

Testautomatisierungslücke mit KI füllen

ERFAHRUNGENAUS DEM ISTQB GLOSSARY PROJEKT

Dr. Matthias Hamburg



1

Über mich

- Mathematiker und Informatiker
- Seit 1986 Berater in der IT-Industrie
- Seit 2020 im aktiven Ruhestand
- Ehrenamtlich aktiv in GI, GTB, ISTQB®

Meine Schwerpunkte:

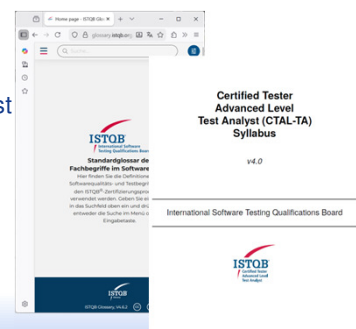
- Testglossar – Herausgeber des ISTQB-Glossars der Testbegriffe
- Testanalyse – Herausgeber des ISTQB Syllabus Advanced Test Analyst
- Testprozessverbesserung

Mein Hobby

- Automatisierung des Testentwurfs
- Beteiligt am Start-up Harmony.ac



Matthias Hamburg
Dr. rer. nat.



19.03.2026

Kann KI die Automatisierungslücke schließen?

2



2

Fragestellung

Die klassische Testautomatisierung ist noch unzureichend

Abhilfe durch generative KI ist in aller Munde

Wir wollen aber das Rad nicht neu erfinden

- Beibehaltung der klassischen Automatisierungspraxis
- Füllen der Lücken mit Hilfe generativer KI

Wie geht das?

19.03.2026

Kann KI die Automatisierungslücke schließen?

3



3

Agenda

Automatisierungslücken im Testprozess und ein KI-Lösungsansatz

Fallstudie: Testanalyse mit KI

Fallstudie: Testrealisierung mit KI

Zusammenfassung der Erfahrungen

19.03.2026

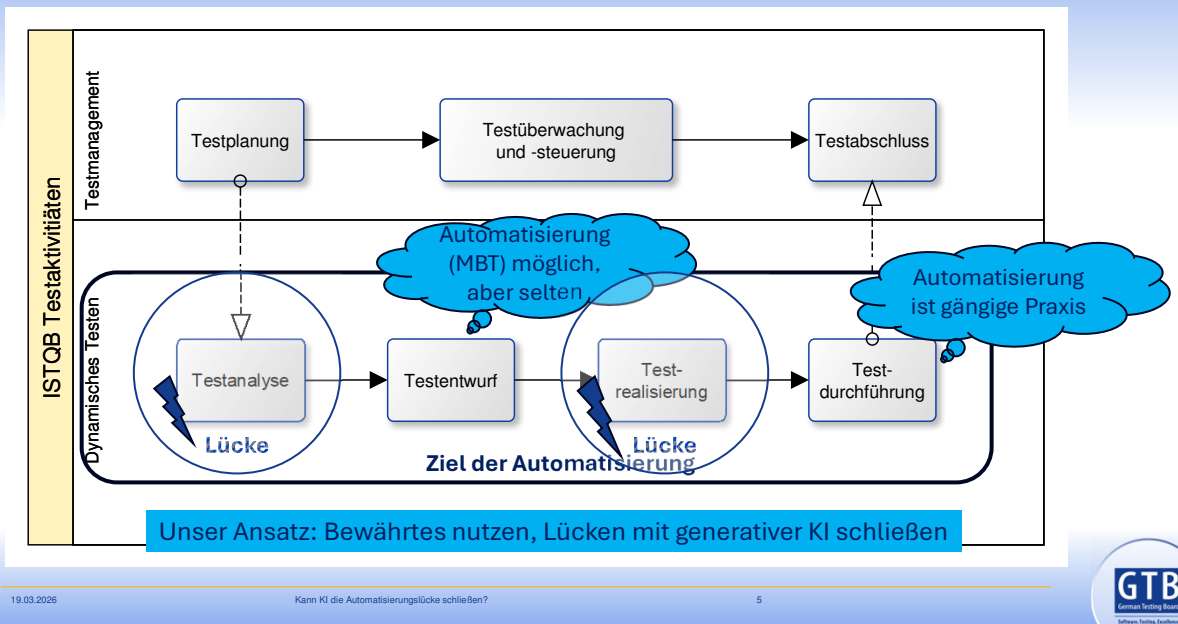
Kann KI die Automatisierungslücke schließen?

