



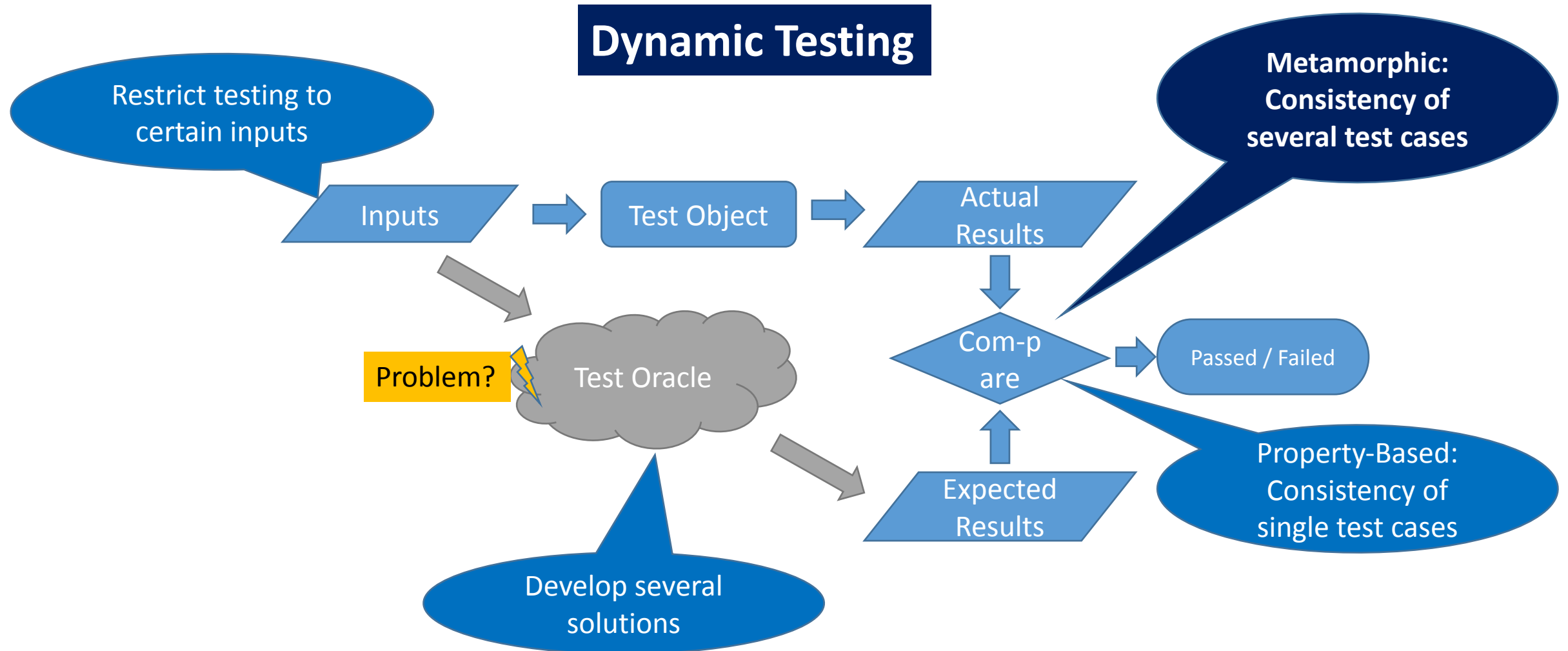
**HUSTEF**  
HUNGARIAN SOFTWARE TESTING FORUM

