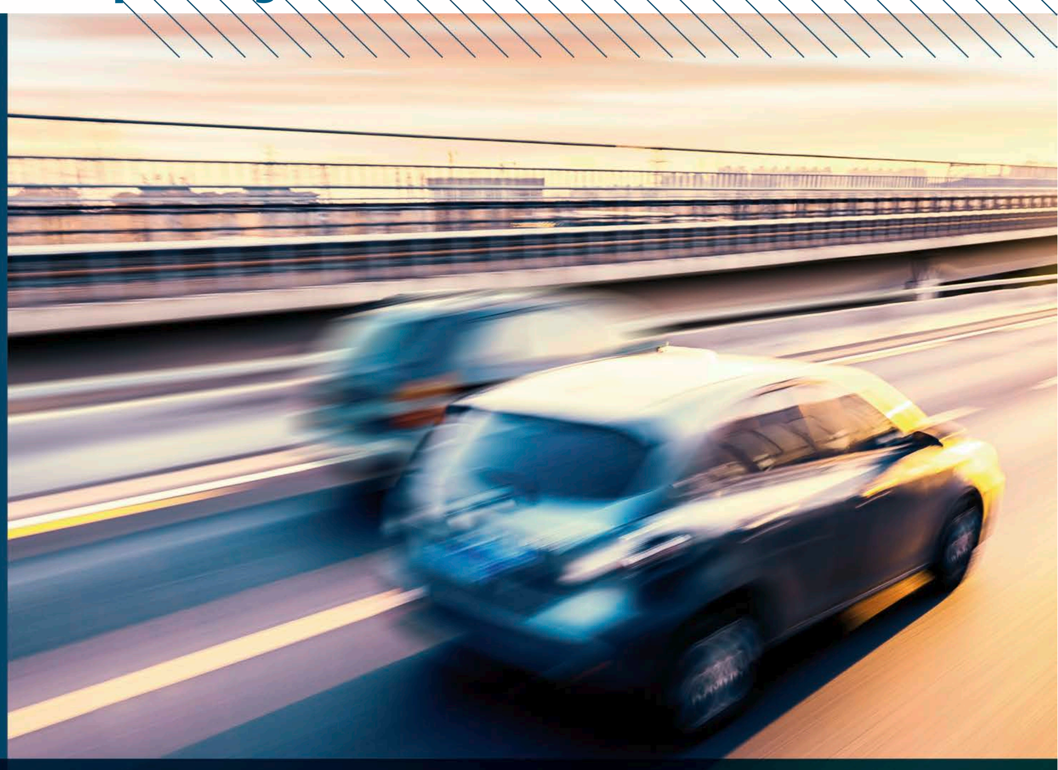
Requirements & Conditions – Stand 8

# CHALLENGES OF A SCENARIO-BASED APPROACH



PEGASUS

## From the testing methodology towards a safety case.

How credible are the evidences provided by scenario-based approaches, Do these evidences support the safety argumentation?

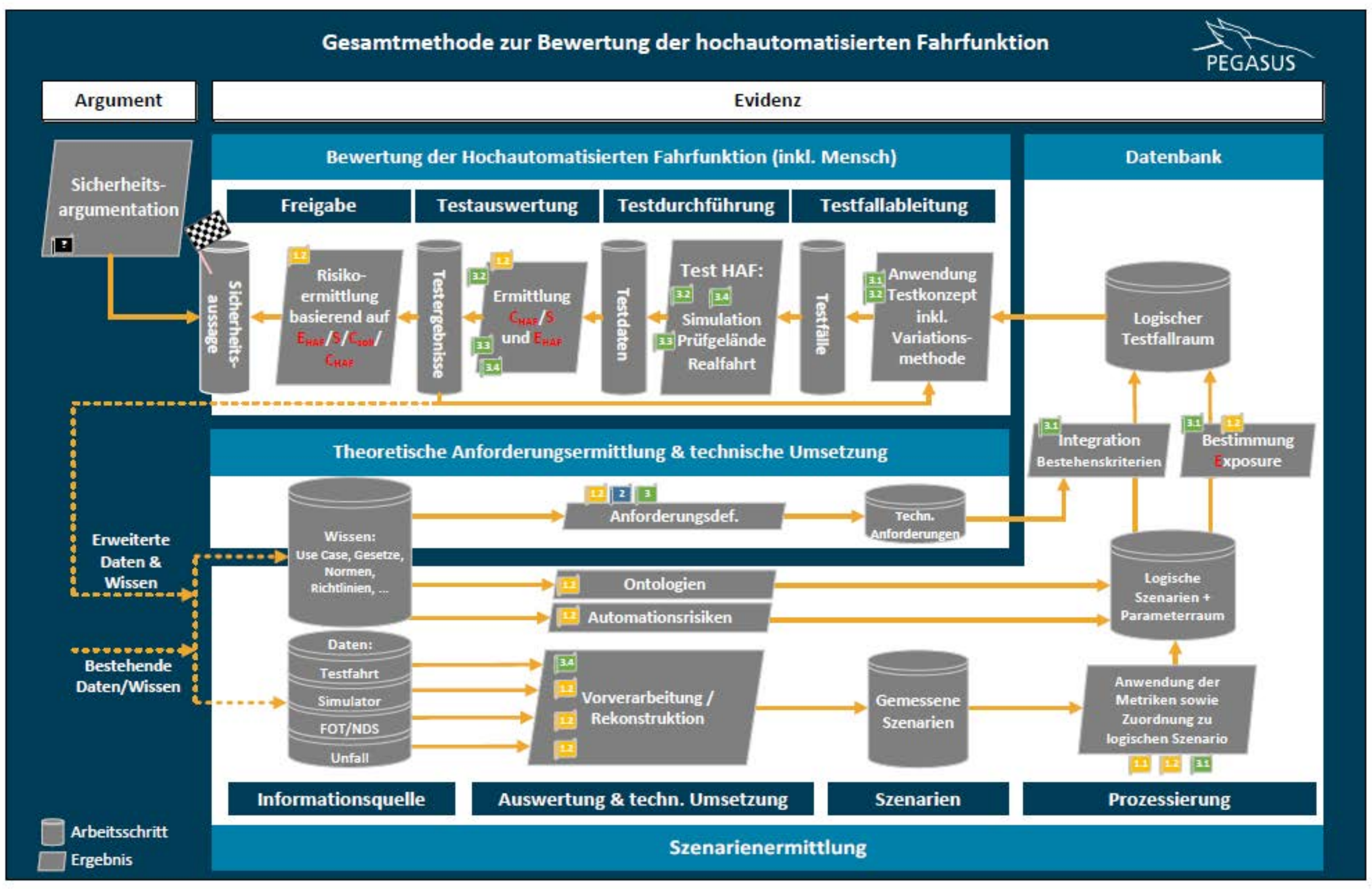
Pegasus provides a grey box testing method by combining obtained scenes with system knowledge in order to provide evidence for the hypothesis that HAF is at least as good as a human driver. To substantiate this, Pegasus provides a method for deriving, conducting and evaluating test cases. The extent to which safety can be demonstrated by using this methodology is scrutinized by means of the following **assessment criteria:**