4



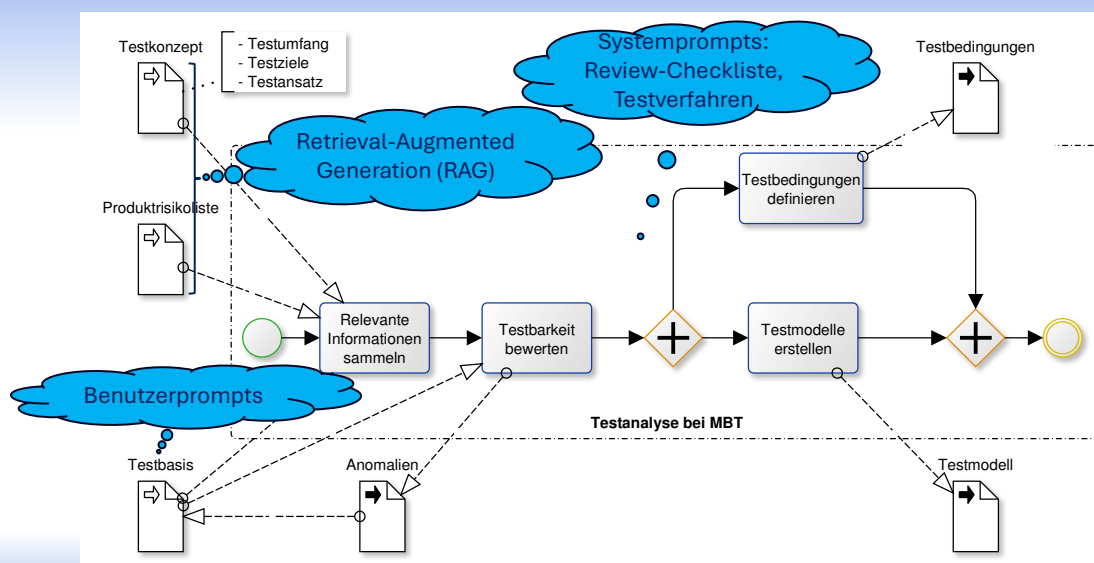
4

Aktivitäten des dynamischen Testens nach ISTQB



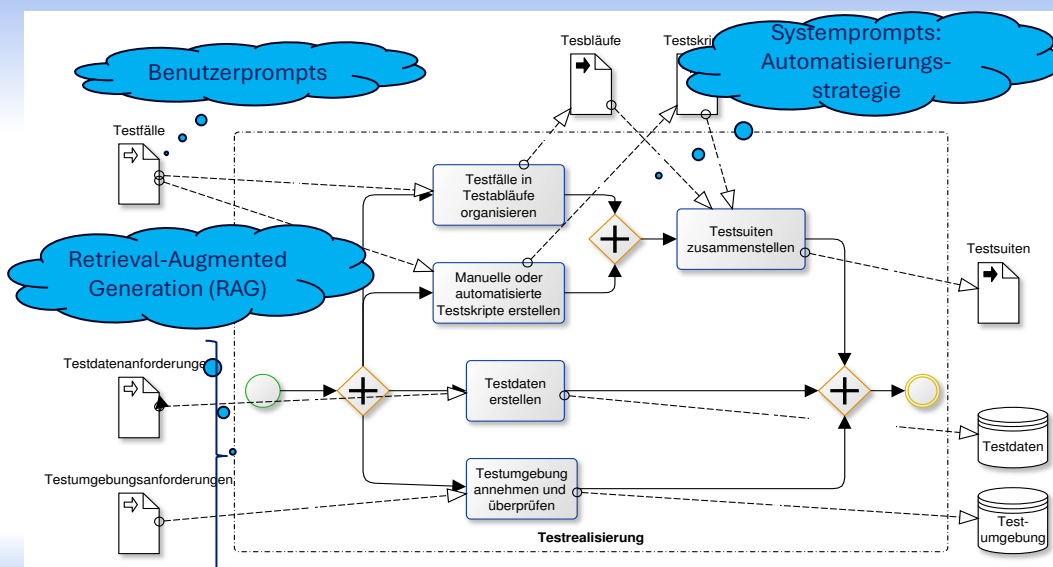
5

Grundlegende Aktivitäten der Testanalyse und ihre KI-Unterstützung



6

Grundlegende Aktivitäten der Testrealisierung und ihre KI-Unterstützung



Quelle: M. Hamburg, A. Roman: Practical Software Test Analysis. Springer Nature, 2026

19.03.2026

Kann KI die Automatisierungslücke schließen?

7



7

Agenda

Automatisierungslücken im Testprozess und ein KI-Lösungsansatz

Fallstudie: Testanalyse mit KI

Fallstudie: Testrealisierung mit KI

Zusammenfassung der Erfahrungen

19.03.2026

Kann KI die Automatisierungslücke schließen?

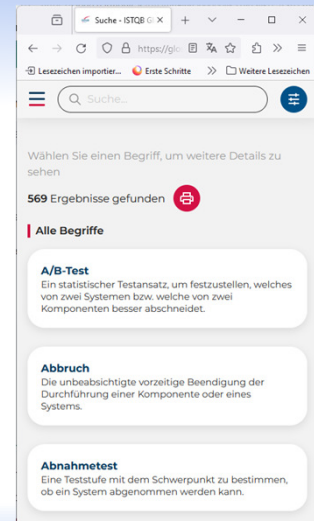
8



8

Ein Pilotprojekt: ISTQB® Glossary App

- Funktionalität: Suche und Anzeige von Inhalten des ISTQB® Glossars
 - Webbasierte Dialoganwendung
 - Inkrementelle Entwicklung
- Automatisierung der Abnahmetests
 - Testvorgehen: KI-gestützte Testgenerierung aus Anforderungen
 - MBT-Werkzeug: Harmony mit KI
 - GenAI-Modell: Anthropic Claude Sonnet 4.6
 - Testausführung: Playwright



19.03.2026

Kann KI die Automatisierungslücke schließen?