# Metamorphic Testing

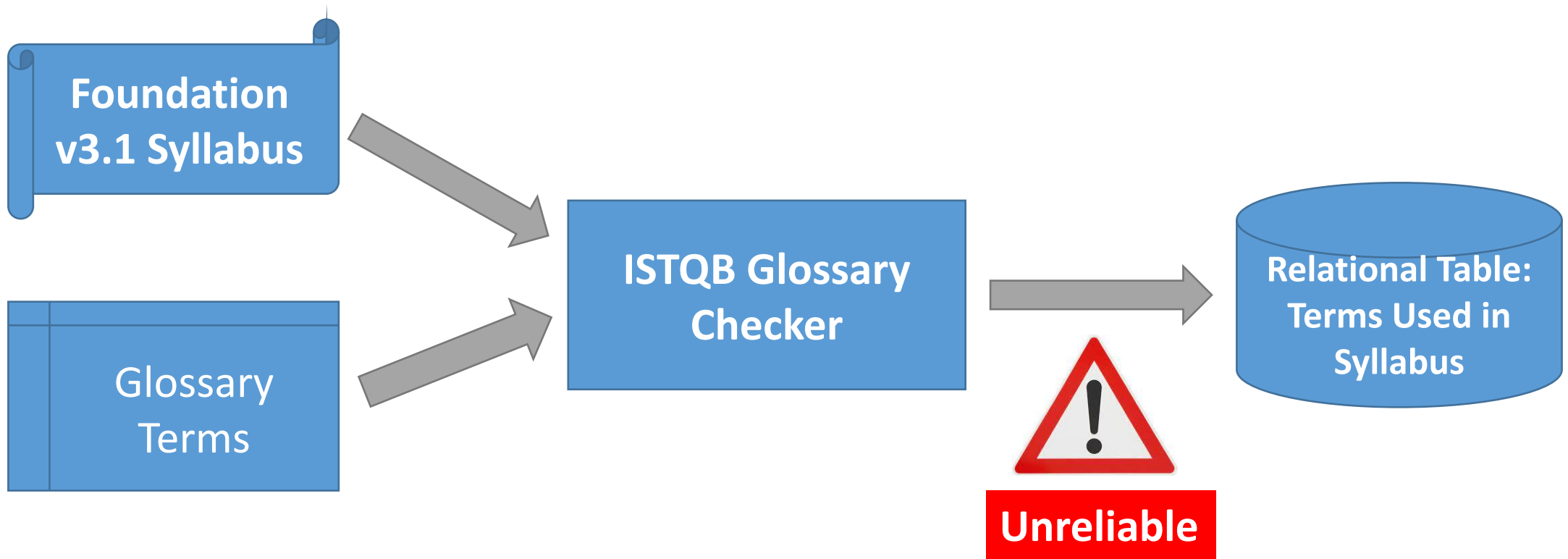
Matthias Hamburg, German Testing Board

October 2022





Classic solutions: See Elaine Weyuker: On Testing Non-testable Programs. The Computer Journal, 1982