1. **Assumption coverage** - states the sufficiency of assumptions w.r.t. the design intent
2. **Unfounded evidence** - refers to arguments without supporting evidence
3. **Unused evidence** - refers to evidences not used for the safety case

**II. Apply assessment criteria**

| | Considered Elements |
|---|---|
| 1 | Information base |
| 2 | Ontology |
| 3 | Criticality |
| 4 | Association |
| 5 | Relation |
| 6 | Variation |

**III. Indicate creditability of Pegasus test methodology**




**I. Identify elements with a potential for affecting the provided evidence**

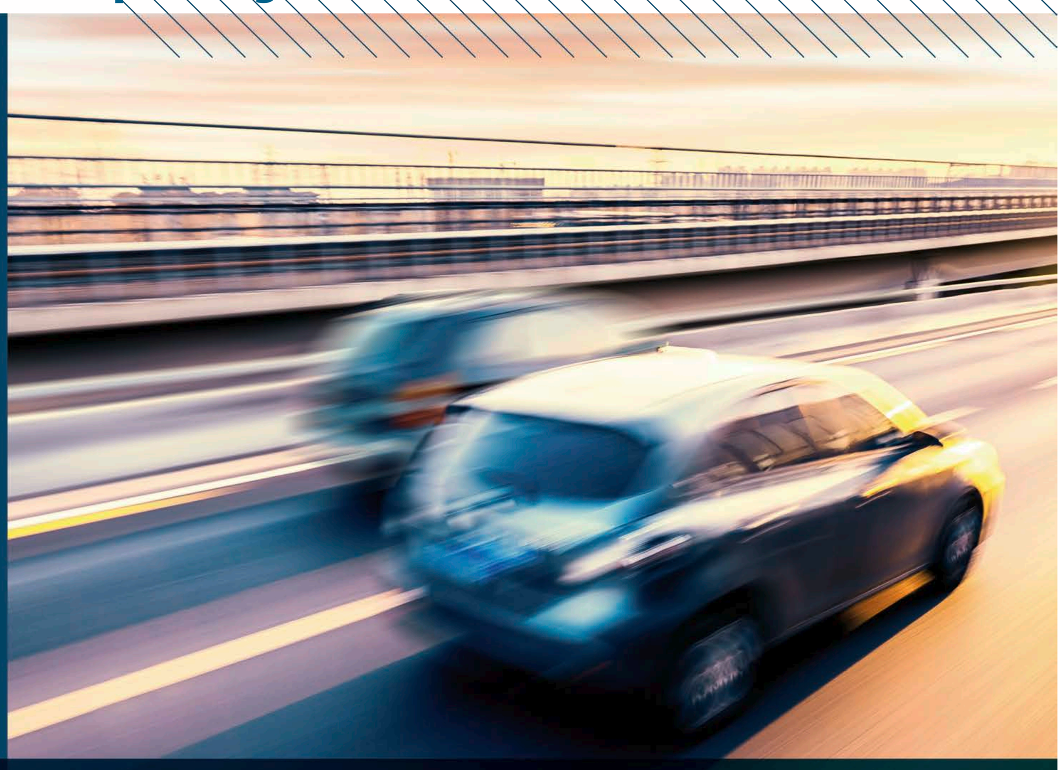
Supported by:

Federal Ministry for Economic Affairs and Energy

on the basis of a decision by the German Bundestag

Requirements & Conditions – Stand 8

# CHALLENGES OF USING A SCENARIO-BASED APPROACH

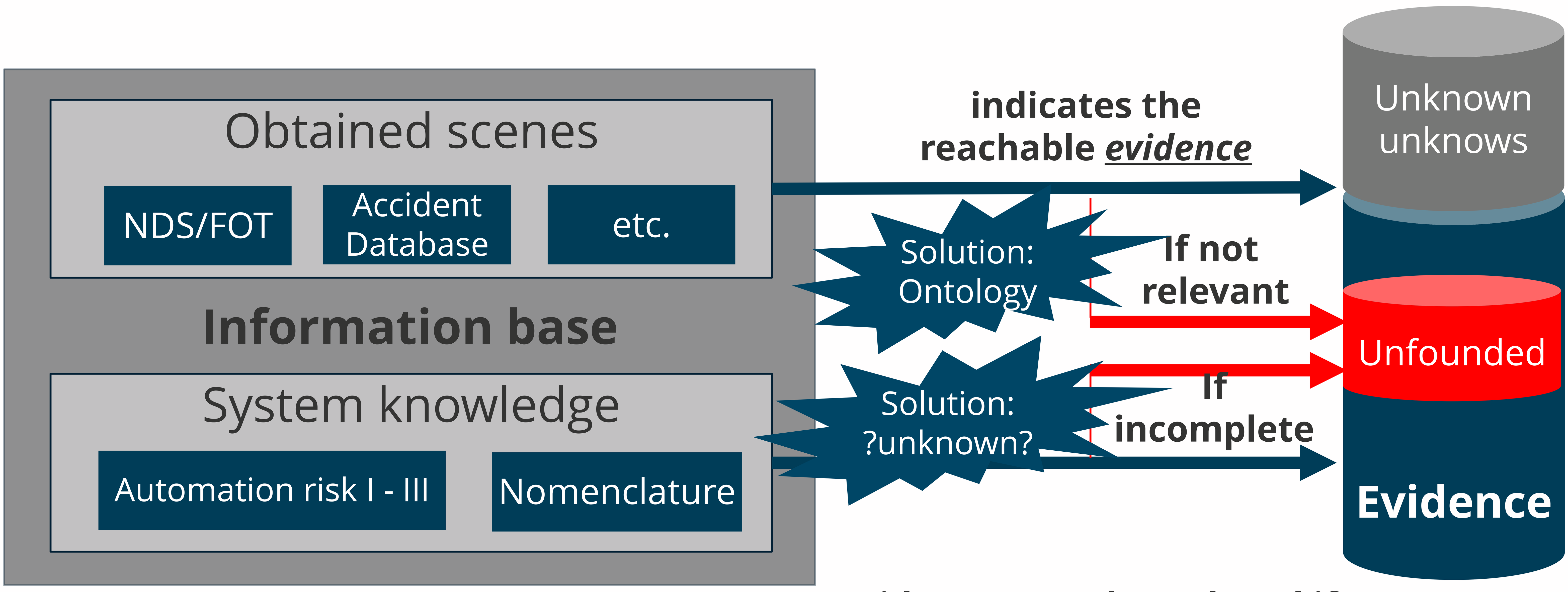


PEGASUS

## From the testing methodology towards a safety case.

How credible are the evidences provided by scenario-based approaches, Do these evidences support the safety argumentation?

1. **Information base**
- is used to provide reference data in order to demonstrate safety,
- which entails an **assumption coverage** being limited by the extent to which scenes are considered where an **unfounded evidence** must be anticipated if reference data is not representative or system knowledge is incomplete.

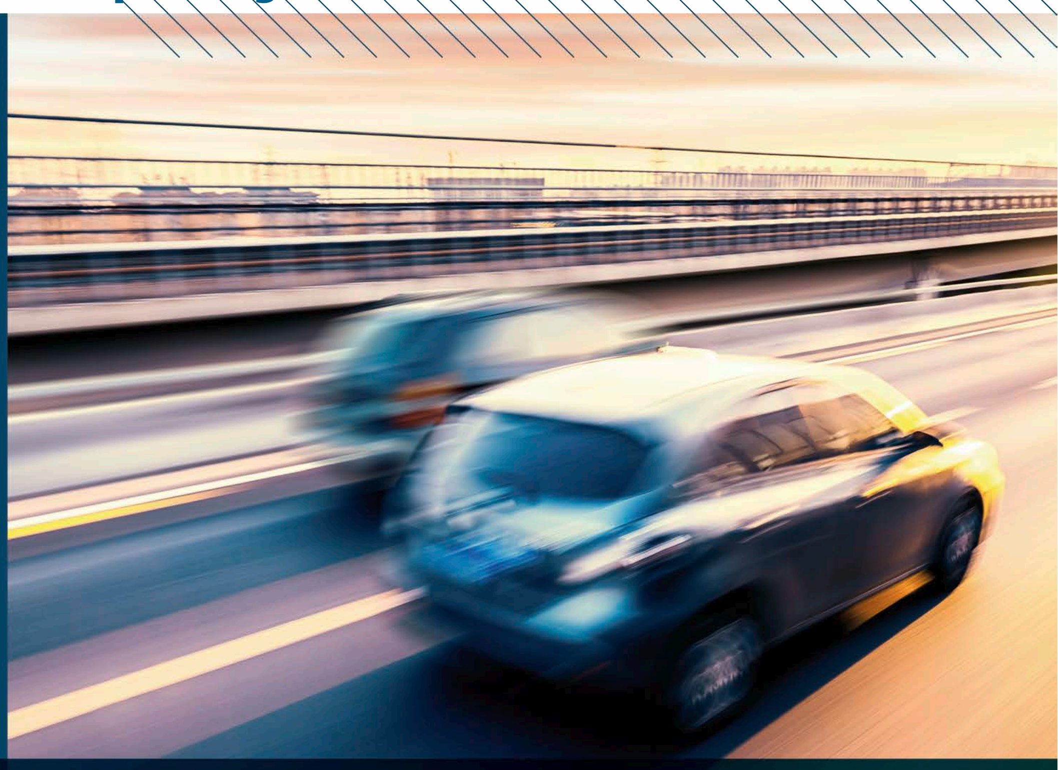



**Evidence must be reduced if System under test (SUT) cannot deal with scenarios derived from system knowledge**

The limitations regarding the assumption coverage pose a **general challenge** since evidence can only be provided for known relationships. In analogy, even the evidence provided by “black box testing“ is dependent on the extent of driving and on its representativity. Neglecting the representativity affects the creditability of both “black box testing” as well as scenario-based testing due to the fact that safety is argued but not demonstrated, which is referred to as unfounded evidence.

Regarding the system knowledge, representative “black box testing” has the advantage that respective scenarios are inherently part of the testing whereas these scenarios must additionally be generated when using a scenario-based testing method.

