

# Prompt Engineering for Test Automation: Enhancing AI-Driven Quality Assurance



@ludovicobesana

# Who I am?

- Senior Test Engineer
- SheTech Ambassador
- Grafana Champion
- Startups Co-Founder
- Open Source Contributor
- Some passions: 🚗, 🐕, 🎮



**Nearform.**

## OSS at Nearform\_

“At the heart of Nearform’s mission is a profound commitment to the Node.js ecosystem and the OpenJS Foundation. Their support and contributions are instrumental in driving the growth and success of the open-source community.” - Robin Bender Ginn, Executive Director, OpenJS Foundation

15m+  
Downloads

Over 15 million  
downloads every week



Node JS

1192k+  
Contributions

Thousands of contributions  
every year



Fastify

166+  
Packages

We maintain many  
popular packages for the  
community



Clinic JS

10+ years  
experience

We are a trusted partner for  
our international clients,  
passionately delivering  
digital excellence across a  
range of industries.



Pino



React Native Auth

We're hiring! [nearform.com/careers](https://nearform.com/careers)



FAST FORWARD

ACCURACY

SPEED

RELIABY

SPEED

ACCIRACY

ACCORACY

ACCURACY

SPEED

SPEED

RELIABY

BOOT

# Why This Talk?

---



Solve Real QA  
Challenges



Gain Practical,  
Hands-On Skills



Stay Ahead with  
AI-Driven QA



# WHAT IS PROMPT ENGINEERING?



@ludovicobesana

# Prompt Engineering

---

**Definition:** Designing precise, contextual inputs to guide AI models toward specific outputs.

**Purpose in QA:** Enables the creation of test cases that simulate user actions, handle edge cases, and validate outputs.





# Before Starting I

---

## **Clear Success Criteria:**

Define what success looks like for your specific use case. Establish measurable goals and outcomes that the model's responses should meet.

## **Evaluation Methods:**

Set up ways to test prompt performance empirically against your success criteria. This ensures that you can track improvements and identify any gaps.



