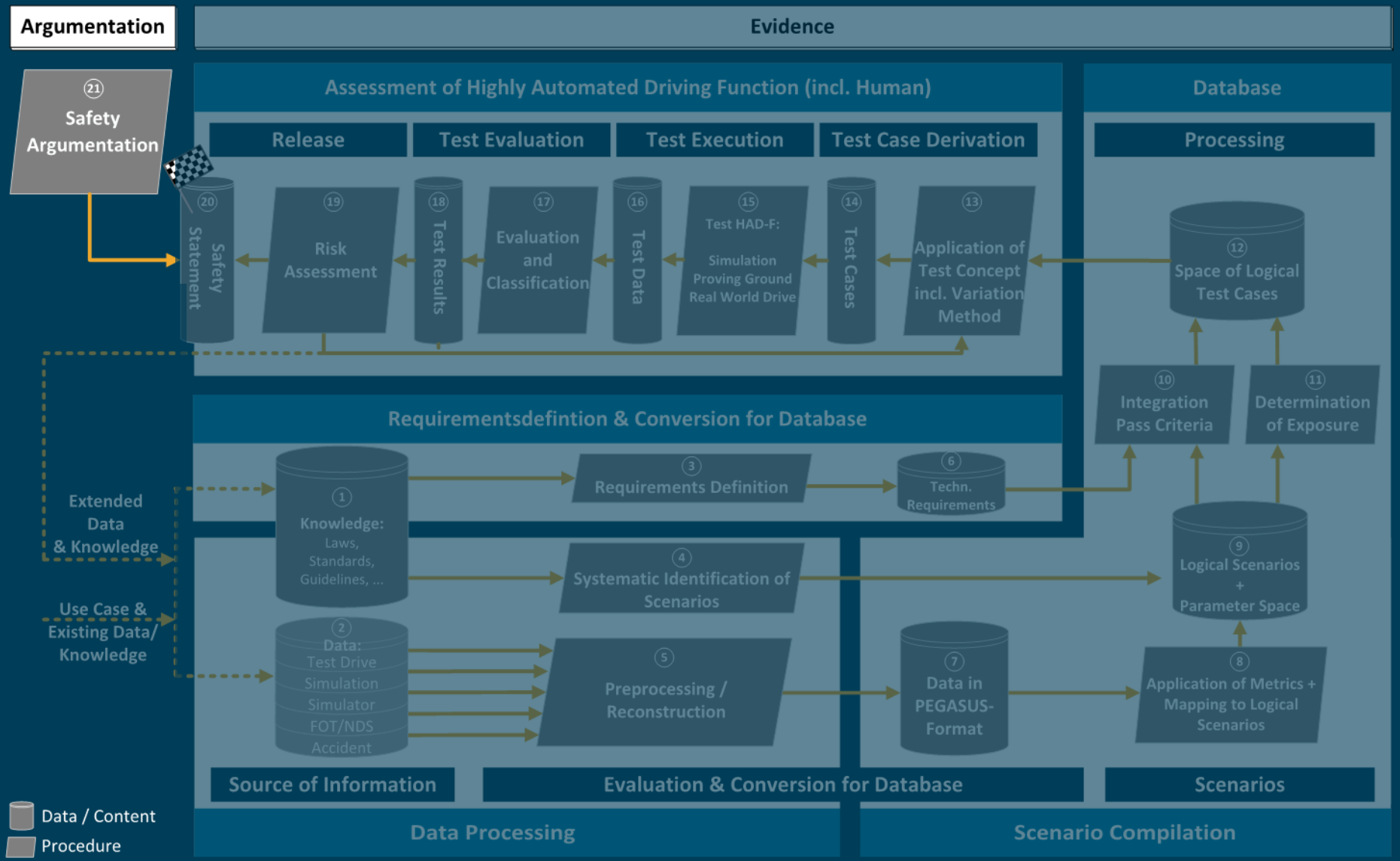
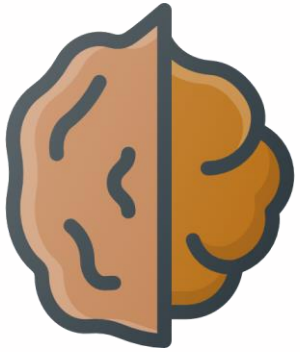
SAFETY ARGUMENT

PEGASUS

Alexander Maus, 14.05.2019

Safety Argumentation - Context





PEGASUS Safety Argumentation – in a nutshell

The PEGASUS Safety Argumentation is a **theoretical framework**.

