

# Per Anhalter durch die ChatGPT-Galaxis

**Nutzen und Grenzen der Verwendung künstlicher Intelligenz  
zur Java-Entwicklung**

Jiri Lundak - 11.12.2023

[jiri.lundak@red-pill.ch](mailto:jiri.lundak@red-pill.ch)  
[LinkedIn](#)

# Agenda

- Wir entdecken eine uns unbekannte Domäne
- Wir schreiben User Stories für ein neues Produkt
- Wir lassen ChatGPT Code für die User Stories generieren
- Wir lassen ChatGPT Tests für den Code und die User Stories generieren
- Wir modifizieren den generierten Code
- Wir lassen ChatGPT Code erklären
- Wir lassen ChatGPT Code modifizieren
- Wir hinterfragen die Ergebnisse




# Metadata

- We use ChatGPT 4 in chat mode
- Data model was trained on as of: September 2021 (still)
- Token Limit: 8000
- Temperature: 0
- We do not use CoPilot (no IDE integration)
- Still no Code Interpreter for Java available from within ChatGPT (only for Python for „Advanced data analysis“)
- „Your mileage may vary!“ (non deterministic road ahead)

# Heuristic

## Too Generic Requirements

### Consequences

- You get a rough idea about a new domain 
- You might be tempted to not look for real user input 
- Specifics are left out 

### Coping Strategy

- You still need real users
- You can get the prototype out to the world, but will need real user feedback quick
- You need to dig deeper into the use case and do background research

# Heuristic

## No Extrapolation / Generalisation

### Consequences

- Always proposes a solution 
- Applies rules without thinking 
- Does not take (wider) context into account that you do not provide 



### Coping Strategy

- When in doubt provide additional context and refine your prompt
- Know your stuff (don't be a shallow, lazy thinker)
- Use your ability to generalise

# Heuristic

## Codebase Out of Date

### Consequences

- You get a rough idea what the import or class name or method could be 
- Old or non-existing imports, classes, methods referenced 








### Coping Strategy

- Take proposed code with a grain of salt
- Verify last version of language, lib or framework available the model is most probably employing (know your release history). It won't tell you.
- Double check in your IDE

# Heuristic

## Made up responses (Hallucination)

### Consequences

- Result looks very credible on the outset  
- It gives you something to ponder about  
- Library for your language might be unavailable 
- Code matches non-existent lib  






### Coping Strategy

- Prompt: „Which of the choices you presented do not exist?“
- Let it generate a Maven or Gradle include (it will make it up also sometimes)
- Integrate fast (in your IDE)

# Heuristic

## Compilers Never Lie

### Consequences

- Going extra mile to integrate proposed solution  
- Time wasted going down wrong path 
- Time wasted undoing things  

### Coping Strategy

- Use your local repo (save working version, rollback quickly when needed)
- When presented alternatives (always go with the first proposed) -> Chances are others down the list are fake solutions



# Heuristic

## Overcomplicates Solutions

### Consequences

- Gives hints at technology one could use 
- Can lead to overblown codebase 
- Code might even compile  

### Coping Strategy

- Ask yourself: „Is the right problem solved?“, „What is the tool/lib used for?“
- Before looking up newer versions of library, check if you need it!

# Heuristic

## Temporary Memory Loss

### Consequences

- Answer is seemingly given (partially) out of context ✘
- Repeating older version of proposed solution (without corrections) ✘
- Solution has nothing to do with question ✘