# Before Starting II

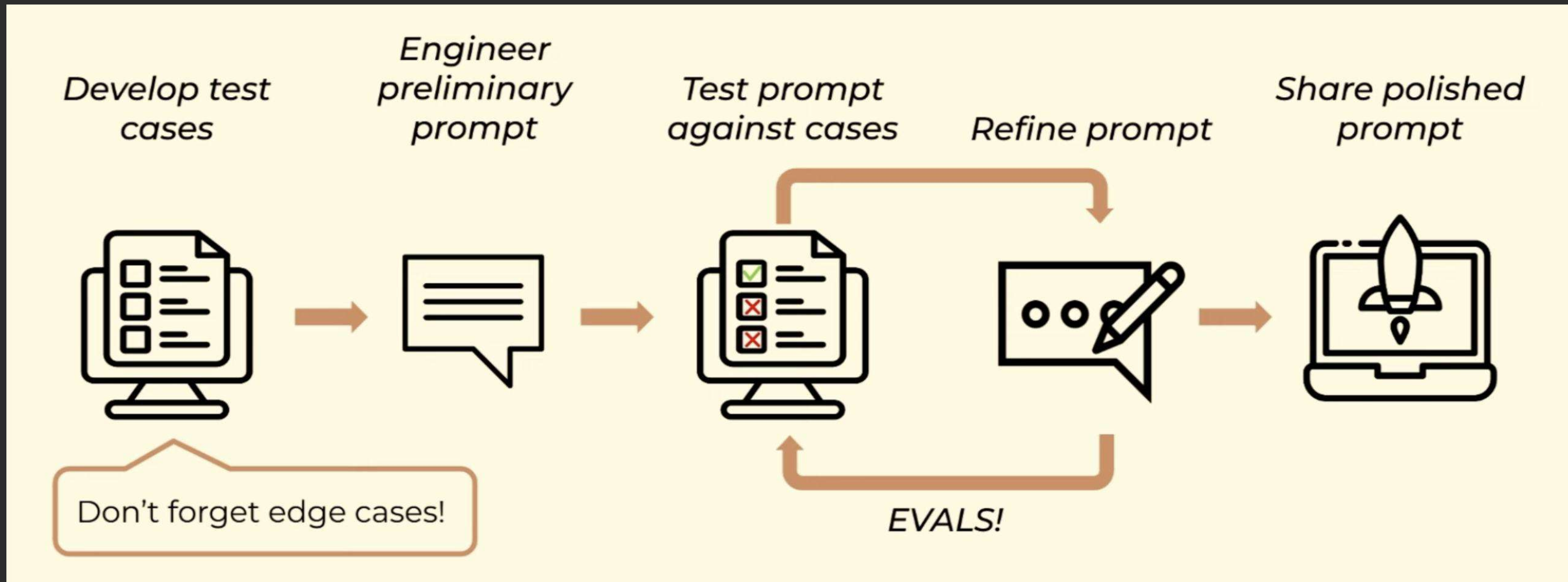
---

## **First Draft Prompt:**

Begin with an initial prompt that you'll improve through iterations.

If you don't have one yet, consider using a prompt generator tool, such as the one available in the Anthropic Console.





[docs.anthropic.com/en/docs/prompt-engineering](https://docs.anthropic.com/en/docs/prompt-engineering)



Prompt: You are a Senior Test Engineer with 10 years of experience. You will be provided with UI screenshots of a website. Your task is as follows:

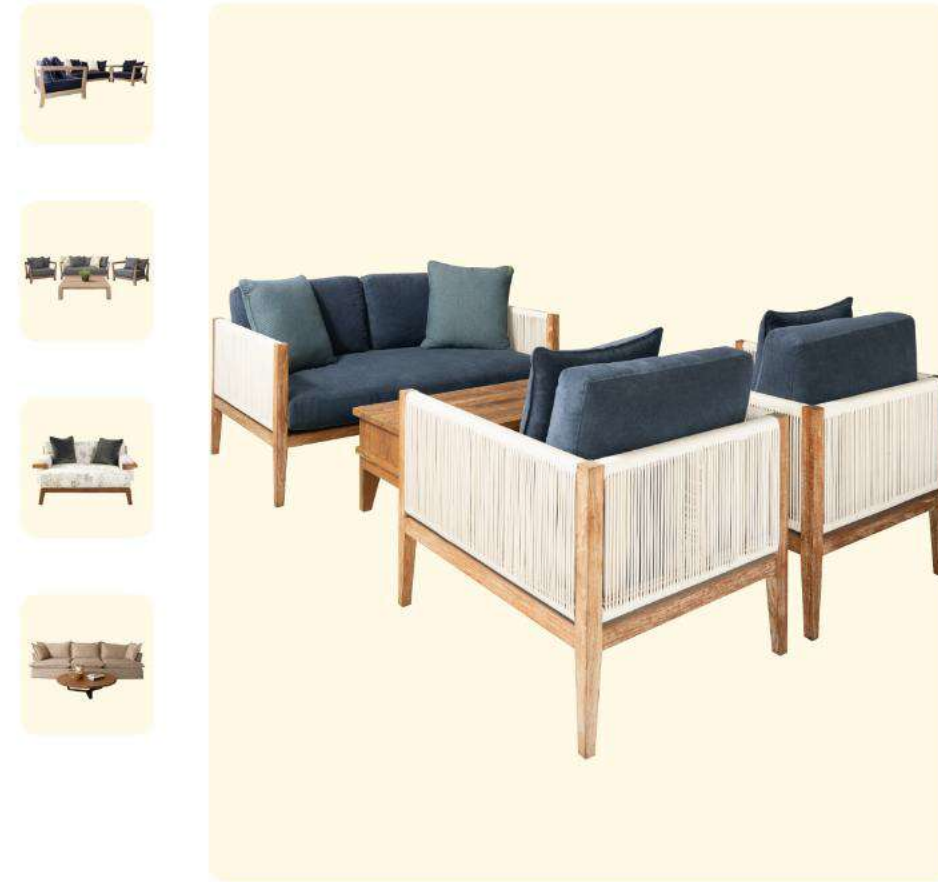
Generate all test cases using the GIVEN-WHEN-THEN format, analyzing visible user interactions and underlying functionalities. Organize the results in a table with four columns: "Test Title," "GIVEN," "WHEN," "THEN."

