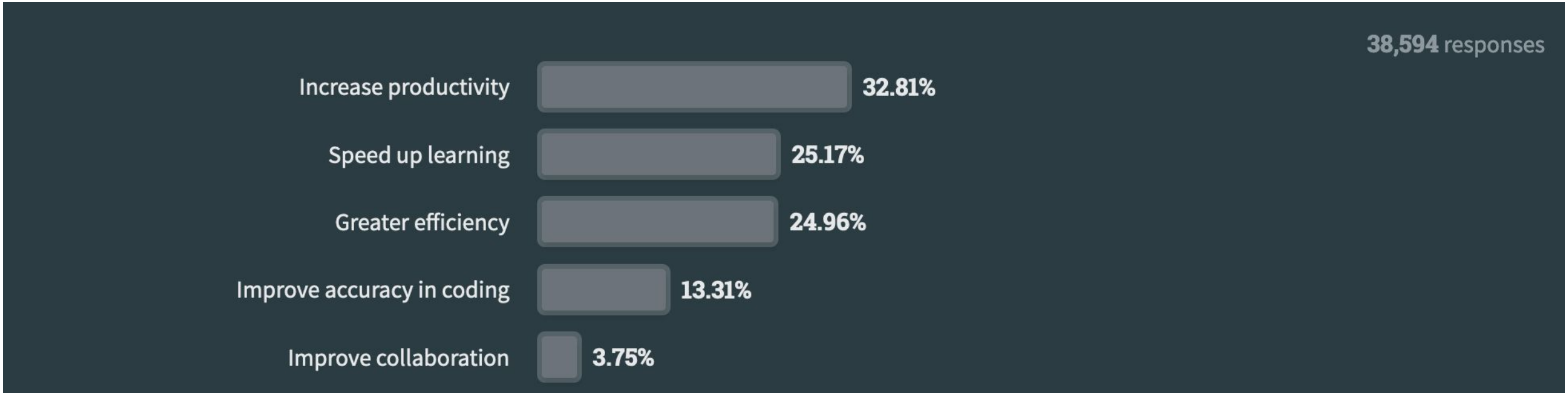
- Which parts of your development RSE workflow will benefit from GenAI tools?