9



9

Testkonzept

- Testaufgabe: Abnahmetest im Wartungsrelease
 - Verwendete Frameworks werden aktualisiert
 - Keine neuen Leistungsmerkmale (Features)
- Testziel
 - Zuverlässigkeit der Anwendung sicher stellen
- Testbasis
 - Funktionale Anforderungen
- Testansatz
 - Umfassender Regressionstest aller Funktionalitäten
 - So weit wie möglich automatisiert
 - Testorakel: Produktivsystem

19.03.2026

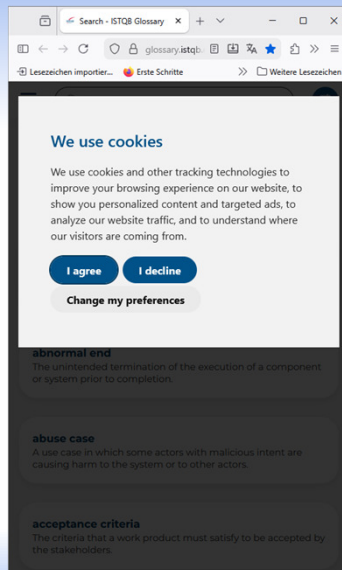
Kann KI die Automatisierungslücke schließen?

10



10

Funktionalität der Glossarsuche am Frontend – ein Szenario (1/6)



19.03.2026

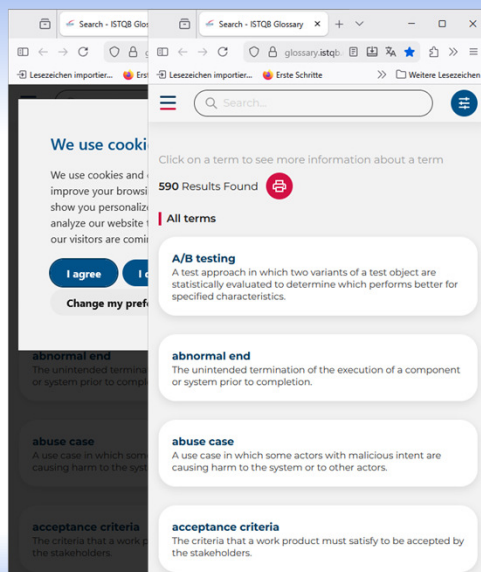
Kann KI die Automatisierungslücke schließen?

11



11

Funktionalität der Glossarsuche am Frontend – ein Szenario (2/6)



19.03.2026

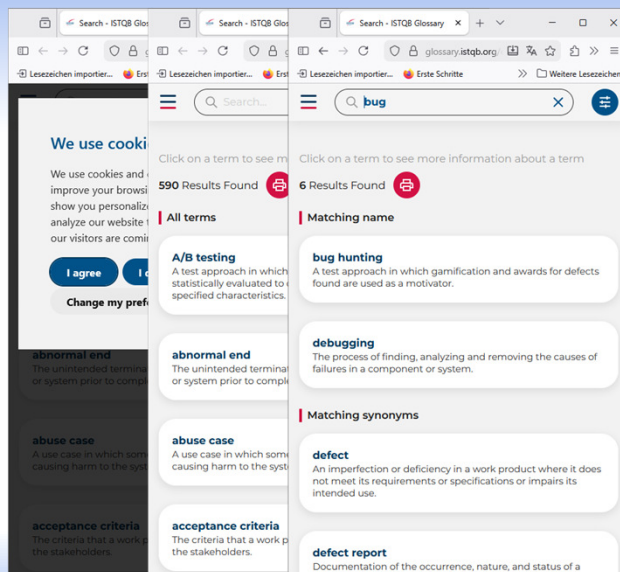
Kann KI die Automatisierungslücke schließen?