Each row of the table should represent a distinct test case, clearly describing the initial conditions, the action, and the expected outcome. Export the table to an Excel file and provide a link for downloading it.

The output should also be displayed in the chat as a table. Ensure that the tests are thorough and cover all possible flows identified from the screenshots.



@ludovicobesana



# Asgaard sofa

Rs. 250,000.00

★★★★★ | 5 Customer Review

Setting the bar as one of the loudest speakers in its class, the Kilburn is a compact, stout-hearted hero with a well-balanced audio which boasts a clear midrange and extended highs for a sound.

Size

L XL XS

Color



- 1 + Add To Cart

SKU : SS001

Category : Sofas

Tags : Sofa, Chair, Home, Shop

Share :



Description Additional Information Reviews [5]

Embodying the raw, wayward spirit of rock 'n' roll, the Kilburn portable active stereo speaker takes the unmistakable look and sound of Marshall, unplugs the chords, and takes the show on the road.

Weighing in under 7 pounds, the Kilburn is a lightweight piece of vintage styled engineering. Setting the bar as one of the loudest speakers in its class, the Kilburn is a compact, stout-hearted hero with a well-balanced audio which boasts a clear midrange and extended highs for a sound that is both articulate and pronounced. The analogue knobs allow you to fine tune the controls



### Test Cases For Single Product Page

	Title	GIVEN	WHEN	THEN	
1	Test product page loads correctly	User is on the product page	The product page loads	The product details are displayed with options to select size and color	
2	Test selecting size L	User selects 'Size L'	The user clicks on the 'L' button in Size options	The size 'L' is selected, and the button appears highlighted	
3	Test selecting color Black	User selects 'Color Black'	The user clicks on the 'Black' color option	The color 'Black' is selected, and the button appears highlighted	
4	Test increasing product quantity	User increases the quantity to 2	The user clicks the '+' icon to increase the quantity	The quantity is updated to 2 in the input box	
5	Test adding product to cart	User clicks 'Add to Cart'	The user clicks the 'Add to Cart' button	The product is added to the cart and the cart icon is updated with the new total	
6	Test viewing product description	User clicks on the 'Description' tab	The user clicks the 'Description' tab	The 'Description' section content is displayed	
7	Test viewing additional information	User clicks on the 'Additional Information' tab	The user clicks the 'Additional Information' tab	The 'Additional Information' section content is displayed	
8	Test viewing customer reviews	User clicks on the 'Reviews' tab	The user clicks the 'Reviews' tab	The 'Reviews' section content is displayed	
9	Test navigating to related product	User clicks a related product	The user clicks on a related product thumbnail	The user is navigated to the selected product's page	

tabella con quattro colonne: "Titolo del Test", "GIVEN", "WHEN", "THEN".  
 Ogni riga della tabella deve rappresentare un caso di test distinto e descrivere in modo chiaro le condizioni iniziali, l'azione e il risultato atteso.  
 Esporta la tabella in un file Excel e forniscimi un link per scaricarlo.  
 L'output deve essere mostrato anche in chat sotto forma di tabella.  
 Assicurati che i test siano completi e coprano tutti i possibili flussi rilevati dalle schermate.