menti.com | use code **5384 8806**

https://www.menti.com/alxy1r19nrye

TU Delft

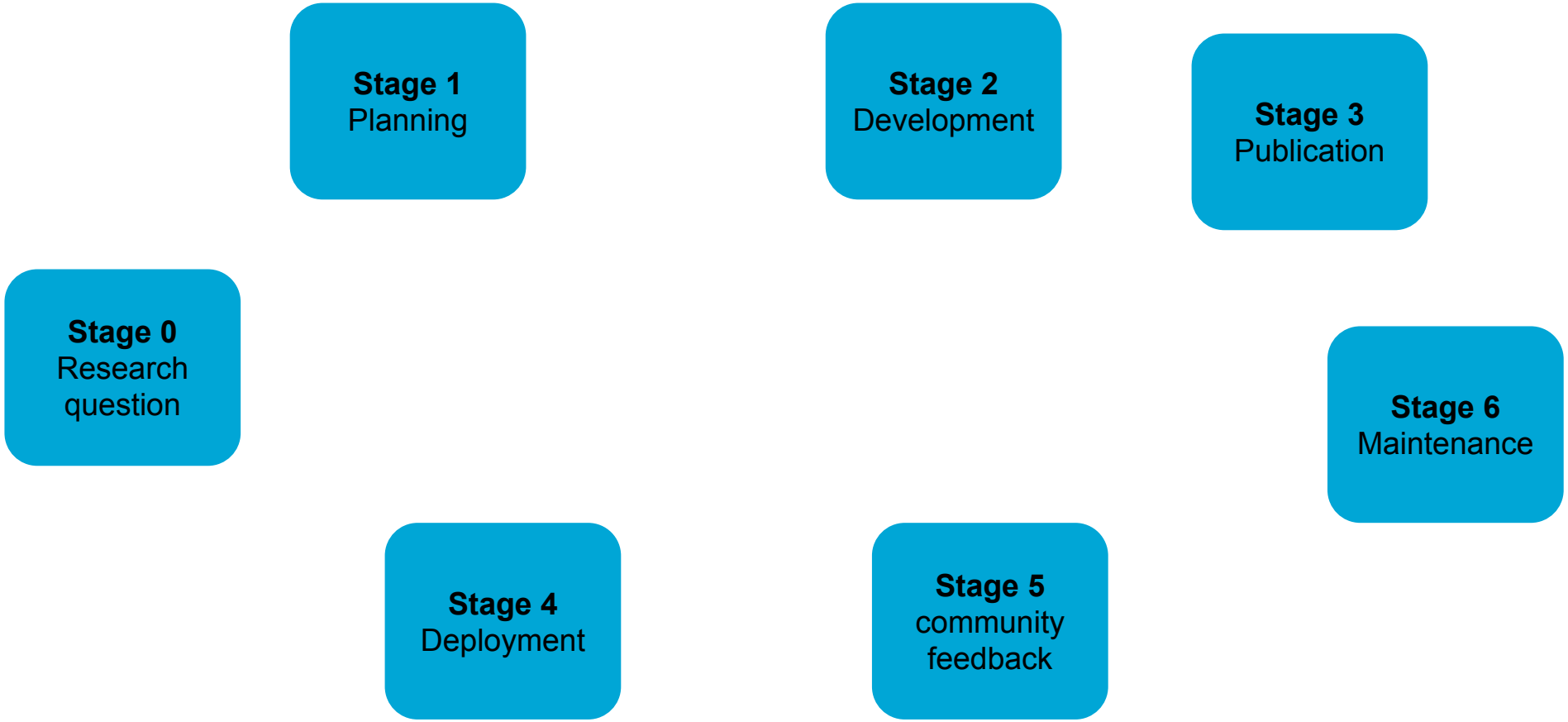
Benefits of AI tools:



source: https://survey.stackoverflow.co/2023/#section-developer-tools-benefits-of-ai-tools

TU Delft

Discussion: Which RSE areas will benefit from GenAI tools?