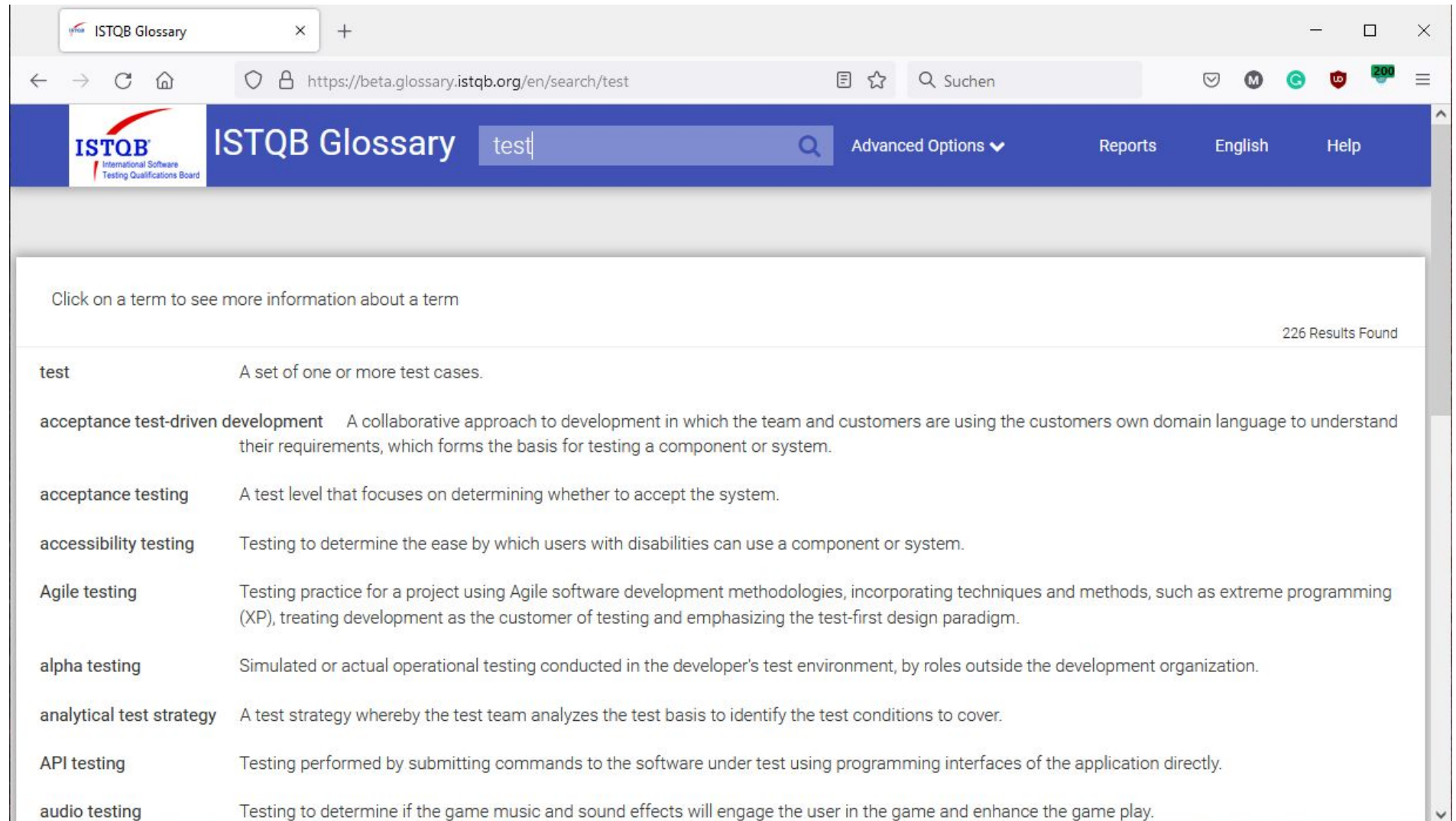


# Source Test Case TC1: Simple Text Search for „test“

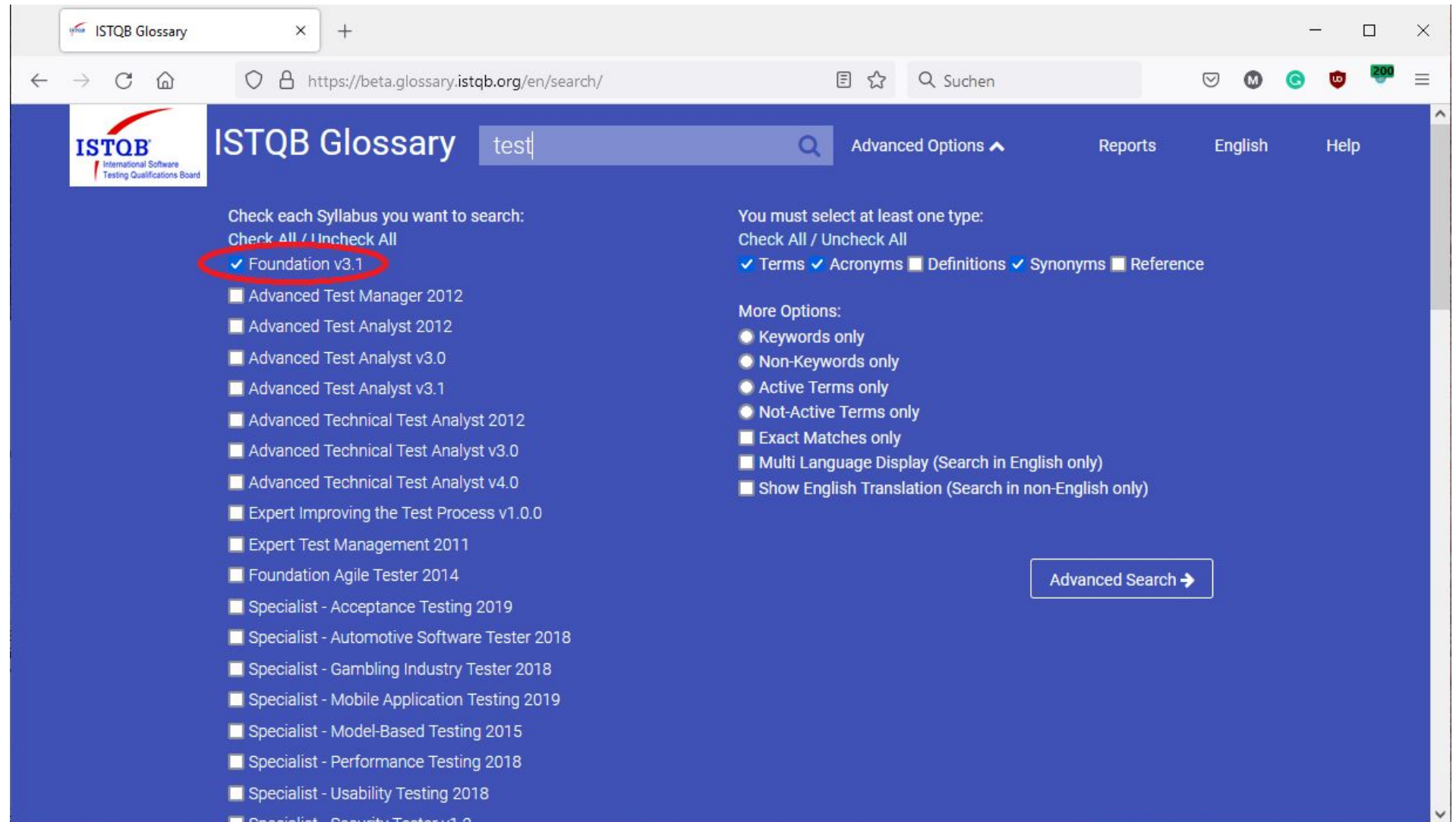
**Input:**  
Simple search for the string „test“ in the entire ISTQB Glossary in English