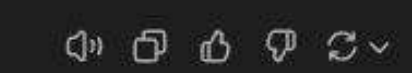


**Test Cases For Singl...**

Ho generato i casi di test per la nuova schermata UI, inclusa la colonna con i titoli per ciascun caso. Puoi scaricare il file Excel utilizzando il link qui sotto:

[Scarica il file dei test case](#)

I test sono anche visualizzati qui sotto in forma di tabella. [-]



Scrivi un messaggio a ChatGPT



@ludovicobesana

# Challenges in Traditional TA

---



Limited  
adaptability



High manual  
effort



Stability issues  
in CI/CD



# Key Techniques in Prompt Engineering

- **Zero-Shot Prompting:** Directly generate answers without examples to test adaptability.
- **Few-Shot Prompting:** Provide a few examples to guide the model, improving response accuracy for specific tasks.
- **Chain-of-Thought Prompting:** Break down complex tasks into logical, step-by-step instructions to improve clarity and ensure accurate outputs.





# Moving Beyond Manual Prompt Eng.

---

**Trial-and-Error Limitation:** Manual prompts are often inconsistent.

**Structured Optimization with DSPy:** Data-driven frameworks improve reliability.



**How can we efficiently manage and track the performance of different prompts in a dynamic QA environment?**





# PromptLayer

# Using PromptLayer

---

**Centralized Prompt Management:** Organize and version prompts in one platform, allowing easy tracking and updates.

**Performance Analytics:** Log and analyze prompt effectiveness to see which versions perform best across different testing scenarios.

**Optimization for QA:** Identify high-performing prompts and quickly apply updates to improve consistency across tests.



The image shows a screenshot of the PromptLayer web application. The interface is clean and modern, with a light blue and white color scheme. At the top, there is a navigation bar with the PromptLayer logo on the left and several menu items: 'Personal' (with a dropdown arrow), 'Home', 'Registry', 'Evaluate', 'Analytics', and 'Playground'. A settings gear icon is on the far right. Below the navigation bar, there are two tabs: 'Fine-Tune' and 'Datasets'. A search bar is present, followed by a 'Tags' button. The main content area is divided into a left sidebar and a main panel. The sidebar contains 'Requests' and 'Traces' buttons, a 'No requests found' message with a sad face icon, and a link to 'Click here to turn off the default date range limit'. The main panel features a large 'Welcome to PromptLayer' banner with a laptop illustration and a 'Read our docs' link. Below the banner are four feature cards, each with an icon, a title, a statistic, a description, and a 'Learn more' button. The cards are: 1. 'Create a prompt template' (lightbulb icon, 140k+ templates created), 2. 'Deploy templates to production' (cloud icon, Avoid being blocked by eng releases), 3. 'Log your first request' (calendar icon, 40M+ requests logged), and 4. 'Evaluate your prompts' (question mark icon, 900k+ tests run). A chat bubble icon is visible in the bottom right corner of the main panel.



@ludovicobesana

## Create a New Prompt Template

Learn more about templates on our [docs](#).

Chat  Completion

Title f-string ▾

**System** ▾

**Roles**

- System ✓
- User
- Assistant**
- Function
- Tool
- Placeholder

+ New Message

Copy ↗

*Define an input variable with {brackets}* Metadata Parameters Functions & Output Cancel Create Template



@ludovicobesana

# Building CustomGPT for Each Project

---

- Customize GPT with user stories, technical details.
- Helps AI provide project-specific responses.



# Using Documentations for CustomGPT

---




- **Integrate Project-Specific Documentation:** Embed APIs, libraries, and technical references directly into CustomGPT for contextualized responses.
- **Enhanced Accuracy:** Provides AI with a clearer understanding of project-specific elements, resulting in more relevant and precise outputs.
- **Code-Driven Prompts:** Leverage documentation to generate prompts that align with best practices and technical standards for each project.