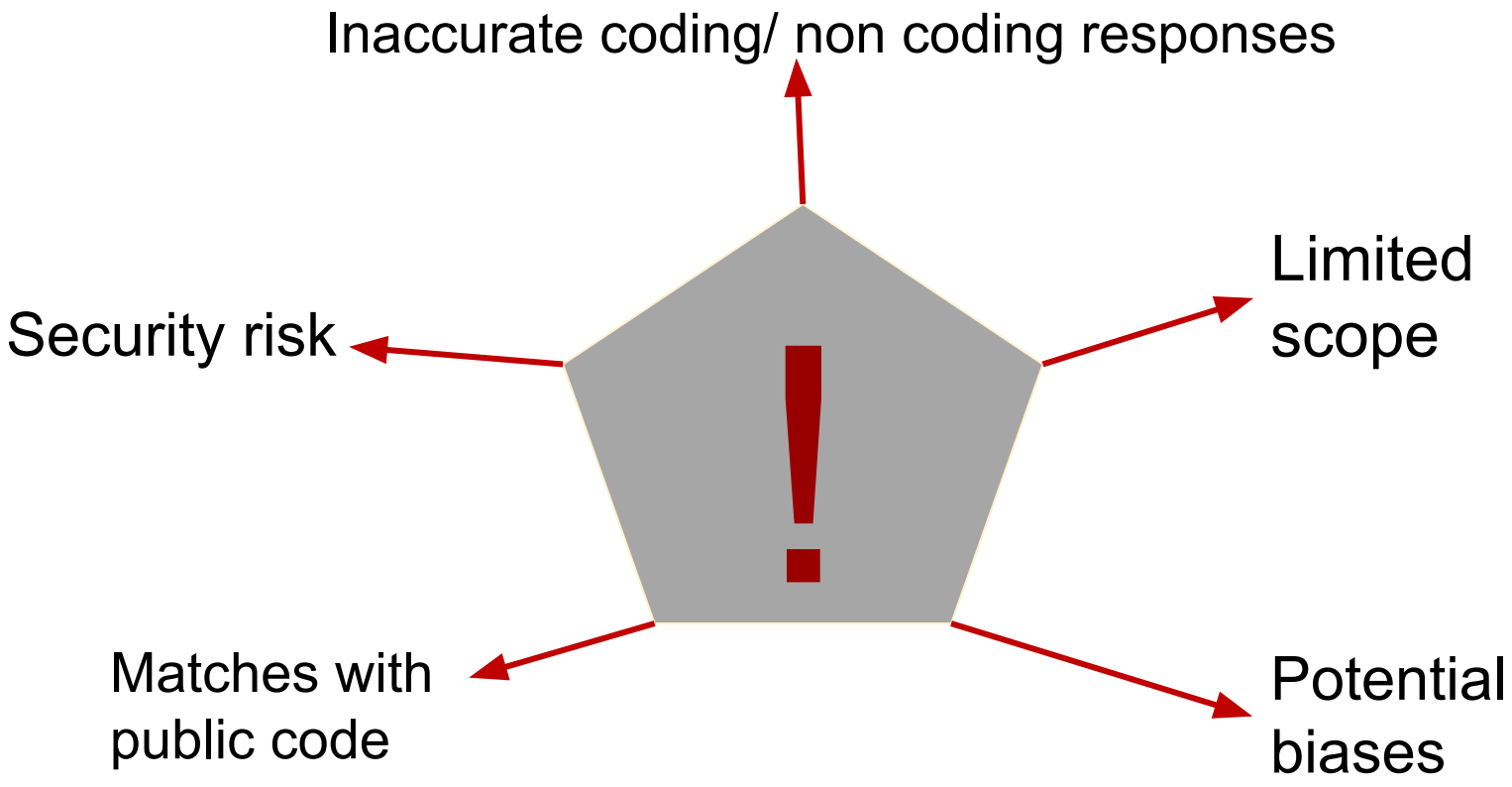


TU Delft

https://docs.github.com/en/codespaces/developing-in-a-codespace/creating-a-codespace-for-a-repository