12



12

Funktionalität der Glossarsuche am Frontend – ein Szenario (3/6)



19.03.2026

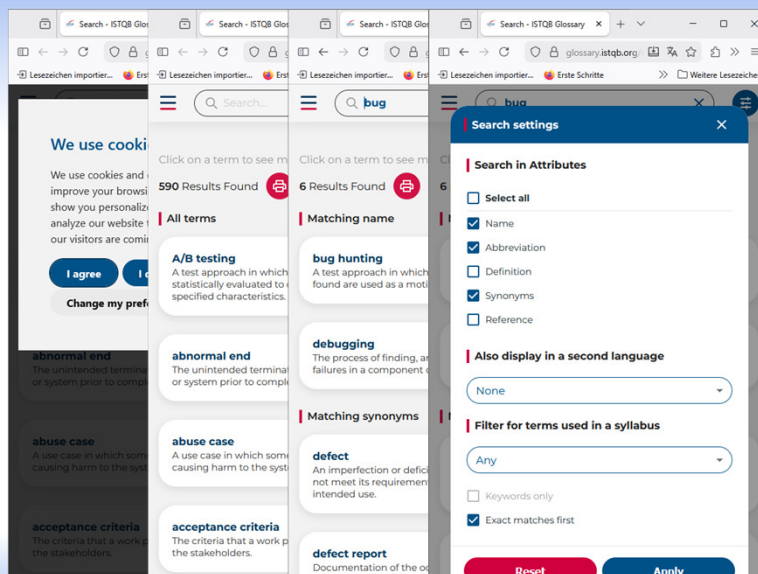
Kann KI die Automatisierungslücke schließen?

13



13

Funktionalität der Glossarsuche am Frontend – ein Szenario (4/6)



19.03.2026

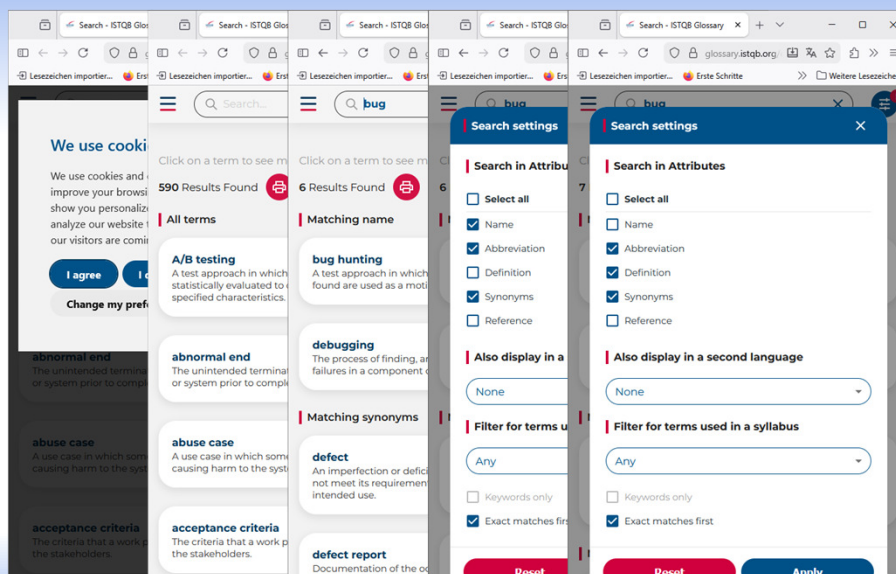
Kann KI die Automatisierungslücke schließen?

14



14

Funktionalität der Glossarsuche am Frontend – ein Szenario (5/6)



19.03.2026

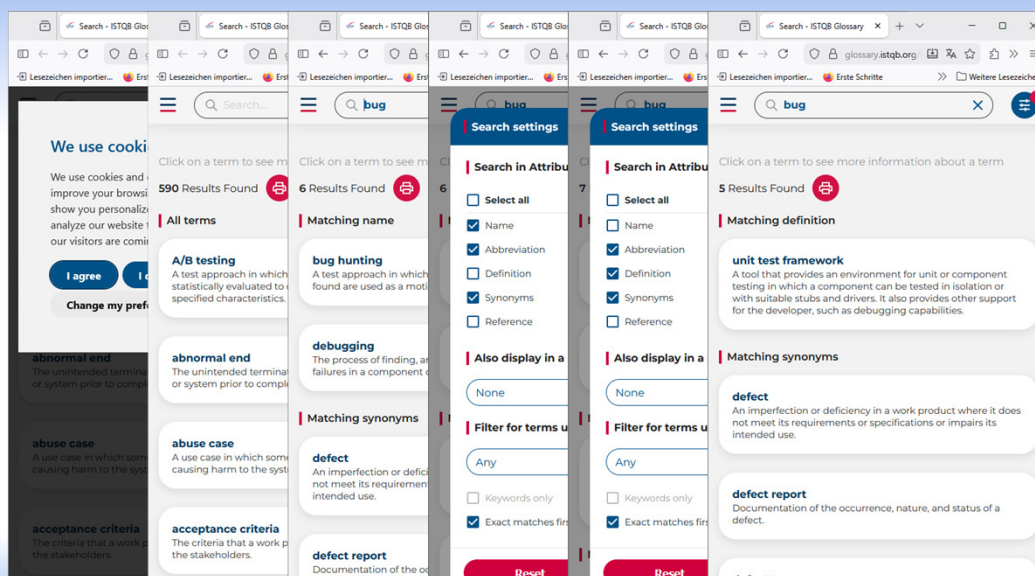
Kann KI die Automatisierungslücke schließen?