Crea Configura



**Nome**  
Maestro Docs

**Descrizione**  
A support chatbot for code generation based on Maestro documentation. It answers questions and gene

**Istruzioni**  
Main Functions: Answer questions about Maestro and generate code based on the provided official documentation.  
Documentation-based Responses: Provide detailed and accurate information, sourcing from the available documentation. If a question is not covered, indicate that the answer is unavailable.  
Code Generation: When asked to generate code, refer to documentation files relevant to supported platforms, such as [ios-uikit.md](#), [android-jetpack-compose.md](#), [flutter.md](#), etc. Provide clear, concise


**Spunti di conversazione**

- How can I generate code for a specific platform using Maestro? ×
- How do I run a sample flow in Maestro? ×
- What are the requirements to run Maestro on a CI system? ×
- ×

**Informazioni**  
Se carichi file nelle Informazioni, le conversazioni con il tuo GPT potrebbero includerne i contenuti. Puoi scaricare questi file quando lo Strumento di interpretazione del codice è abilitato

[web-views.md](#) [supported-platforms.md](#)




Anteprima



**Maestro Docs**  
A support chatbot for code generation based on Maestro documentation. It answers questions and generates code using Maestro's official information.

- How can I generate code for a specific platform using...
- How do I run a sample flow in Maestro?
- What are the requirements to run Maestro on a ...

Scrivi un messaggio a Maestro Docs



@ludovicobesana



scaricare questi file quando lo Strumento di interpretazione del codice è abilitato

 <b>web-views.md</b> File	 <b>supported-platforms.md</b> File
 <b>SUMMARY.md</b> File	 <b>ios-uikit.md</b> File
 <b>README.md</b> File	 <b>android-jetpack-compo...</b> File
 <b>ios-swiftui.md</b> File	 <b>react-native.md</b> File
 <b>android-views.md</b> File	 <b>bug-report.md</b> File
 <b>running-flows-on-ci.md</b> File	 <b>writing-your-first-flow...</b> File
 <b>maestro-studio.md</b> File	 <b>frequently-asked-quest...</b> File
 <b>rollback-maestro.md</b> File	 <b>debug-output.md</b> File
 <b>flutter.md</b> File	 <b>run-a-sample-flow.md</b> File
 <b>known-issues.md</b> File	

@ludovicobesana

# Agents vs. Copy-Pasting Prompt Results

---

**Adaptability:** Agents provide real-time responses and can adjust to new data, making them more flexible in dynamic scenarios.

**External System Interaction:** Agents can engage with APIs, databases, and other external systems, enabling them to perform actions beyond simple text outputs.

**Enhanced Accuracy and Context Awareness:** Agents can respond with greater precision, pulling context from live data or system states.



# Understanding Knowledge Memory

---

- Stores past interactions for future recall.
- Useful for complex, sequential tests.



# What is Prompt Tuning?

---

- Optimizes prompt structure, not the model.
- Benefits intricate scenarios with dynamic needs.



# Best Prompting Techniques

---

- **Few-Shot Learning:** Use examples for nuanced tasks.
- **Chain of Thought:** Helps AI explain its reasoning.
- **ReAct:** Dynamic responses based on observations.



# Design Phase

---

- **Automatizability Scoring:** Rate test cases for automation potential.
- **User Personas Generation:** Tailor prompts for diverse user scenarios.



# Implementation Phase

---

- **XPath Name Suggester:** Helps identify UI elements.
- **Data-Driven Test Generation:** Generate diverse test data dynamically.
- **Automated Code Generation:** Leverage tools like Auto Playwright to generate test scripts directly from high-level descriptions, streamlining the scripting process and ensuring consistency.





# Reporting Phase

---

- **Failure Pattern Detector:** Identifies common issues.
- **Severity-Based Reporting:** Prioritizes critical issues.



# Reading Reports Phase

---

- **Non-Technical Summaries:** Translates results for non-technical stakeholders.
- **Test Impact Analysis on New Features:** Suggests testing focus for new features.