Supported by:
Federal Ministry for Economic Affairs and Energy
on the basis of a decision by the German Bundestag



Requirements & Conditions – Stand 8

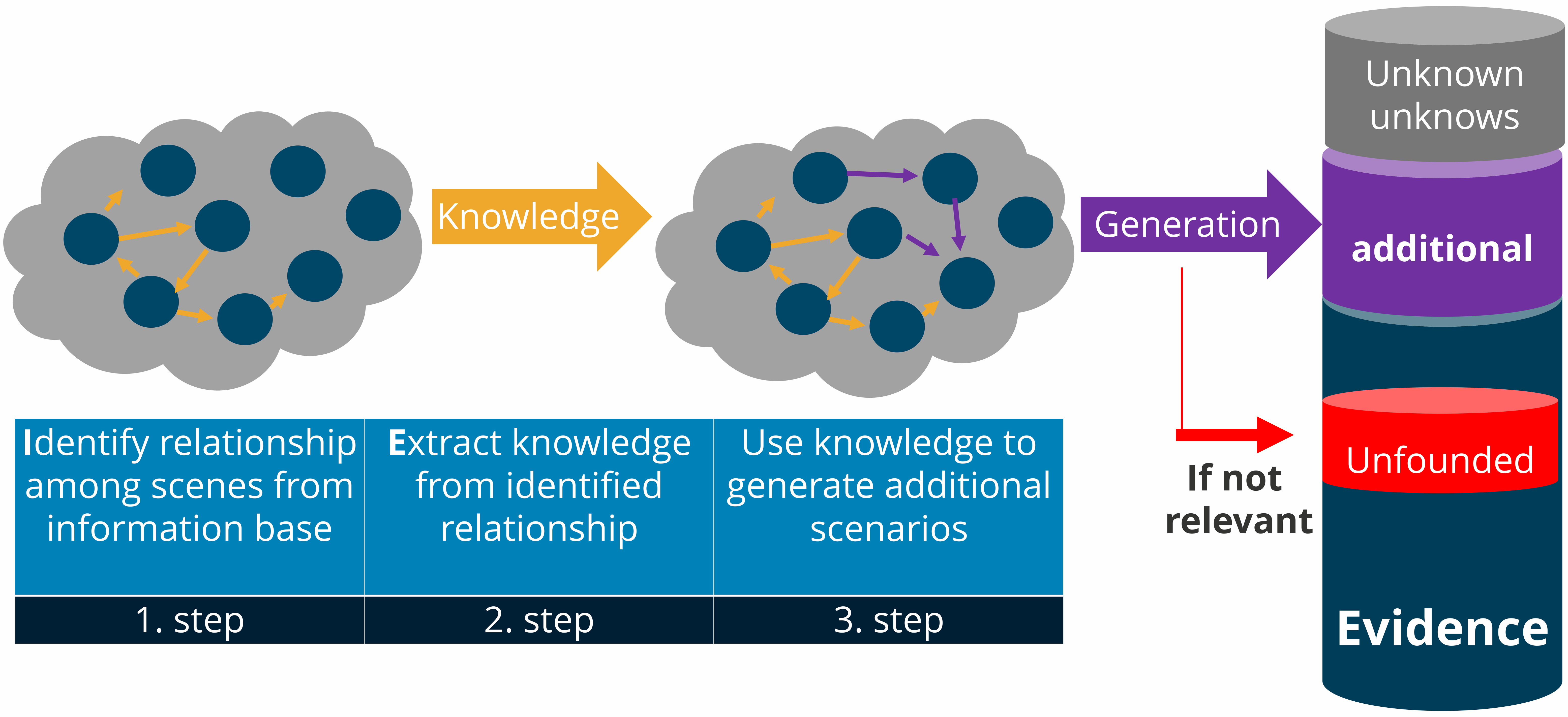
# CHALLENGES OF USING A SCENARIO-BASED APPROACH

PEGASUS

## From the testing methodology towards a safety case.

How credible are the evidences provided by scenario-based approaches, Do these evidences support the safety argumentation?

2. **Ontology**
- Is used to identify the relationship among the observed traffic scenes so that additional scenarios can be generated,
- which may entail an increased **assumption coverage** but at the expense of an **unfounded evidence** if the generated scenarios are without relevance.

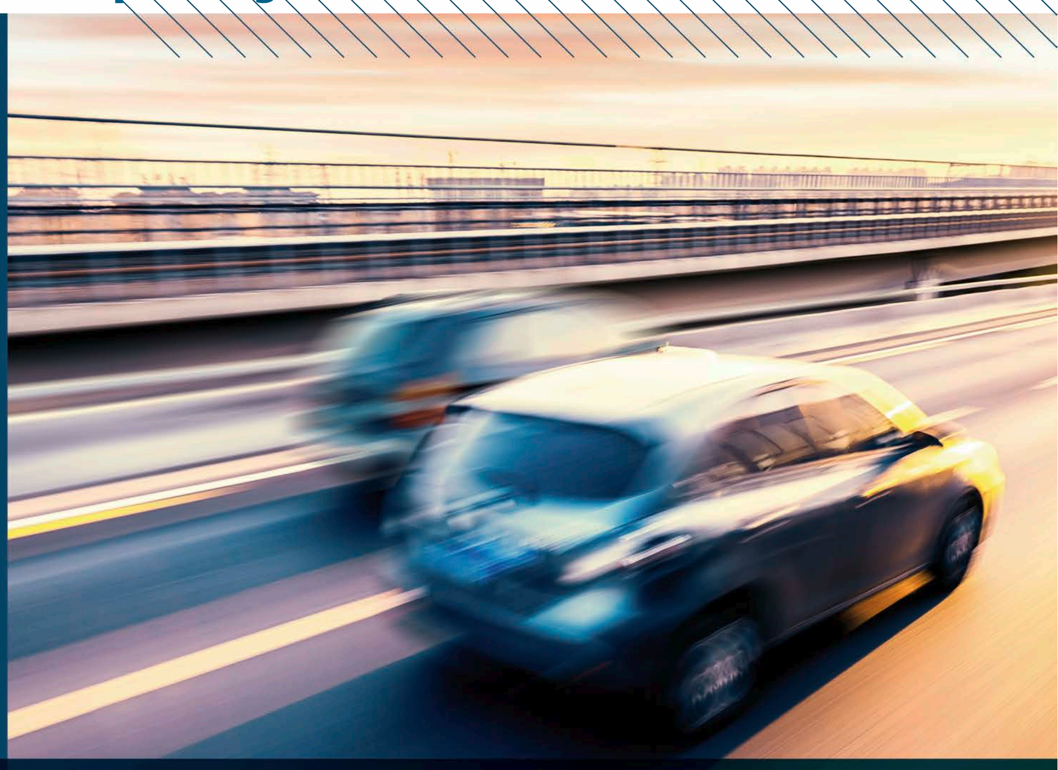



The challenge of using an ontology for the generation of new/additional scenarios consists in extracting the relationship from observed scenes in a way that its outcome is representative. Unfortunately, as long as a scenario did not occur, its representativity is usually unknown. Leaving this unaddressed may entail an unfounded evidence due to the fact that safety is demonstrated but its contribution to the safety case is left unclear. To circumvent this, we propose using the available system knowledge in order to argue the relevance of generated scenarios. In case no argumentation can be found, the generated scenarios are not allowed to contribute to the safety case.