It proposes **five layers** for this upon which **elements** are located.

Dimensions of **integrity** and **relevance** are proposed as quality criteria for the five layers with respect to their elements.

The central assumption of the PEGASUS Safety Argumentation is: if a chain of arguments, which was created following the method of the PEGASUS Safety Argumentation, stands up to a **critical examination**, an approval recommendation can be given.

PEGASUS Safety Argumentation – fundamental paradigms

Structuring

A layered structure attempts to bring more clarity into the continuing discussions regarding an approval, particularly surrounding highly automated driving functions.

Formalisation

Formalisation within the PEGASUS Safety Argumentation means making the individual elements of a layer explicit by means of a defined, standardised and ideally established notation.

Coherence

Only once there is success in reasonably linking elements, can one show how safety and reliability can be verified and how an argumentation can be made for socially accepted, highly automated mobility in the larger context.

PEGASUS Safety Argumentation – deliveries

PEGASUS

Proof of concept

An exemplary demonstration of how a PEGASUS Safety Argumentation might look like.

Delivery: Visualization

Implementation

Software tools to „build“ a Safety Argumentation

Delivery: Not within PEGASUS Project.

Documentation

Theoretical framework to support an approval recommendation particularly aimed at highly automated driving functions.

Delivery: A written document explaining the PEGASUS Safety Argumentation.

PEGASUS Safety Argumentation – a hierarchical concept

Highly Automated Driving Functions (HAD-F) are widely accepted in the public.

There is an understanding of what factors foster acceptance of HAD-F.

Top level goals are set to be met in order to achieve acceptance of HAD-F.

Logical structure is described to reach that goals.

The strategies are implemented using methods and tools.

Results become evident when they can be traced back to the achievement of a goal.

Layers of the Argumentation

Highly Automated Driving Functions (HAD-F) are widely accepted in the public.

There is an understanding of what factors foster acceptance of HAD-F.

0

Motivation

Context

Top level goals are set to be met in order to achieve acceptance of HAD-F.

1

Logical structure is described to reach that goals.

2

The strategies are implemented using methods and tools.

3

Results become evident when they can be traced back to the achievement of a goal.

4

Argumentation / Approval Recommondation

leading questions

Highly Automated Driving Functions (HAD-F) are widely accepted in the public.

There is an understanding of what factors foster acceptance of HAD-F.

0

Why do we need this evidence?

Top level goals are set to be met in order to achieve acceptance of HAD-F.

1

Why is the result evident?

Logical structure is described to reach that goals.

2

Why was the result achieved?

The strategies are implemented using methods and tools.

3

How was the result achieved?

Results become evident when they can be traced back to the achievement of a goal.

4

What is the result?

Integrity of the argumentation

Highly Automated Driving Functions (HAD-F) are widely accepted in the public.

There is an understanding of what factors foster acceptance of HAD-F.

0

e.g. validity

Top level goals are set to be met in order to achieve acceptance of HAD-F.

1

e.g. completeness

Logical structure is described to reach that goals.

2

e.g. traceability

The strategies are implemented using methods and tools.

3

e.g. adequacy, reliability/ reproducibility

Results become evident when they can be traced back to the achievement of a goal.

4

e.g. representativeness / correctness

Roles & involvement

Society

Regulator

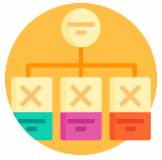
Industry

Highly Automated Driving Functions (HAD-F) are widely accepted in the public.

There is an understanding of what factors foster acceptance of HAD-F.

Top level goals are set to be met in order to achieve social acceptance of HAD-F.