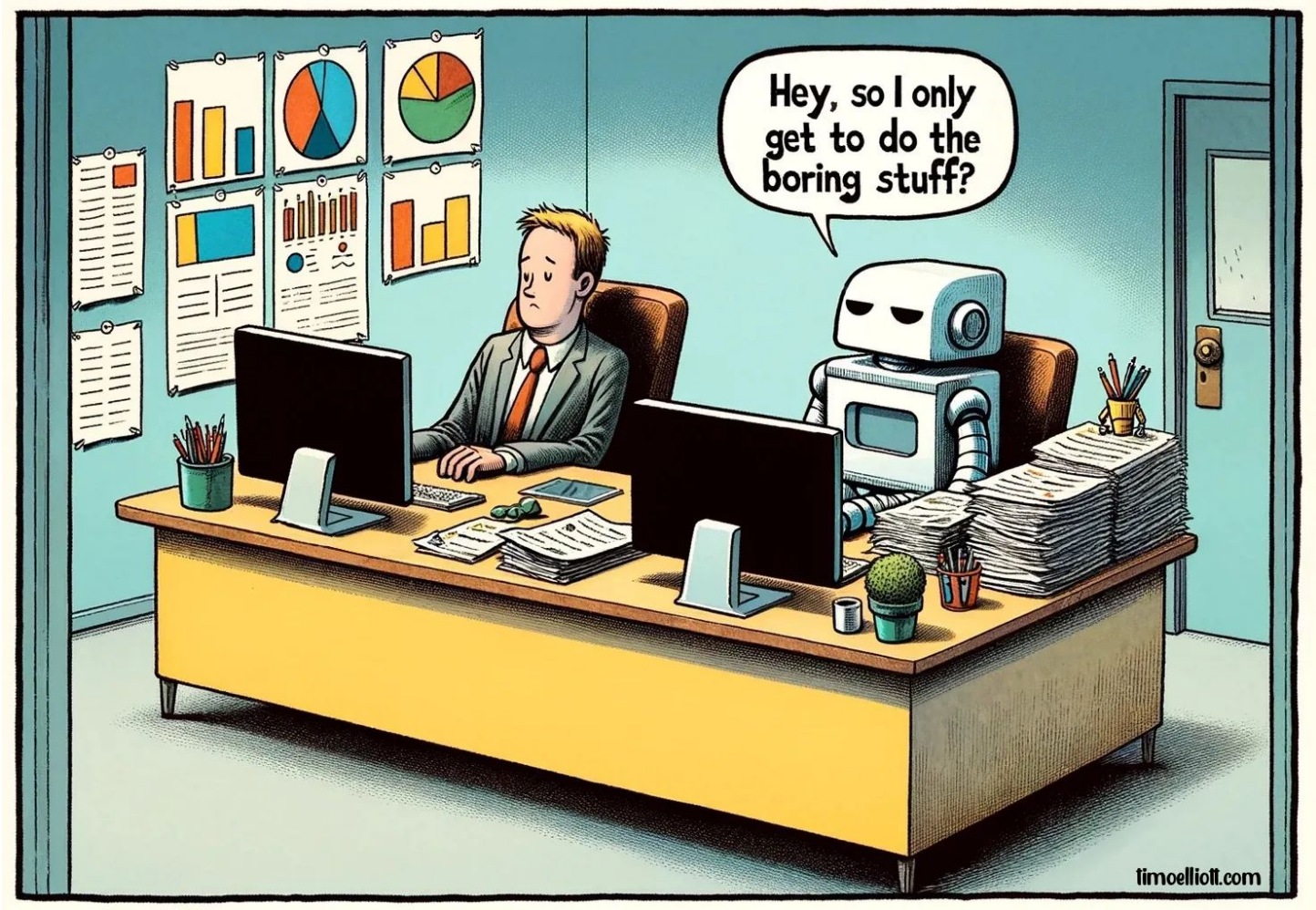
Three modalities and its limitations

1. Github CLI
2. Web browser
3. IDE editor



https://docs.github.com/en/copilot/github-copilot-chat/copilot-chat-in-ides/about-github-copilot-chat-in-your-ide#limitations-of-github-copilot-chat

Tendency of sloppy thinking?