Supported by:
Federal Ministry for Economic Affairs and Energy
on the basis of a decision by the German Bundestag

Requirements & Conditions – Stand 8

# CHALLENGES OF USING A SCENARIO-BASED APPROACH

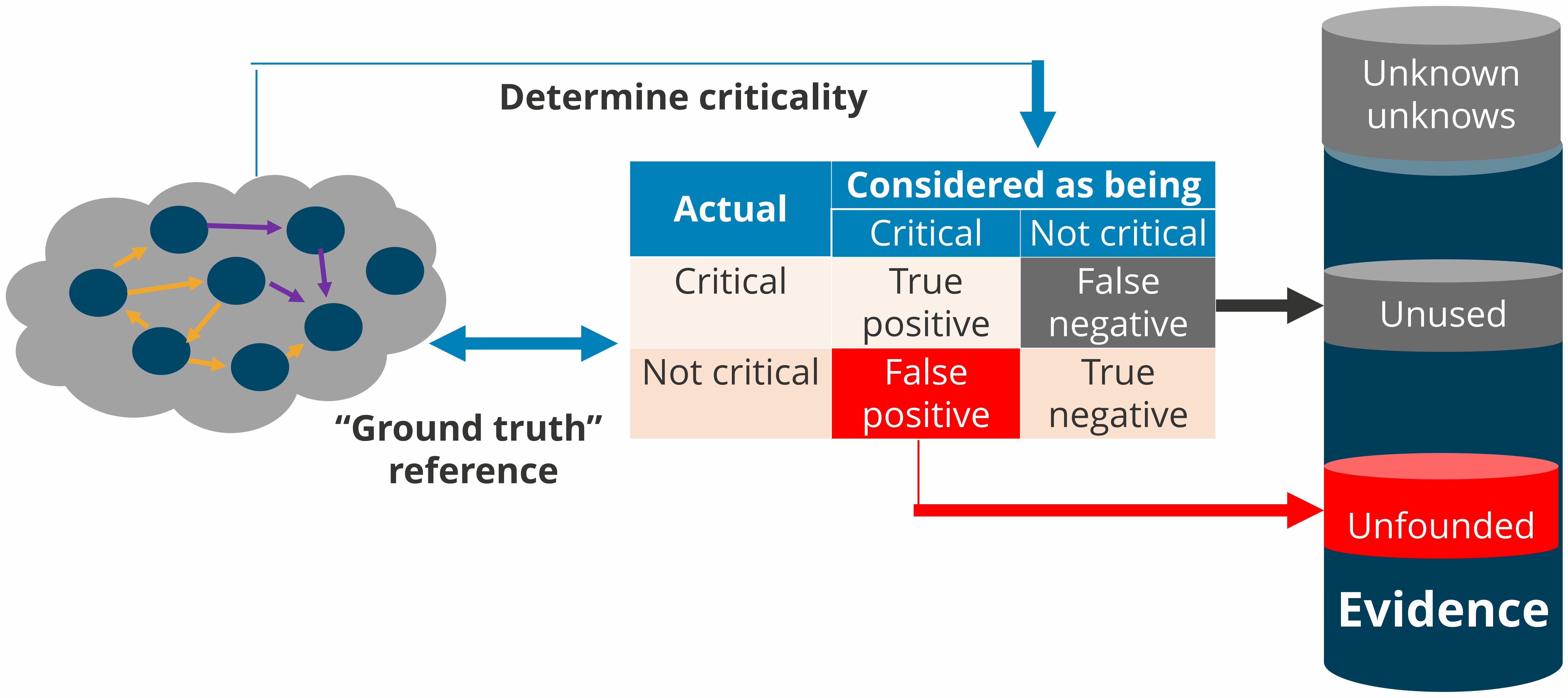


PEGASUS

## From the testing methodology towards a safety case.

How credible are the evidences provided by scenario-based approaches, Do these evidences support the safety argumentation?

3. **Criticality**
- is used to identify scenes with a potential for causing accidents,
- which entails either an **unfounded evidence** in the presence of *false positives or an* **unused evidence** in case of *false negatives.*




The challenge of determining the criticality consists in trading between false positives and false negatives causing an unfounded and an unused evidence, respectively. In case of false positives, a scene with limited or, even, without criticality is considered critical, which entails an wrong argument on the demonstrated safety. In contrast, evidence is left unused if a critical scene is considered uncritical and thus omitted during the testing. Regarding this tradeoff, it needs to be noted that the criticality is mandatory for providing an argument on why scenes need be tested.

Supported by:
Federal Ministry for Economic Affairs and Energy
on the basis of a decision by the German Bundestag