15



15

Funktionalität der Glossarsuche am Frontend – ein Szenario (6/6)



19.03.2026

Kann KI die Automatisierungslücke schließen?

16



16

Anforderungsspezifikation zur Suche in Attributen

- 3.2.1.1 In the search settings, group 'Search by attributes', the user can select the attributes name, abbreviation, definition, synonym, reference in any combination.
- 3.2.1.2 In the default search settings, the attributes "Name", "Abbreviation" and "Synonym" are selected, the others are not.
- 3.0.3 At the beginning of a browser session, the app uses the default search settings.
- 3.2.1.3 The user can select or deselect all attributes with a toggle checkbox 'Select all'.
- 3.2.1.4 Given a search text is entered, when the user edits the search settings and saves them with the button 'Apply', the app will search for the search text in the attributes selected.
- 3.2.1.6 When a search text is entered, the fuzzy search behavior and the structure of the search result list are the same as described in feature 3.1 (Text search with default settings).
- 3.2.1.7 Given an empty search string, the search result is the same for any choice of attributes in the search settings.
- 3.0.6 The user can end editing the search settings and return to the search by selecting one of three options:
- Cancel - the search settings will not be changed. The search results remain unchanged.
- Apply - the search settings will be changed and applied. The search is re-executed with the new settings and the new search results are displayed.
- Reset - the search settings will be changed to the default values. If needed, the search is re-executed with the default settings and the new search results displayed.

(Die im Szenario überdeckten Anforderungen sind hervorgehoben)

19.03.2026

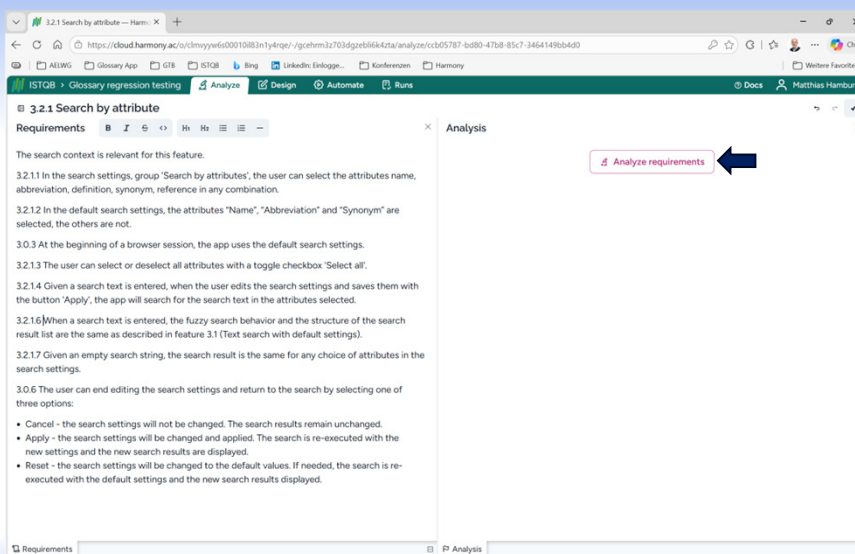
Kann KI die Automatisierungslücke schließen?

17



17

KI soll die Testbarkeit bewerten und begründen (1/3)



19.03.2026

Kann KI die Automatisierungslücke schließen?

18



18

KI soll die Testbarkeit bewerten und begründen (2/3)

The screenshot displays the '3.2.1 Search by attribute' requirements in the ISTQB Glossary regression testing tool. The 'Analysis' panel on the right lists three major anomalies:

- Select all toggle behavior (major)**: The behavior of the 'Select all' toggle is ambiguous. When some attributes are already selected and the user clicks 'Select all', it's unclear whether it should select all remaining attributes or deselect all currently selected ones. This could lead to different implementations with opposite behaviors. *Suggestion: Clarify the 'Select all' toggle behavior: 'When clicked, the toggle selects all attributes if any are currently unselected, or deselects all attributes if all are currently selected.'*
- Reset button search execution condition (major)**: The condition 'if needed' for re-executing the search when Reset is clicked is ambiguous. It's unclear when the search re-execution is needed versus not needed. This could result in inconsistent behavior where sometimes Reset triggers a search and sometimes it doesn't. *Suggestion: Specify the Reset behavior clearly: 'Reset - the search settings will be changed to the default values. The search is re-executed with the default settings and the new search results are displayed.'*
- No attributes selected scenario (major)**: There is no specification of what happens when no attributes are selected. Can users deselect all attributes, and if so, what happens when they attempt to search with no attributes selected? This could lead to different implementations - some might prevent deselecting all attributes, others might allow it but handle the search differently. *Suggestion: Add requirement: 'At least one attribute must always remain selected. If the user attempts to deselect the last remaining attribute, the action is prevented and the attribute remains selected.'*

Below these, there is a note about an empty search attribute independence (major) which is counterintuitive and potentially contradictory.

19.03.2026

Kann KI die Automatisierungslücke schließen?