**Actual Result:**  
List of 226 terms out of 590 in total

We assume it passed



MR1: Adding a filter for terms used in a Syllabus must yield a subset of the original result list.



# Follow-up Test Case TC2

Input:  
Same as for TC1 + filter  
for Foundation v3.1  
Syllabus

Actual result:  
List with 112 terms.

We verify MR1 with a  
comparator

**MR1 Passed.**

The screenshot shows a web browser window with the URL <https://beta.glossary.istqb.org/en/search/test>. The page title is "ISTQB Glossary" and the search term is "test". The page displays a list of 112 search results. The first few results are:

| Term                          | Definition   |
|-------------------------------|--|
| test                          | A set of one or more test cases.   |
| acceptance testing            | A test level that focuses on determining whether to accept the system.   |
| accessibility testing         | Testing to determine the ease by which users with disabilities can use a component or system.  |
| alpha testing                 | Simulated or actual operational testing conducted in the developer's test environment, by roles outside the development organization.  |
| beta testing                  | A type of acceptance testing performed at an external site to the developer's test environment by roles outside the development organization.  |
| black-box test technique      | A test technique based on an analysis of the specification of a component or system.   |
| change-related testing        | A type of testing initiated by modification to a component or system.  |
| checklist-based testing       | An experience-based test technique whereby the experienced tester uses a high-level list of items to be noted, checked, or remembered, or a set of rules or criteria against which a product has to be verified. |
| component integration testing | Testing in which the test items are interfaces and interactions between integrated components.   |

Let us make TC2 to a source test case of a new metamorphic relation.

MR2: Adding the filters for Keywords only/Non-Keywords only must partition the output of the source test case.

The image displays two screenshots of the ISTQB Glossary search interface. Both screenshots show the search term 'test' and the search results filtered by 'Foundation v3.1' (indicated by a red circle). The top screenshot shows the search results filtered by 'Keywords only' (indicated by a red circle). The bottom screenshot shows the search results filtered by 'Non-Keywords only' (indicated by a red circle). Both screenshots also show the search results filtered by 'Terms', 'Acronyms', and 'Synonyms' (indicated by red circles).

# Follow-up Test Cases TC3 and TC4

**Input:**  
Same as for TC2 +  
filters for Keywords  
only / Non-keywords  
only

**Actual results:**  
Keywords only: 64  
Non-keywords only: 42

**We verify MR2:**  
Expected: 112 in total.

**MR2 failed:**  
**6 terms are missing!**

Click on a term to see more information about a term

64 Results Found

|                          |  |
|--------------------------|--|
| acceptance testing       | A test level that focuses on determining whether to accept the system.   |
| alpha testing            | Simulated or actual operational testing conducted in the developer's test environment, by roles outside the development organization.  |
| beta testing             | A type of acceptance testing performed at an external site to the developer's test environment by roles outside the development organization.  |
| black-box test technique | A test technique based on an analysis of the specification of a component or system.   |
| change-related testing   | A type of testing initiated by modification to a component or system.  |
| checklist-based testing  | An experience-based test technique whereby the experienced tester uses a high-level list of items to be noted, checked, or remembered, or a set of rules or criteria against which a product has to be verified. |