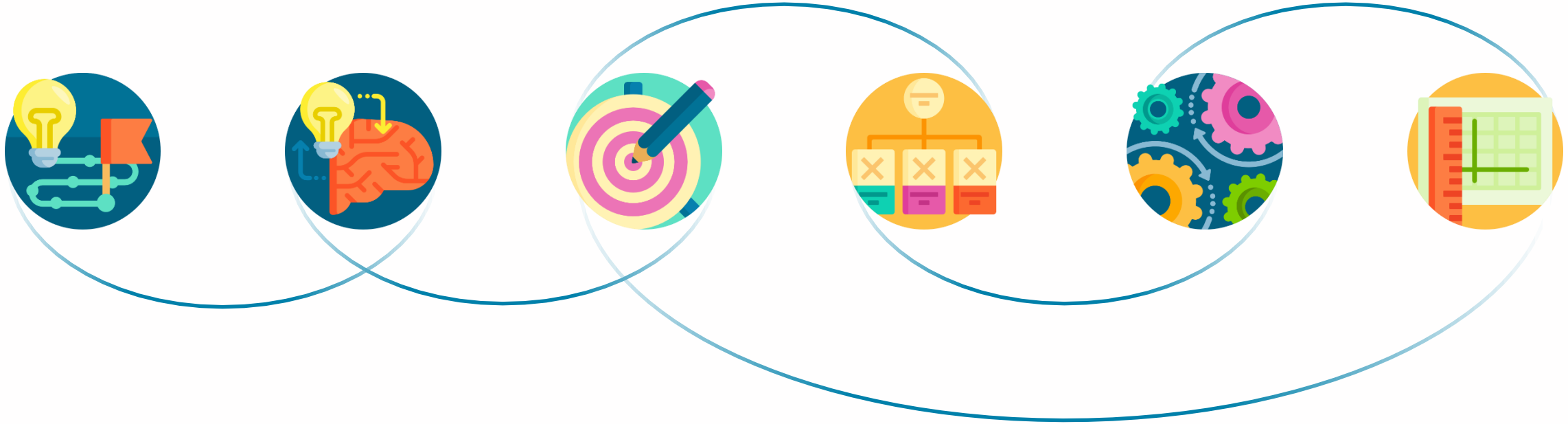
Logical structure is described to reach that goals.

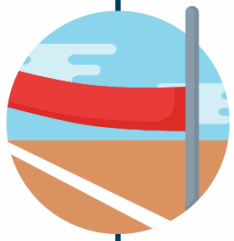


The strategies are implemented using methods and tools.

Results become evident when they can be traced back to the achievement of a goal.

Process





The central assumption of the PEGASUS Safety Argumentation is: if a chain of arguments, which was created following the method of the PEGASUS Safety Argumentation, stands up to a critical examination, an approval recommendation can be given.

Some open issues (not to be solved within PEGASUS)

Highly Automated Driving Functions (HAD-F) are widely accepted in the public.

There is an understanding of what factors foster acceptance of HAD-F.

0

Further research needed

Top level goals are set to be met in order to achieve acceptance of HAD-F.

1

Are the NHTSA design principles sufficient?

Logical structure is described to reach that goals.

2

What do OEMs already do to reach that goals?

The strategies are implemented using methods and tools.

3

What methods and tools already exist?

Results become evident when they can be traced back to the achievement of a goal.

4

What results can be achieved already?