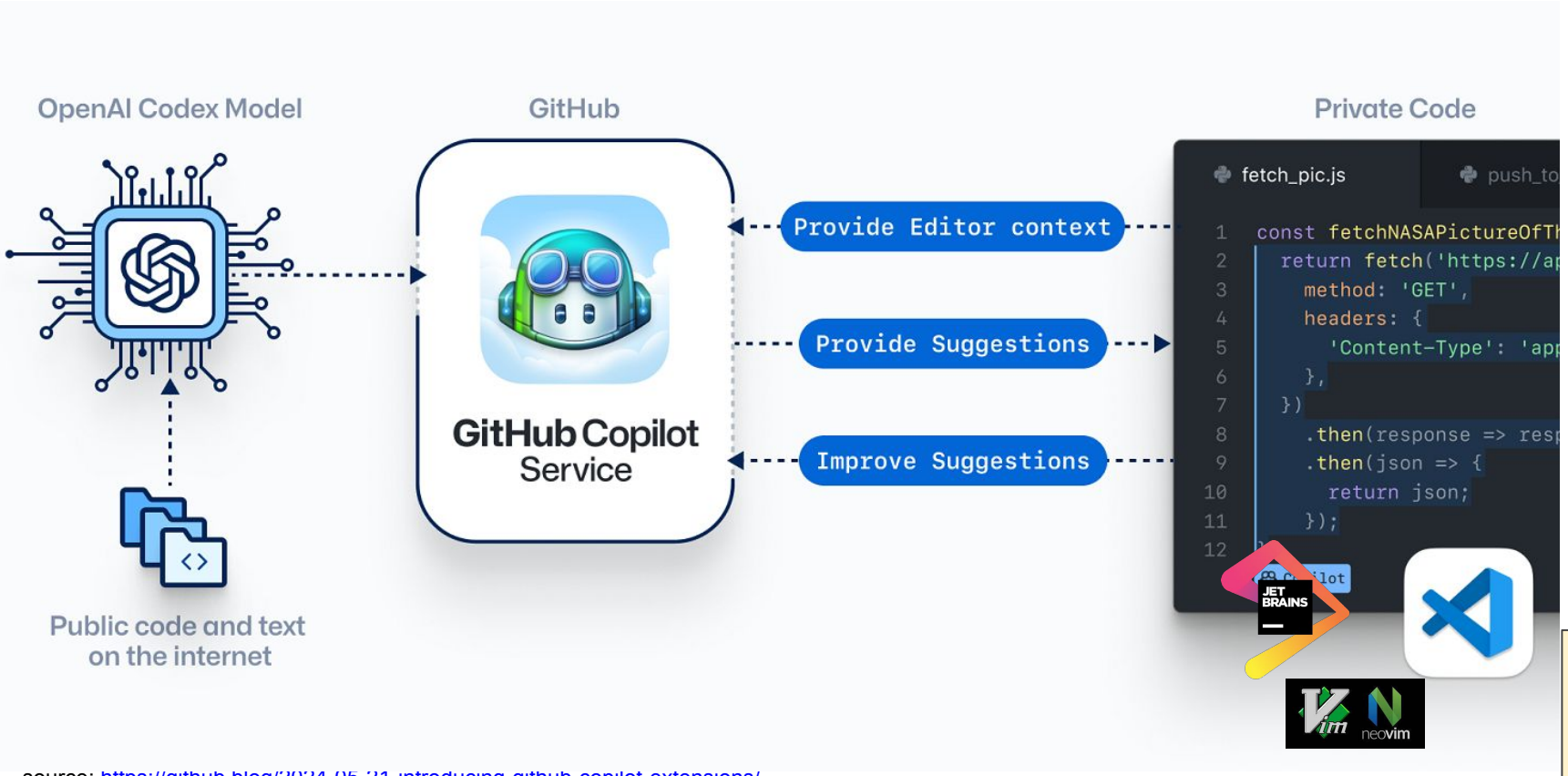
https://timoelliott.com/blog/cartoons/artificial-intelligence-cartoons

TU Delft

Part 4
Demo: Github Copilot

Collaborative notes:
https://edu.nl/qx9qj

Copilot promise: researchers can stay in the flow longer, uplevel their skills, and innovate faster.