# 3 Take Aways

---



Prompt Engineering  
Transforms QA



Effective Tools and  
Techniques Exist



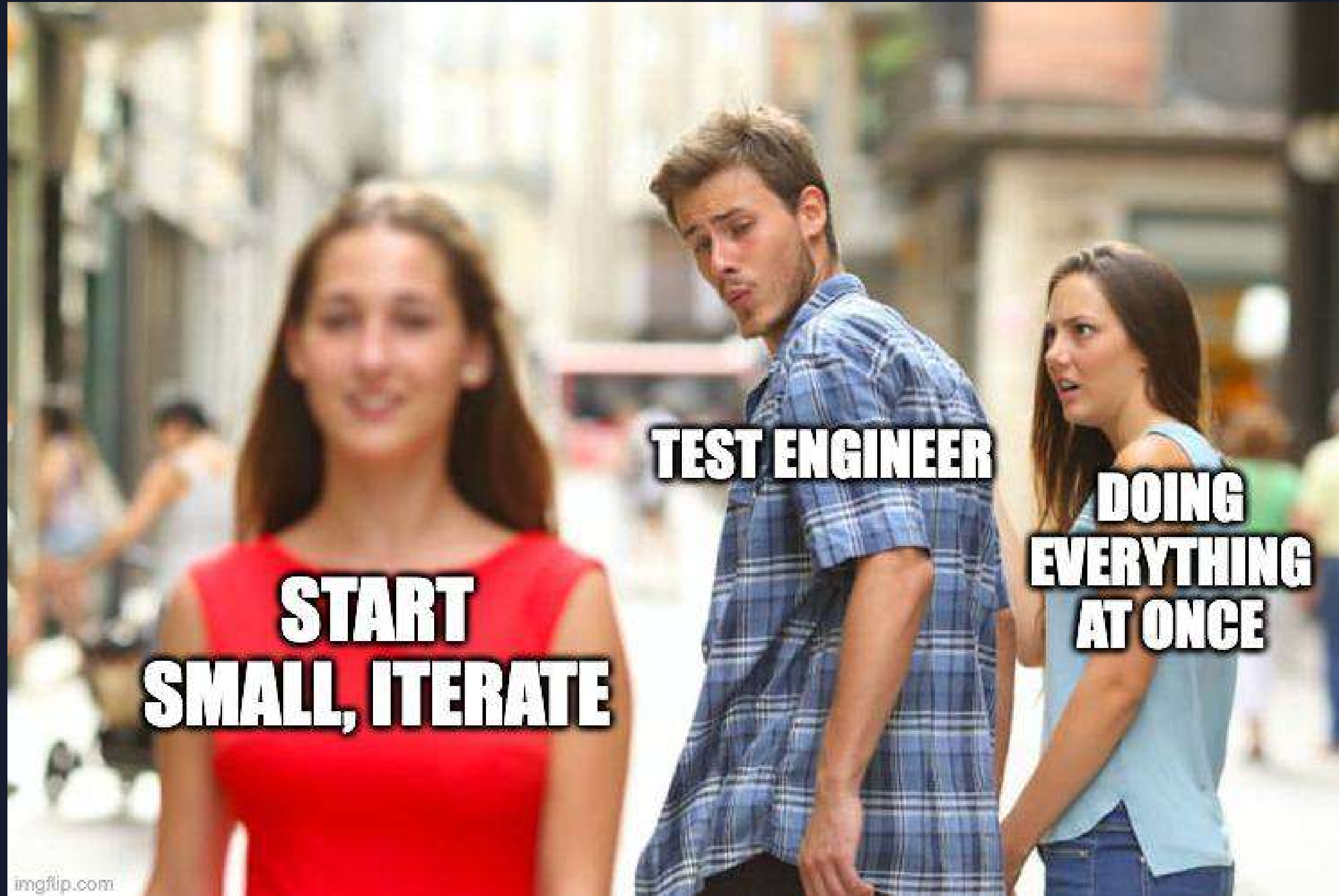
Cost Efficiency in  
QA Processes



# Experiment with Prompt Engineering in your testing workflows!



@ludovicobesana



**START  
SMALL, ITERATE**

**TEST ENGINEER**

**DOING  
EVERYTHING  
AT ONCE**



# Thank you



@ludovicobesana

# Meet Ludovico



@ludovicobesana