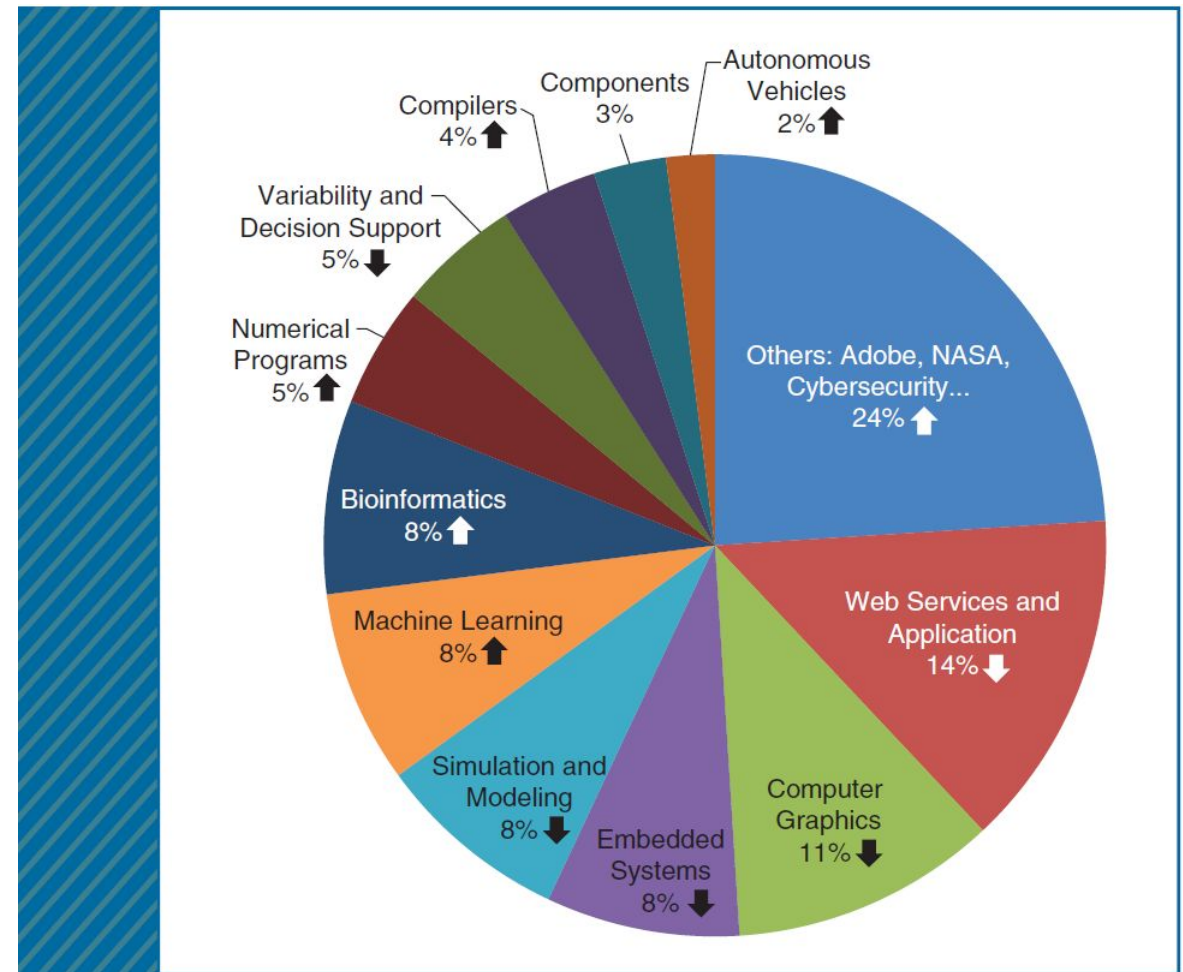
Click on a term to see more information about a term

42 Results Found

|                          |  |
|--------------------------|--|
| test                     | A set of one or more test cases.   |
| accessibility testing    | Testing to determine the ease by which users with disabilities can use a component or system.                          |
| decision testing         | A white-box test technique in which test cases are designed to execute decision outcomes.                              |
| exhaustive testing       | A test approach in which the test suite comprises all combinations of input values and preconditions.                  |
| experience-based testing | Testing based on the tester's experience, knowledge and intuition.   |
| high-level test case     | A test case with abstract preconditions, input data, expected results, postconditions, and actions (where applicable). |