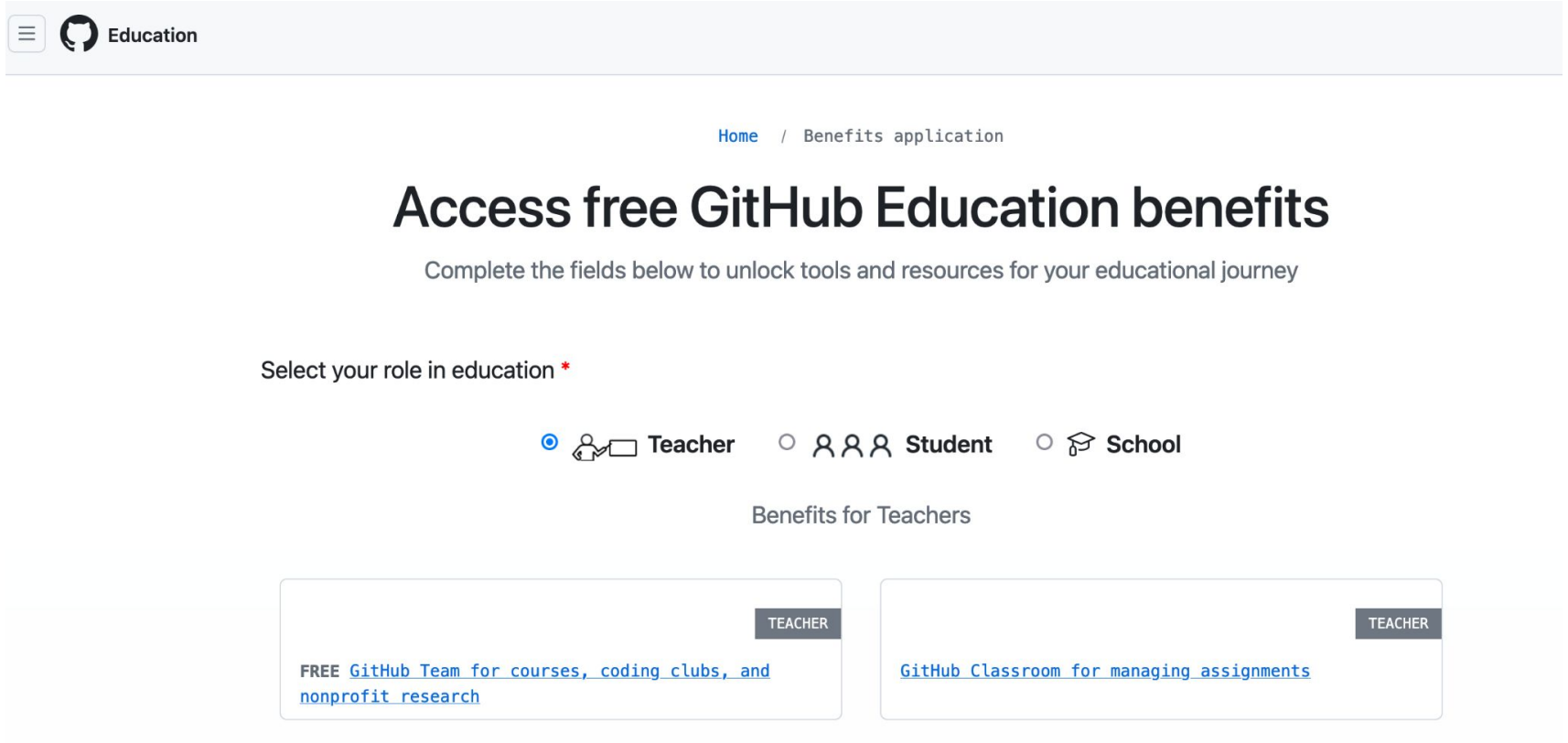
source: https://github.blog/2024-05-21-introducing-github-copilot-extensions/

| *"Verified students, teachers, and maintainers of popular open source projects on GitHub are **eligible to use Copilot Individual for free**." [1]*

! WARNING: Its not an easy process to get verified

[1]: https://docs.github.com/en/copilot/copilot-individual/about-github-copilot-individual#about-billing-for-github-copilot