5 questions, 5 layers, 5 paradigms

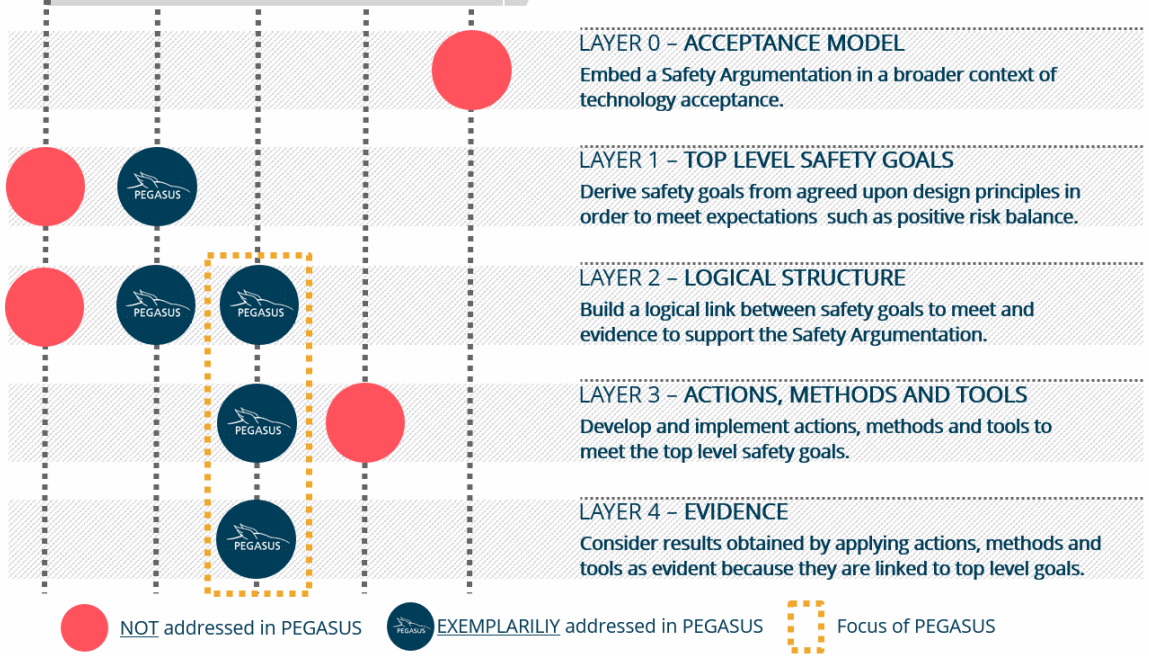
What do people expect from highly automated mobility in the larger context? — *e.g. convenience, mobility benefits, data privacy*

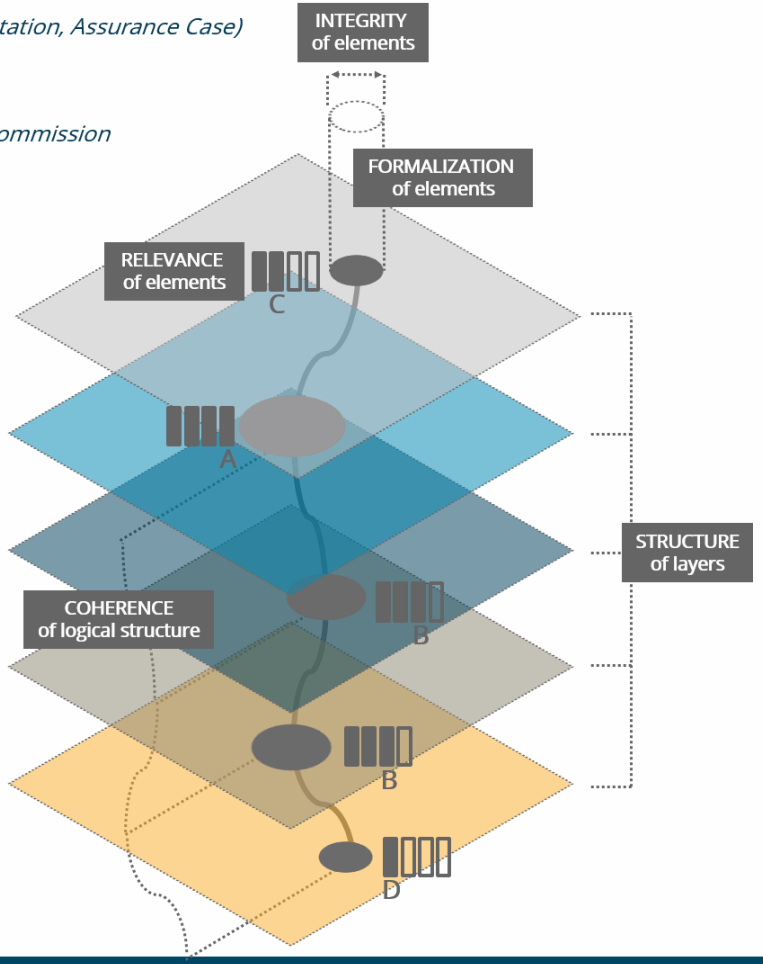
How to monitor a HAD-F over its entire life cycle? — *e.g. field monitoring*

How to support a Safety Argumentation with evidence? — *e.g. safety case (Goal Structuring Notation, Assurance Case)*