Requirements & Conditions – Stand 8

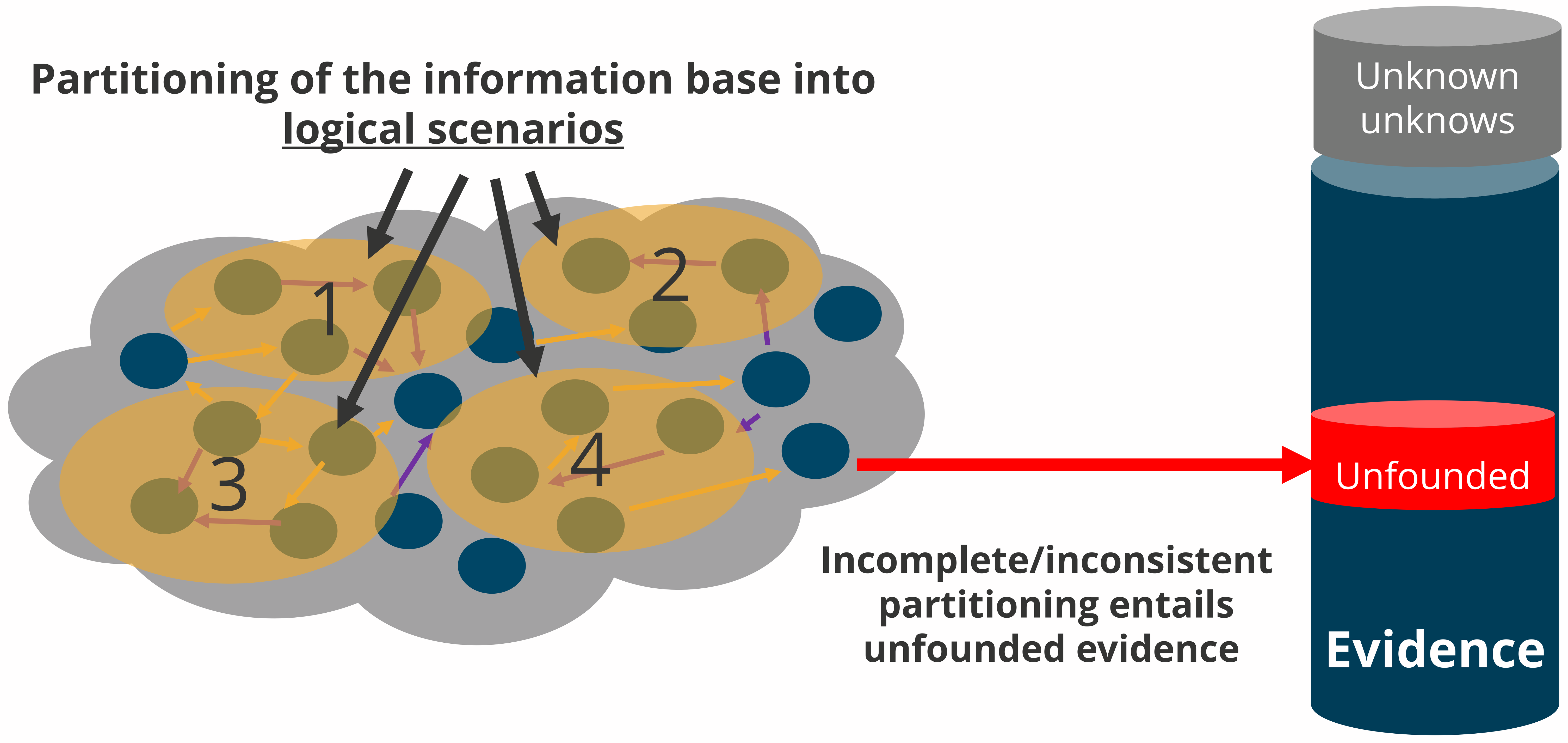
# CHALLENGES OF USING A SCENARIO-BASED APPROACH

PEGASUS

## From the testing methodology towards a safety case.

How credible are the evidences provided by scenario-based approaches, Do these evidences support the safety argumentation?

4. **Association**
- is used to partition the observed scenes into logical scenarios
- which may entail an **unfounded evidence** if the logical scenario relaxes the dynamics of the observed scene with the result that safety is argued but no supporting evidence is provided.

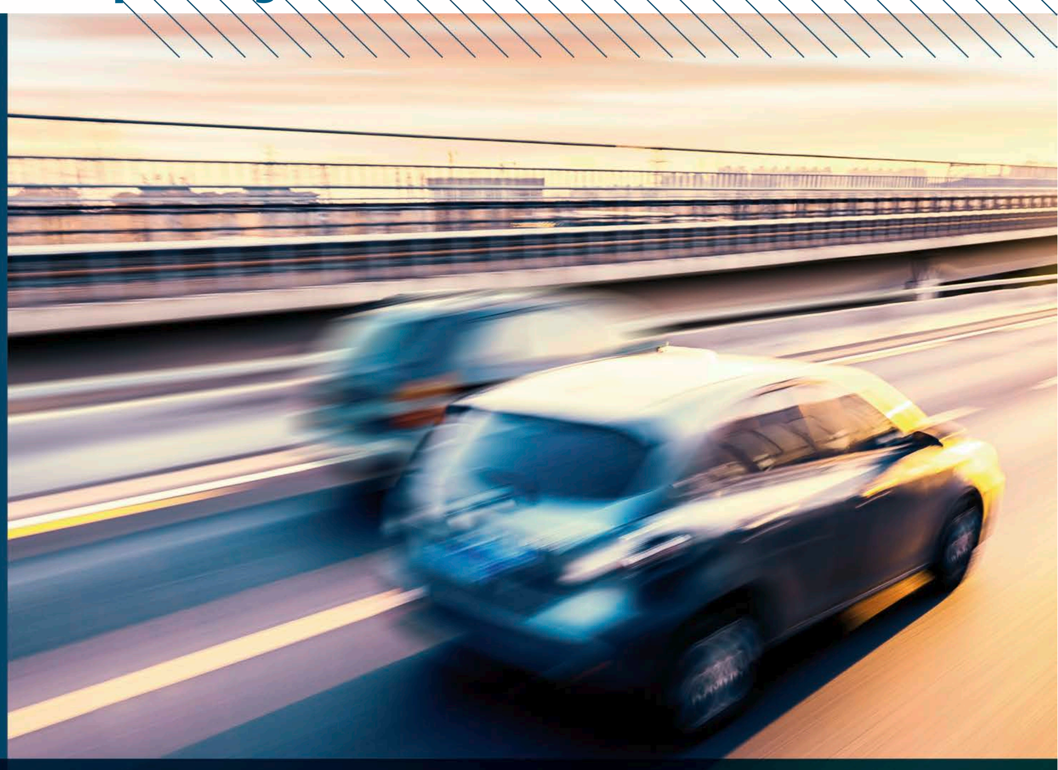



The association poses the challenge of partitioning the observed scenes such that the resulting logical scenarios do not simplify the observed traffic dynamics. In light of a simplified scene, the tests derived from these logical scenarios will be made easier to pass in comparison with the observed scene being partitioned. From this follows that safety is argued but evidence is provided only for a relaxed argument, which entails an unfounded evidence being provided for the safety case. In contrast, “black box testing” does not suffer from this issue since the observed scenes are tested without applying any changes.