Join Github Education



Collaborative notes:
https://edu.nl/qx9qj

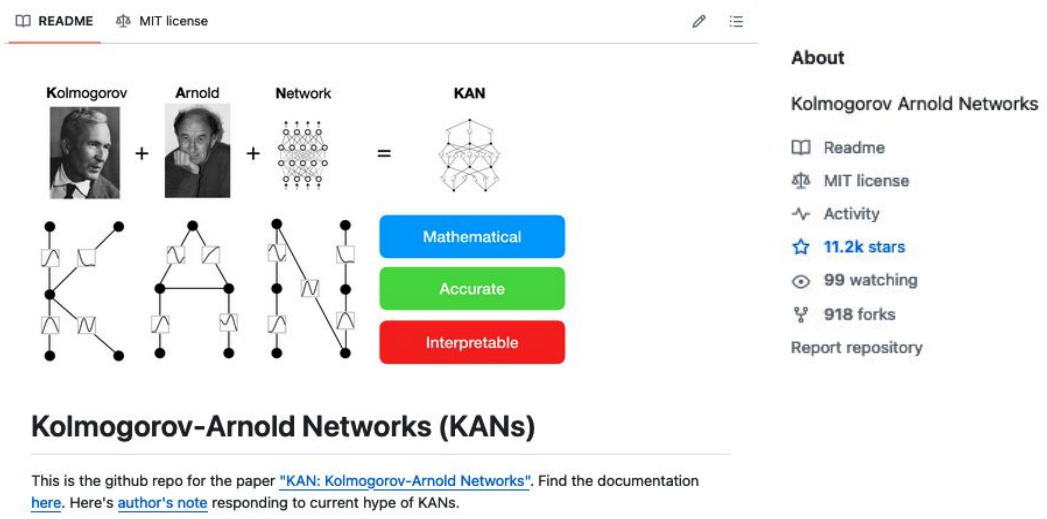
edu.nl/qx9qj

TU Delft

Live demo: How to best interact with Github Copilot

- Submitting prompts
- Using keywords in your prompt
- Slash commands
- Chat variables
- Asking right questions/ prompts:
  - Ask general software questions
  - Write code
  - Setup a new project
- Additional ways to access Copilot Chat

Repo: https://github.com/KindXiaoming/pykan



source: https://docs.github.com/en/copilot/github-copilot-chat/copilot-chat-in-ides/using-github-copilot-chat-in-your-ide#example-prompts

TU Delft

Mini-demo: Halford Unit test

TU Delft

Mini-demo: Structure codebase using conventions -- Cookiecutter vs GitHub Copilot

- Task 1: Install cookiecutter [1]
- Task 2: Find a template for ML projects
- Task 3: Use cookiecutter functionality

- Task 4: Open new repository in VSCode
- Task 5: Check Github Copilot status
- Task 6: Ask Copilot Chat [2]
- Task 7: Execute suggested commands
- Task 8: Review and discuss outputs

[1]: https://cookiecutter-pypackage.readthedocs.io/en/latest/tutorial.html
[2]: https://code.visualstudio.com/docs/copilot/overview#_jumpstart-your-project
Repo: https://github.com/carlosug/ai4rse-workshop-TUD-symp2024

https://miro.medium.com/v2/resize:fit:640/format:webp/0*zAXvoC2SceDxvzBP.jpg

TU Delft

Mentimeter

Thinking about how your workflow and process changes over time, how similar or different do you anticipate your workflow to be 1 year from now as a result of AI tools you are currently using?

- very different
- somewhat different
- neither different nor similar
- somewhat similar
- very similar

TU Delft

Mentimeter

How much do you trust the accuracy of the output from AI tools as part of your RSE workflow?
Highly trust

Somewhat trust

Neither trust nor distrust

Somewhat distrust

Highly distrust

TU Delft

Activity with mentimeter round discussion (if time)