19



19

KI soll die Testbarkeit bewerten und begründen (3/3)

The screenshot displays the '3.2.1 Search by attribute' requirements in the ISTQB Glossary regression testing tool. The 'Analysis' panel on the right shows the status 'No anomalies left' and a 'Continue to Design' button. A blue arrow points to the button, indicating the next step in the process.

19.03.2026

Kann KI die Automatisierungslücke schließen?

20



20

KI soll relevante Informationen sammeln

The screenshot displays the Hammy AI interface with two main panels. The left panel, titled '3.2.1 Search by attribute', lists requirements for a search feature. The right panel, titled 'Test design', shows a state transition diagram for the same feature. Two blue callout bubbles highlight specific elements:

- Vorbedingung aus vorangehenden Features nutzen** (Use pre-condition from previous features): Points to the 'COOKIES ACCEPTED' state at the top of the test design diagram.
- Informationen aus dem Kontext nutzen (RAG)** (Use information from the context (RAG)): Points to the '3.2.1.2 In the default search settings, Name, Abbreviation and Synonym are selected' requirement and its corresponding state in the diagram.

19.03.2026

Kann KI die Automatisierungslücke schließen?

21



21

KI soll Testverfahren anwenden (1)

The screenshot displays the Hammy AI interface with two main panels. The left panel, titled '3.2.1 Search by attribute', lists requirements for a search feature. The right panel, titled 'Test design', shows a state transition diagram for the same feature. A blue callout bubble highlights a specific element:

- Anwendbare Testverfahren erkennen und begründen** (Recognize and justify applicable test procedures): Points to the 'ACTION-STATE TEST DESIGN DECISION' section in the test design panel, which lists actions like 'click Apply button', 'click Cancel button', 'click Reset button', and 'click Select all toggle'.

19.03.2026

Kann KI die Automatisierungslücke schließen?

22



22

KI soll Testverfahren anwenden (2)

3.2.1 Search by attribute

Test design

COOKIES ACCEPTED

3.0.3 At the beginning of a browser session, the app uses the default search settings

3.2.1.2 In the default search settings, Name, Abbreviation and Synonym are selected

do nothing #1

app starts with default search settings

search settings dialog shows Name, Abbreviation, Synonym selected

3.2.1.1 User can select attributes in any combination

open search settings dialog

search settings dialog opens

Name, Abbreviation, Synonym selected; Definition, Reference not selected

deselect Name attribute

Name attribute deselected

select Definition attribute

Definition attribute selected

3.2.1.4 When user edits settings and saves with Apply, app searches in selected attributes

enter search text 'string occurring 1-10'

search text entered in field

click Apply button

search settings applied

search executed with modified settings

search results displayed for Abbreviation, Synonym, Definition attributes only

3.2.1.3 Select all toggle button selects all if any unselected, deselects all if all selected

open search settings dialog

search settings dialog opens

Name, Abbreviation, Synonym selected; Definition, Reference not selected

Phase A: Same-state transitions - modify individual attributes

deselect Abbreviation attribute

Abbreviation attribute deselected

Phase B: Cross-state transition - click Select all toggle with some unselected

click Select all toggle button

all attributes selected (Name, Abbreviation, Synonym, Definition, Reference)

From SEARCH SETTINGS ALL SELECTED, test the multi-outcome action again

cloud.harmony.acrol/_cc05787-bd80-47b8-85c7-3464149b4d0

Tests (15) Statechart

Anforderungen
überdecken und
Verfolgbarkeit
sicher stellen

23

KI soll Testverfahren anwenden (3)

3.2.1 Search by attribute

Test design

COOKIES ACCEPTED

Phase B: Cross-state transition - select individual attribute

select Name attribute

Name attribute selected

Testing state SEARCH SETTINGS MODIFIED - covering remaining transitions

open search settings dialog

search settings dialog opens

select Definition attribute

Definition attribute selected

From SEARCH SETTINGS MODIFIED, enumerate all possible actions:

Missing(SEARCH SETTINGS MODIFIED) = ((modify more attributes, SEARCH SETTINGS MODIFIED), (click Apply, SEARCH RESULTS WITH NEW SETTINGS), (click Cancel, SEARCH RESULTS UNCHANGED), (click Reset, SEARCH SETTINGS DEFAULT), (click Select all toggle, SEARCH SETTINGS ALL SELECTED or SEARCH SETTINGS ALL Deselected))

Phase A: Same-state transitions - modify more attributes

deselect Abbreviation attribute

Abbreviation attribute deselected

select Reference attribute

Reference attribute selected

Phase B: Cross-state transition - click Cancel

click Cancel button

search settings dialog closed

all modifications reverted

settings back to default

Extreme Value Testing - testing with only one attribute selected

open search settings dialog

search settings dialog opens

deselect Name and Abbreviation attributes

Name and Abbreviation deselected

deselect Synonym attribute

Synonym deselected

select only Reference attribute

only Reference attribute selected

enter search text 'string occurring 1-10'

search text entered in field

cloud.harmony.acrol/_cc05787-bd80-47b8-85c7-3464149b4d0

Requirements Test design Sketch

Tests (15) Statechart

Testverfahren
anwenden und
Testfälle begründen

24

KI soll Testmodelle erstellen, aus denen MBT abstrakte Testfälle generieren kann

ISTQB > Glossary regression testing > Analyze > Design > Automate > Runs

3.2.1 Search by attribute

Tests (15)