Supported by:
Federal Ministry for Economic Affairs and Energy
on the basis of a decision by the German Bundestag

Requirements & Conditions – Stand 8

# CHALLENGES OF USING A SCENARIO-BASED APPROACH

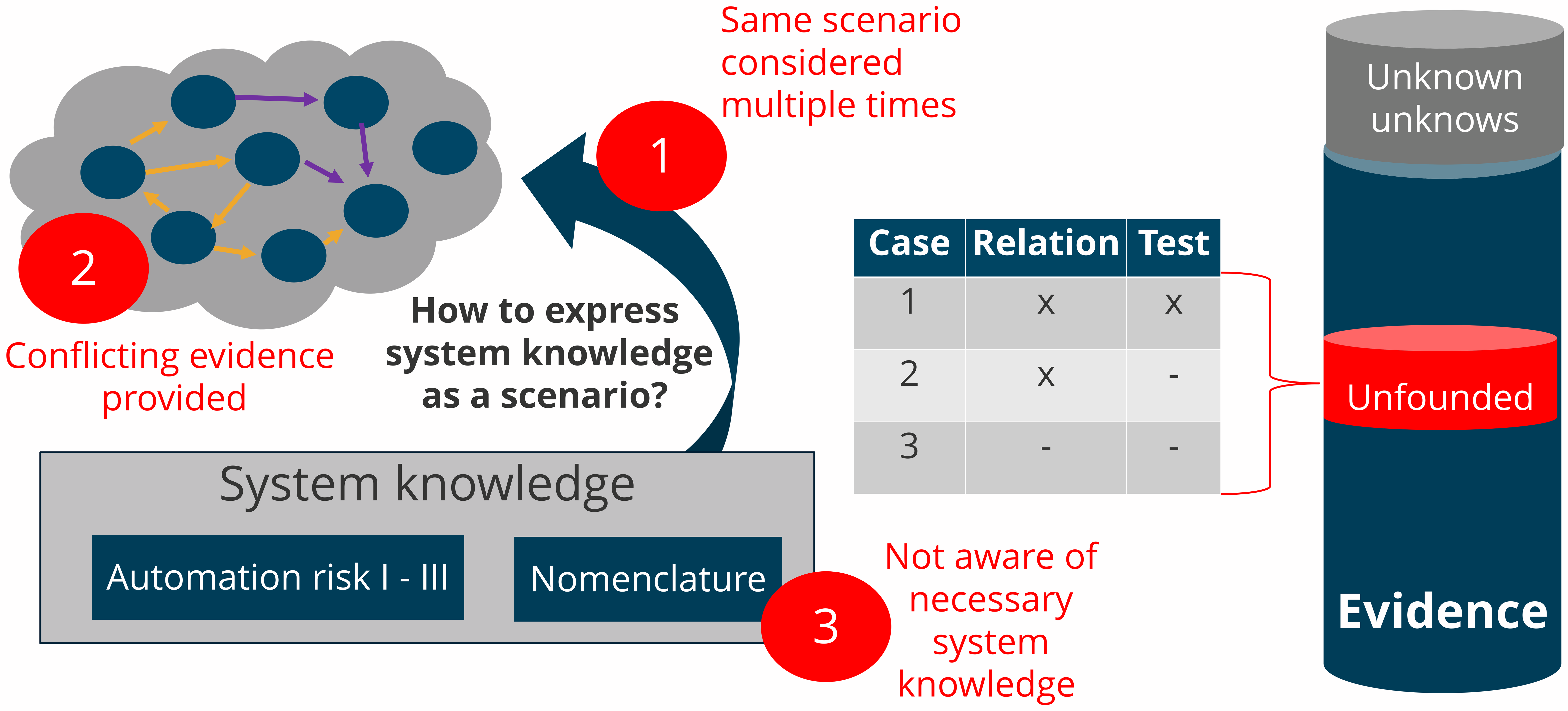


PEGASUS

## From the testing methodology towards a safety case.

How credible are the evidences provided by scenario-based approaches, Do these evidences support the safety argumentation?

5. **Relation**
- is used to express the trigger for a functional insufficiency in terms of a scenario,
- which may entail an **unfounded evidence** if multiple evidences are provided for the very same scenario due to the fact that scenarios are being tested with and without information about risks of automation.