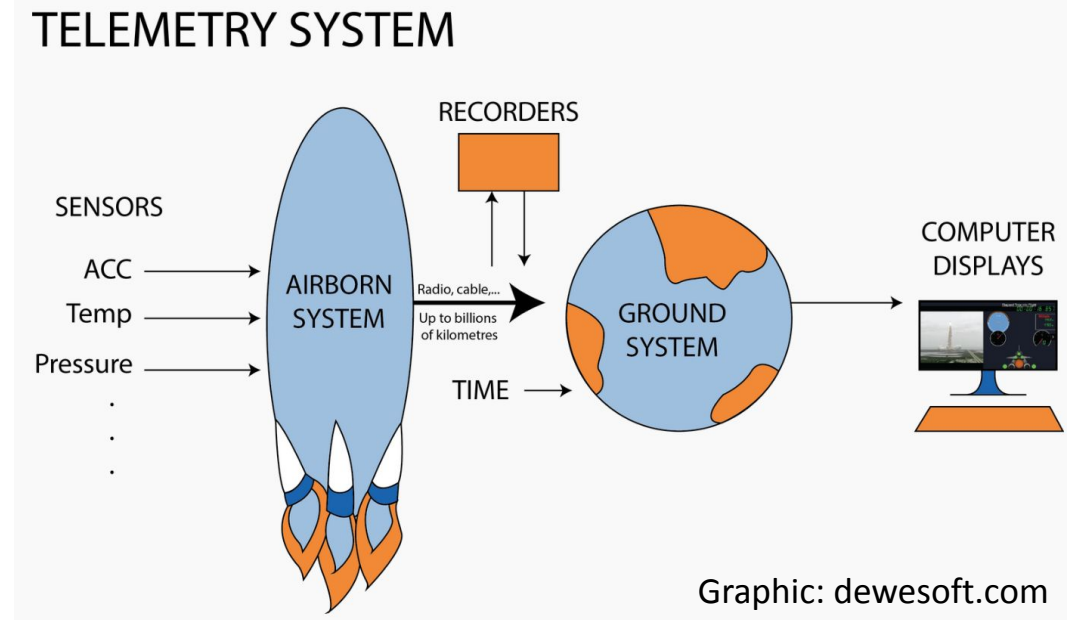


- Introduced in 1998 by Chen et al.
- Survey conducted in 2018: 84 case studies published
- Included in ISO/IEC/IEEE 29119-4:2021 (Test Techniques)
- Not discussed (yet) in ISTQB Advanced Test Analyst

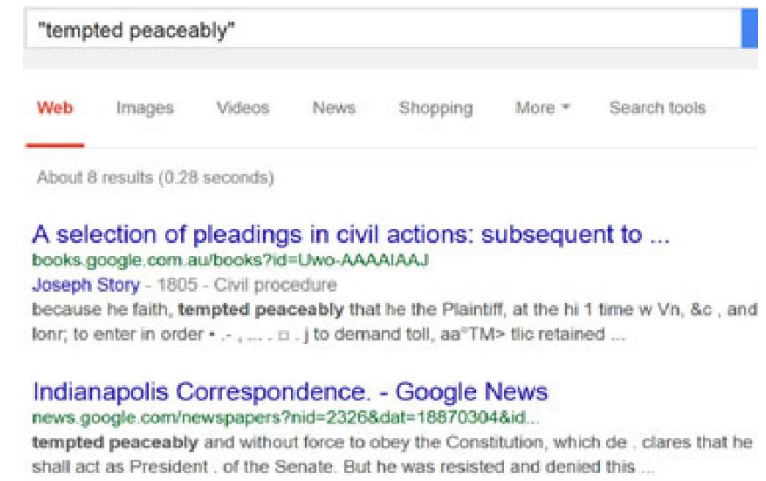


Distribution of 84 case studies by application domains  
Source: Segura et al., 2020

- Test object: NASA Data Access Toolkit
  - High-performance query and research user interface
  - Huge database of telemetry data of the NASA missions
- Oracle problem:
  - Are the query results accurate?
- Metamorphic relations:
  - Equivalent queries must deliver the same outputs
- Test result:
  - A large number of defects detected
  - Successfully combined with MBT and test automation in an agile software lifecycle



- Test objects: The most popular search engines
  - Google, Bing, Chinese Bing and Baidu
- Oracle problem:
  - The database is huge and changes all the time
- Metamorphic relations:
  - Restriction to the url domains in the result list
  - Restriction to the titles in the result list
  - Permutations of conditions combined with „AND“
  - Etc.
- Test result:
  - None of the search engines is 100% reliable
  - In English is Google the most reliable, in Chinese it is Baidu
  - The testers reported several specific defects to the developers