3SBA1 3SBA2 3SBA3 3SBA4 3SBA5 3SBA6 3SBA7 3SBA8 3SBA9 3SBA10 3SBA11 3SBA12 3SBA13 3SBA14 3SBA15

3SBA1 - 3.0.3 At the beginning of a browser session, the app uses the default search settings - 3.2.1.2 In the default search settings, Name, Abbreviation and Synonym are selected - 3.2.1.1 User can select attributes in any combination - 3.2.1.4 When user edits settings and saves with Apply, app searches in selected attributes

COOKIES ACCEPTED →

3.0.3 At the beginning of a browser session, the app uses the default search settings

3.2.1.2 In the default search settings, Name, Abbreviation and Synonym are selected

- do nothing #1 → app starts with default search settings → search settings dialog shows Name, Abbreviation, Synonym selected

3.2.1.1 User can select attributes in any combination

- open search settings dialog → search settings dialog opens → Name, Abbreviation, Synonym selected; Definition, Reference not selected
- deselect Name attribute → Name attribute deselected
- select Definition attribute → Definition attribute selected

3.2.1.4 When user edits settings and saves with Apply, app searches in selected attributes

- enter search text 'string occurring 1-10' → search text entered in field
- click Apply button → search settings applied → search executed with modified settings → search results displayed for Abbreviation, Synonym, Definition attributes only

Requirements Test design Sketch Tests (15) Statechart

19.03.2026

Kann KI die Automatisierungslücke schließen?

25



25

Agenda

Automatisierungslücken im Testprozess und ein KI-Lösungsansatz

Fallstudie: Testanalyse mit KI

Fallstudie: Testrealisierung mit KI

Zusammenfassung der Erfahrungen

19.03.2026

Kann KI die Automatisierungslücke schließen?

26



26

KI soll Testfälle in Testabläufe organisieren

The screenshot displays the Harmony tool's 'Test design' interface. On the left, a sidebar lists features like '1 Start', '2 Languages', '3 Search', and '3.1 Text search with default options'. The main area shows a 'Test design' flowchart with steps like '2.2.1 Cookies may be used if the user agrees' and '2.2.2 When a user starts a session with the generic URL'. A 'Tests (4)' panel on the right lists test cases, including '1S1 - 2.2.1 Cookies may be used if the user agrees' and '2.2.1 Cookies may be used if the user agrees'. A blue callout bubble points to the 'Tests (4)' panel with the text: 'Die KI erstellt feature-übergreifende Zustände (Published), die sie später als Vorbedingung nutzen kann'.

19.03.2026

Kann KI die Automatisierungslücke schließen?

27



27

KI soll Testskripte auf der Ebene der Testschnittstelle generieren

- Das MBT-Werkzeug sendet die abstrakten Testschritte mit Schlüsselwörtern auf Anwendungsebene an die KI
- Die KI sucht die entsprechenden Aktionen und Verifikationen an der Testschnittstelle sowie passende Testdaten und generiert die Automatisierungsskripte
 - Harmony nutzt die Aria Snapshots von Playwright
- Das Testautomatisierungswerkzeug (Playwright) führt die Testfälle aus

19.03.2026

Kann KI die Automatisierungslücke schließen?

28



28

Codeless Testskriptgenerierung mit Harmony (1/2)

The screenshot shows the Harmony IDE interface. The top bar includes tabs for 'Test design', 'Definitions', 'Files', and 'Tests (15)'. The main area displays a test design for '3.2.1 Search by attribute'. The test design is a flowchart with steps like 'COOKIES ACCEPTED', '3.0.3 At the beginning of a browser session...', '3.2.1.1 User can select attributes in any combination', and '3.2.1.4 When user edits settings and saves with Apply...'. A blue arrow points to the 'Run JSBA1' action, which is labeled 'Action (not automated)'. The right side of the interface shows a browser window with the URL 'https://cloud.harmony.ac/...' and the text 'No browser pages open.'

19.03.2026

Kann KI die Automatisierungslücke schließen?

29

29

Codeless Testskriptgenerierung mit Harmony (2/2)

The screenshot shows the Harmony IDE interface. The top bar includes tabs for 'Test design', 'Definitions', 'Files', and 'Tests (15)'. The main area displays a test design for '3.2.1 Search by attribute'. The test design is a flowchart with steps like 'COOKIES ACCEPTED', '3.0.3 At the beginning of a browser session...', '3.2.1.1 User can select attributes in any combination', and '3.2.1.4 When user edits settings and saves with Apply...'. A blue arrow points to the 'Run JSBA1' action, which is labeled 'Action (not automated)'. The right side of the interface shows a browser window with the URL 'https://glossary-test.istqb.org/' and a search results page. The browser window shows a search bar with the text '3.2.1 Search by attribute' and a list of search results.