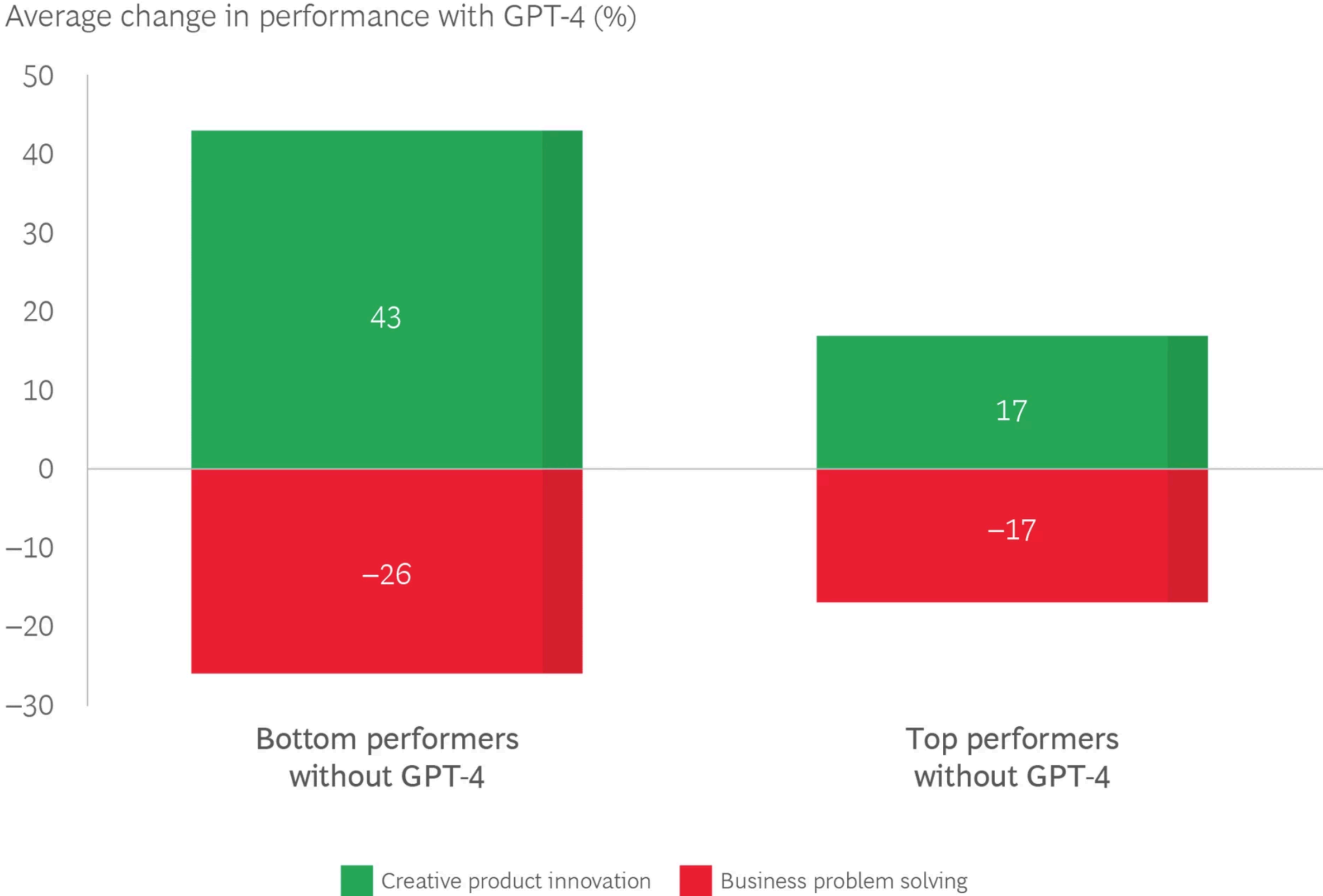
### Coping Strategy

- Restate context
- Repeat current state of code to make it remember
- Correct to answer
- Note: This problem will be alleviated soon (chatgpt-4-1106-preview) holds context of up to 128'000 tokens (instead of just 8'000)

# Question

Does ChatGPT make you better?

Exhibit 2 - Generative AI's Performance Effect Is Strongest When Baseline Proficiency Is Lowest



Source: <https://www.bcg.com/publications/2023/how-people-create-and-destroy-value-with-gen-ai>

# Conclusions

## How to leverage ChatGPT today?

### What it is good for

- Get acquainted with new stuff (programming languages, libraries, a new domain)
- Initial set of user stories (needs to get refined)
- Initial skeleton of a little application (refining it in dialog)
- Initial set of acceptance tests
- Snippets of logic (size limited by token limit -> resort to uploading files is not an efficient way to go)
- Let it explain some code (hopefully the naming of classes and methods makes sense, as it infers what the program does from the naming) -> try some obfuscated code to see where the limits are. 😊

# Conclusions

## How to leverage ChatGPT today?

### Where the use is dubious

- Everything that involves numbers is at risk getting lost in the model (ChatGPT is still bad at math). -> but that is no surprise (it is a language model!)
- As ChatGPT never says NO, insisting in a specific direction can be misleading. So, when you know what you do, say what you don't like. You will get so many apologies. 😊
- Never let ChatGPT generate Gradle/Maven dependency definitions! -> Most of the time there are newer versions around. Just look them up directly on the net.
- Background checking is key. ChatGPT's model is still shallow. Even when it answers authoritatively it might be wrong. -> Alternative: know an expert!

# Conclusions

## How to leverage ChatGPT today?

### When not to use it

- For large scale tasks -> It is still not made for it (Token limit strikes, workarounds are cumbersome)
- When the domain needs speciality knowledge (like astronomy 😊) -> ChatGPT has ingested many examples of login dialogs
- Finding border cases needs deeper domain knowledge -> ChatGPT does not have that yet
- When I do not have the domain expertise to be able to question and judge the results -> Your customer is still your best friend



**Thank you!**

**Let's have a human chat now ;-)**