Possible positive impact

Increasing productivity and efficiency in coding

Improving code documentation and reusability

Reviewing code in real-time, increasing quality of publicly available code

Eliminate repetition

Keep you in the flow

Possible negative impact

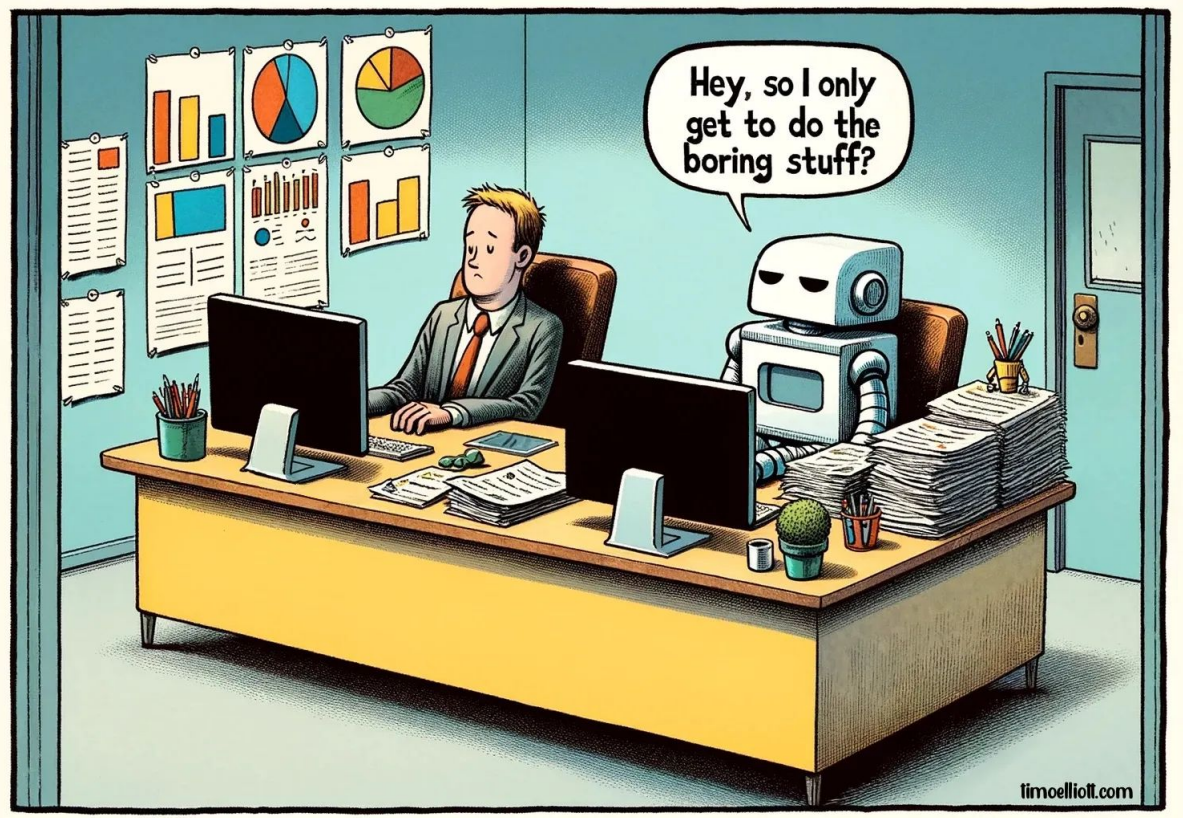
may lack critical skills for assessing code - sloppy thinking

Deprecated suggestions software libraries or APIs which have since been deprecated, so that code dependencies no longer work, and neither does the code

source: https://osf.io/preprints/socarxiv/zns7g
mentimeter? Good, bad and mixed - ugly table

TU Delft

Wrap-up: Build software faster using AI tools, but think twice before integrate into your workflow



3 key takeaways:

- **GenAI has plenty of potential to aid with RSE, and to help us get productive more quickly**
- **GenAI is a tool, and effective use still requires skill and good judgement**
- **After “the hype cycle” we will have a new generation of tools, but not as dramatic a revolution as predicted**

TU Delft

Add topics of your interest for the next Gen-AIforSE workshop

Collaborative notes:
https://edu.nl/qx9qj

edu.nl/qx9qj

TU Delft

Workshop AIforRS

Thank you for your attention!

Date: 4 June 2024

Organized by: Research Data and Software Team (RDS)
TU Delft University | Library

TU Delft

Image artificially generated using openart.ai with prompt "illustrate a workshop on AI tools for research software in Delft University"