(a)

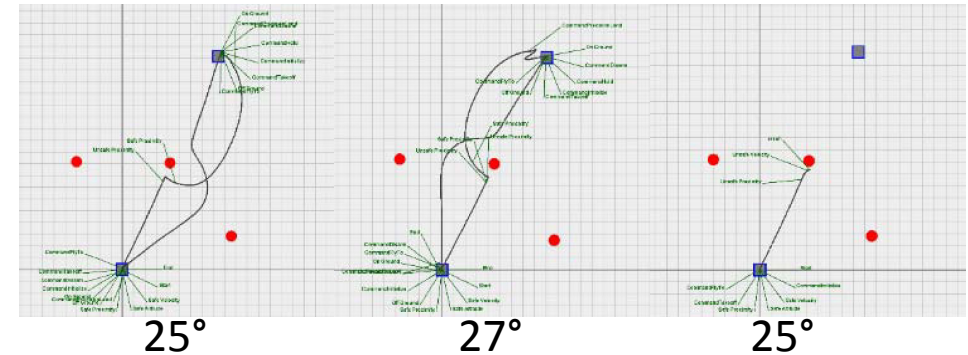


(b)

Aus: Zhou et al. (2016)

- The expected result is often unspecified
  - The behavior cannot be predicted exactly
  - The behavior may change over time without changing the software
- Metamorphic testing is a test technique of choice
  - Dealt with by the new ISTQB® Syllabus „Certified Tester AI Testing“
  - Included in the standard ISO/IEC TR 29119-11:2020 (Testing of AI-based Systems)

- Test object: Drone AI controller
  - Tested in a simulated environment
- Oracle problem:
  - Does the autonomous drone behave safely and as expected?
- Metamorphic relations:
  - Variations of the exposure to light; Rotation; Translation; Changing the distance of the obstacles; Geometrically equivalent obstacle groups
- Test result:
  - Failures in certain situations: Collision with obstacles, failing to recognize the landing pad, unstable behavior, etc.



Follow-up test cases rotating the scene by three similar angles. Source: Lindvall et al.

- Test objects: Autonomous vehicle control systems
  - Three top-performing deep neural networks of the Udacity self-driving car challenge

- Oracle problem:

- What is the correct driving behavior in a specific situation?

- Metamorphic relations:

- Same driving task under various environmental and camera conditions: rain, fog, blurring, picture scale, contrast, brightness, rotation, translation, shear, etc.

- Test result:

- Hundreds of failures that might lead to severe or even fatal accidents
  - Neuron coverage is relevant for effective defect detection