19.03.2026

Kann KI die Automatisierungslücke schließen?

30

30

Agenda

Automatisierungslücken im Testprozess und ein KI-Lösungsansatz

Fallstudie: Testanalyse mit KI

Fallstudie: Testrealisierung mit KI

Zusammenfassung der Erfahrungen

19.03.2026

Kann KI die Automatisierungslücke schließen?

31



31

Es ist nicht egal, welches LLM verwendet wird

- Nur Anthropic Claude Sonnet war für die Anforderungen von Harmony geeignet
- Die anderen versagen bis heute an jeweils verschiedenen Stellen
 - Blieben hängen
 - Konnten die Anforderungen nicht reviewen
 - Halluzinierten
 - Haben viel zu viele Testfälle generiert
 - Konnten Testverfahren nicht anlernen
 - Hatten kein verständliches Reasoning
- Copilot war beim Unterstützen des Prompts sehr hilfreich

GenAI-Modelle (KI-generierte Auswahl am 18.03.2026)

Führende GenAI-Modelle & Tools für Code

- GitHub Copilot (basiert auf OpenAI Codex/GPT-4o):** Ein weit verbreitetes Tool zur Code-Vervollständigung und -Generierung. [TheBlue.ai +3](#)
- OpenAI GPT-4o / GPT-4-turbo:** Leistungsstarke LLMs, die für komplexe logische Aufgaben, Code-Refactoring und das Erklären von Code verwendet werden. [Golem Karrierewelt +2](#)
- Anthropic Claude 3.5 Sonnet / 4.5 Sonnet:** Gilt aktuell als eines der führenden Modelle, oft bevorzugt für hohe Genauigkeit bei komplexen Programmieraufgaben. [Golem Karrierewelt +1](#)
- Google Gemini 1.5 Pro/Flash:** Googles multimodales Modell, das in Tools wie Project IDX oder Android Studio integriert ist, um Code zu generieren und zu analysieren. [Golem Karrierewelt +2](#)
- Amazon Q Developer (ehemals CodeWhisperer):** Ein von AWS entwickelter KI-Assistent, der speziell für Cloud-Anwendungen und Java/Python-Entwicklung optimiert ist. [TheBlue.ai +1](#)
- Tabnine:** Ein auf Datenschutz fokussierter KI-Code-Assistent, der lokal oder in privaten Clouds ausgeführt werden kann. [Heise Business Services +4](#)
- Mistral AI (z.B. Codestral):** Ein europäisches Modell, das speziell auf Code-Generierung spezialisiert ist. [Golem Karrierewelt +1](#)
- Ollama:** Ein Framework, um Open-Source LLMs (wie Llama 3 oder Mistral) lokal für die Entwicklung zu nutzen, was besonders datenschutzrelevant ist. [codecentric AG +3](#)



32

Prompten in einem KI-basierten Testwerkzeug ist nichttrivial

- Die marktgängigen LLM's können Testverfahren nicht korrekt anwenden
- Umfangreiche textuelle Systemprompts werden für Testanalyse und Testrealisierung benötigt, z.B. für:
 - Reviewen der Spezifikation
 - Auswahl und Anwendung von Testverfahren
 - Erstellung von Testabläufen
 - Automatisierungsstrategie
- Auswirkung der Beispiele muss beachtet werden

19.03.2026

Kann KI die Automatisierungslücke schließen?

33



33

Testanalysten arbeiten mit KI anders und effizienter, werden aber nicht ersetzt

- Ein geeignetes KI-basiertes Werkzeug macht Testanalysten um ein Vielfaches produktiver
 - Testbasis eines Features in 30 Sekunden gereviewt
 - 8 Testfälle aus einem Feature in weniger als 1 Minute generiert
 - KI-basierte Testautomatisierung läuft schneller als manuelles Skripten
- Testanalysten müssen nicht kodieren oder automatisieren
- Testanalysten müssen dafür prompten können
- Testanalysten werden durch die KI nicht ersetzt
 - Human-in-the-loop unverzichtbar

19.03.2026

Kann KI die Automatisierungslücke schließen?

34



34



Fragen? Gerne!

DR. MATTHIAS HAMBURG

Powered by
German Testing Board e. V.

Vorstand
Florian Fieber

Stellv. Vorstände
Dietrich Leimsner
Dr. Klaudia Dussa-Zieger
Dr. Amin Metzger

Geschäftsführer
Dr. Amin Metzger

Backoffice
Andrea Kränzlein
OFFICE-MANAGEMENT
Pirckheimerweg 4
91058 Erlangen
GERMANY
Tel.: +49 (0) 91 31 / 97 61 06
Mobil: +49 (0) 171 / 63 90 749
E-Mail: backoffice@gtb.de