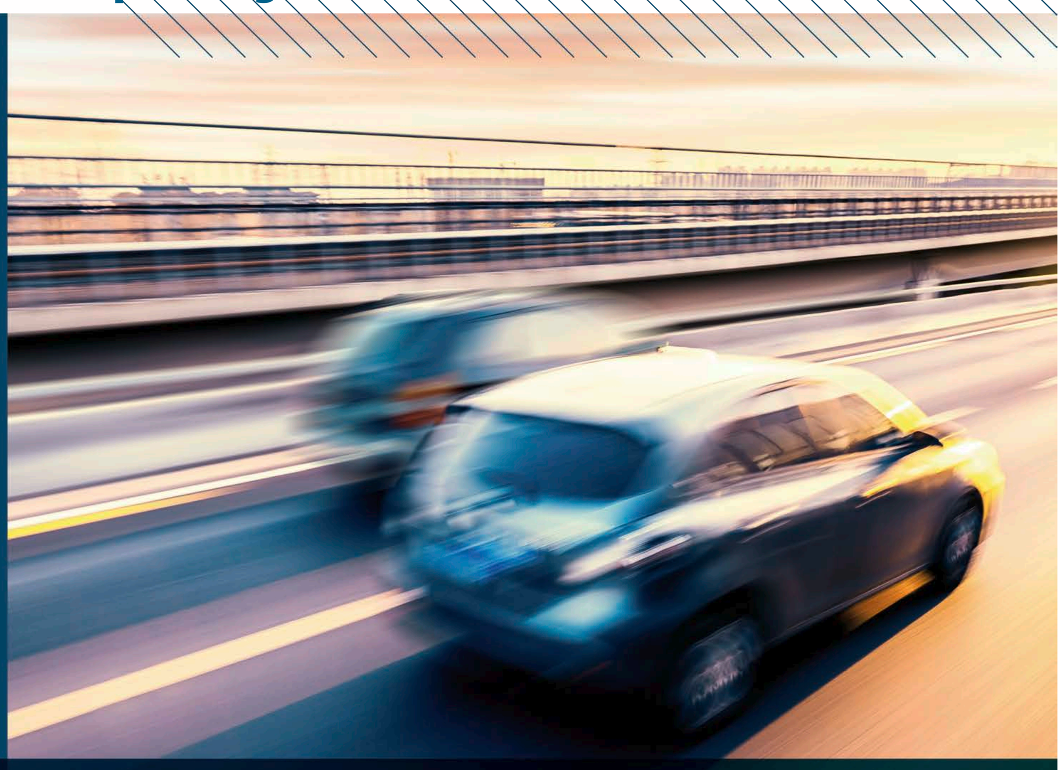

| Case | Relation | Test |
|---|---|---|
| 1 | x | x |
| 2 | x | - |
| 3 | - | - |

The relation entails the challenge of transferring system knowledge into a scenario because of a semantic gap between both and/or a lack of context information that is required to trigger the automation risk. Leaving these challenges unaddressed causes an unfounded evidence because the system is tested under relaxed conditions. In contrast, “black box testing” does not suffer from the described challenges since the triggers for an automation risk are inherently part of the observed scene. However, although the “black-box approach” can be used to test the ability of “system under test” to deal with automation risks, there is no guarantee that the necessary triggers occur merely because of excessive testing. In this respect, the scenario-based testing method has the advantage that system knowledge can reproducibly be tested once respective triggers have been identified.

Supported by:
Federal Ministry for Economic Affairs and Energy
on the basis of a decision by the German Bundestag

Requirements & Conditions – Stand 8

# CHALLENGES OF USING A SCENARIO-BASED APPROACH



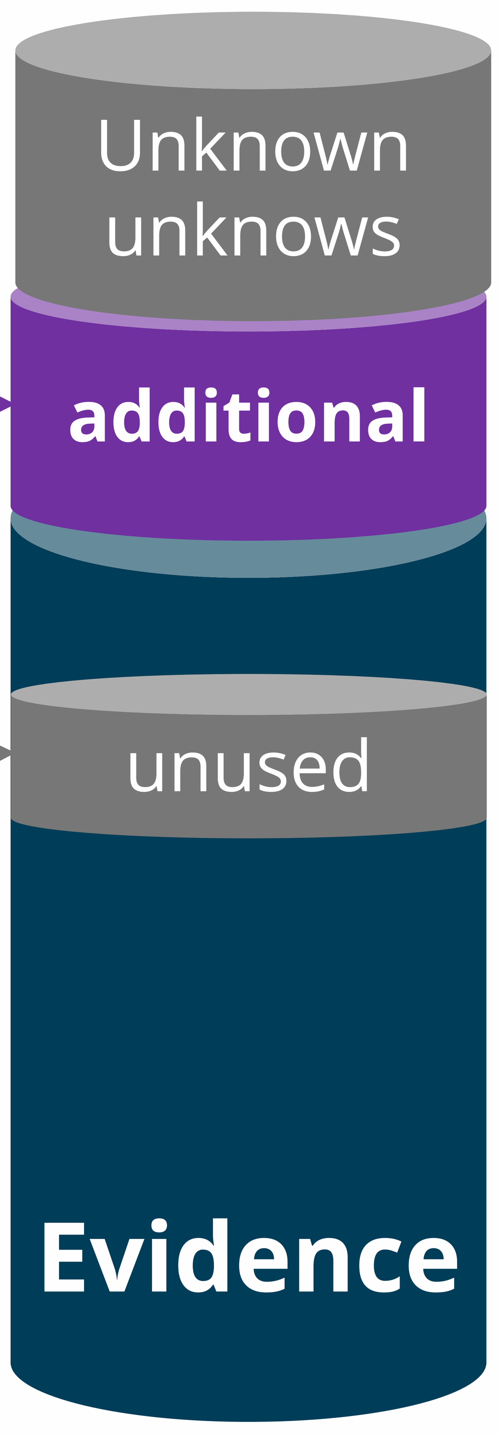
PEGASUS

## From the testing methodology towards a safety case.

How credible are the evidences provided by scenario-based approaches, Do these evidences support the safety argumentation?

6. **Variation**
- is used to test the stability of the "system under test" by changing parameters,
- which may entail an increased assumption coverage or an unfounded evidence if scenarios with or without relevance are generated, respectively.

| | | |
|---|---|---|
| **Relevant** | **Increases the assumption coverage since new scenarios can be tested that are otherwise unknown** | **Can be used to deal with unused evidence if not tested before** |
| Not relevant | Scenarios being not relevant can be tested for verifying the stability of the SUT but evidences provided by these tests are not allowed to contribute to the safety case. | |
| | Different from data base (new) | Part of data base (not new) |
| | Scenarios resulting from variation | |




The challenge of the variation consists in making changes to the test specification such that the resulting scenarios are still relevant. In cases where new scenarios being relevant are provided, the assumption coverage can be increased. However, if relevance is presumed but does not exist, unfounded evidence must be anticipated. To counteract this, we propose using system knowledge in order to argue the representativity of scenarios that are a result of the variation. If no argument can be found, the varied scenario can, nonetheless, be tested but only for testing the stability of the "system under test", which are not allowed to contribute to the safety case unless their representativity can otherwise be argued.

Supported by:
Federal Ministry for Economic Affairs and Energy
on the basis of a decision by the German Bundestag