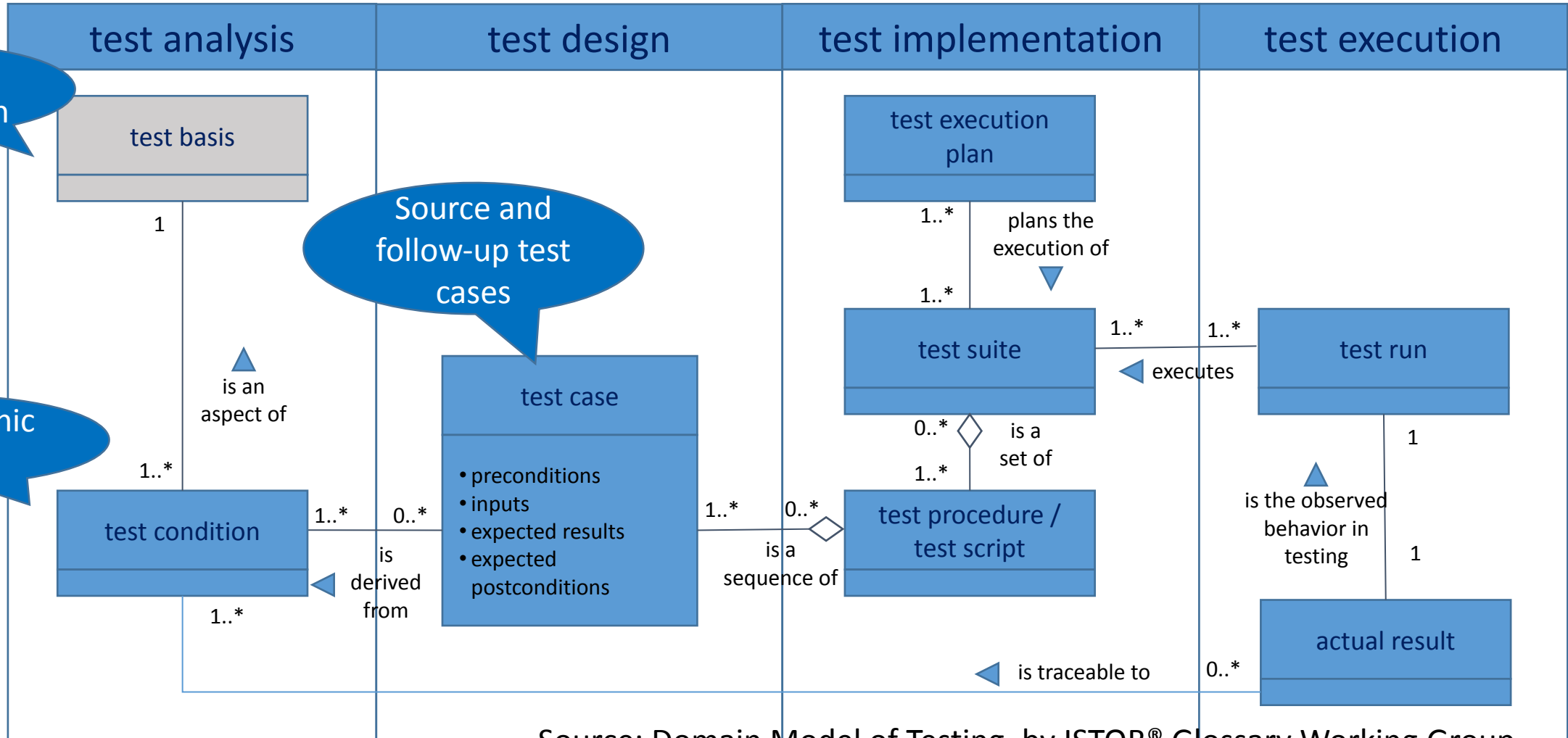


Source test case

Rain added

Source: [deeplearningtest.github.io/deepTest/](https://deeplearningtest.github.io/deepTest/)

# Metamorphic Testing Is a Test Technique



Source: Domain Model of Testing, by ISTQB® Glossary Working Group

- ... determines the effectiveness of defect detection
- ... is an essential creative activity
  - Knowledge of the application domain
  - Functional decomposition
  - Formal specification of metamorphic relations
  - Highest possible code coverage
- ... has two different approaches
  - Input-oriented: Which change of the input leads to a verifiable change of the output?
  - Output-oriented: When are the expected outputs identical / equivalent / partitioning etc.?





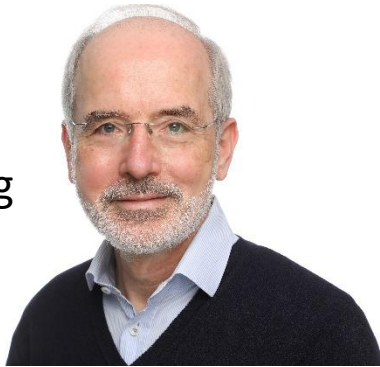
- ISO/IEC/IEEE 29119-4 (2021):
  - “There is currently no industry agreed approach to calculating coverage for metamorphic testing.”
- We need alternatives to the coverage-based end criteria
  - Option 1: Defect detection rate  
(e.g. number of failures per 100 follow-up test cases)
  - Option 2: White-box coverage  
(e.g. neuron coverage)

- A relatively new test technique
- Especially suitable for test objects with a relevant oracle problem
  - Web services, artificial intelligence
- Integrates well with MBT, Agile and DevOps
- The defect detection effectiveness is empirically proven
- The efficiency depends on the degree of automation

**Does metamorphic testing deserve to be included  
in the canon of test techniques?**

Thank you very much for your attention!

Matthias Hamburg



**HUSTEF**  
HUNGARIAN SOFTWARE TESTING FORUM

**German Testing Board**

**President**

Florian Fieber

**Vice-Presidents**

Horst Pohlmann

Dietrich Leimsner

**Executive Director**

Dr. Armin Metzger

**Back Office**

Andrea Kränzlein

OFFICE-MANAGEMENT

Koldestrasse 8 b

91052 Erlangen

GERMANY

Tel.: +49 (0) 91 31 / 97 61 06

Mobile: +49 (0) 171 / 63 90 749

E-Mail: [backoffice@gtb.de](mailto:backoffice@gtb.de)