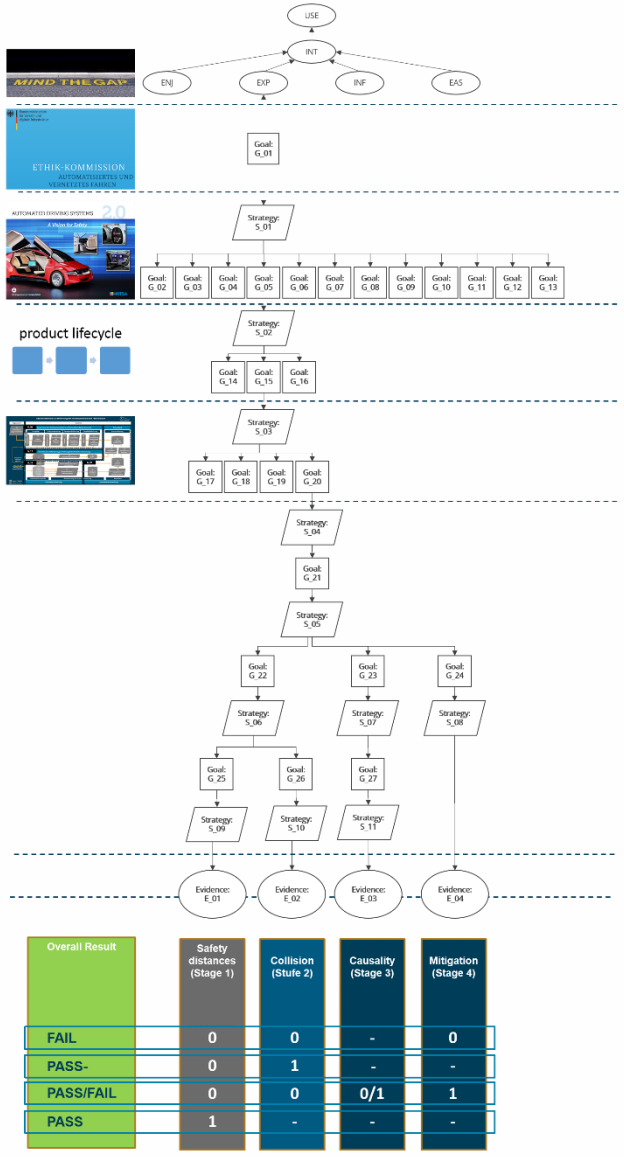
Which elements should a Safety Argumentation contain sufficiently? — *e.g. NHTSA design principles, German ethics commission*

How to ensure necessarily compliance with all relevant standards and directives needed for homologation? — *e.g. ISO 26262, ISO/PAS 21448:2019(E)*





Example



element	definition	relevance	integrity
G_01	HAD-F reduces the level of harm until it is completely prevented (positive balance of risks).	A	high, no other goals on that level of hierarchy to contradict
S_01	Define principles to follow in order to reach G_01 (decompose).	C	high, decomposition seems to be a reasonable and is in line with the NHTSA approach for example
G_05	Process and procedure for verifcation and validation of HAD-F behavioral performance with the prescribed ODD is documented and considered	A	medium, unless there is a an agreed upon list of safety goals on that level of hirarchy it is very likely that there is a certain degree of overlap among goals on that level of hierarchy
S_02	Adapt process and procedure for verifcation and validation of HAD-F behavioral performance with the prescribed ODD to life cycles stages of the HAD-F.	C	high, it seems reasonable to cover the whole life cycle of a product for safeguarding
G_15	Adapt to Product Development	C	medium, further clarification of life cyde stages respectively particular requirements might be needed here
S_03	Devide process and procedure for verifcation and validation of HAD-F nominal performance with the prescribed ODD for Product Development into sub processes and sub procedures	C	high, there is a practical need for decomposition in order to operationalize processes and procedures
G_20	Assess behavioral performance of HAD-F with the prescribed ODD	C	high, one of the main goals of PEGASUS
...	...	...	...
E_01	Safety distances are respected (in particular testcase)	C	medium, check for violation of safety distances is implementable, further clarification might be needed
E_02	No collision (in particular testcase)	C	high, check for collision is practicable
E_03	Collision not caused by HAD-F	C	low, lack of implementable definition of causation
E_04	HAD-F mitigated collision	C	medium, further clarification of mitigation strategy to assess might be needed