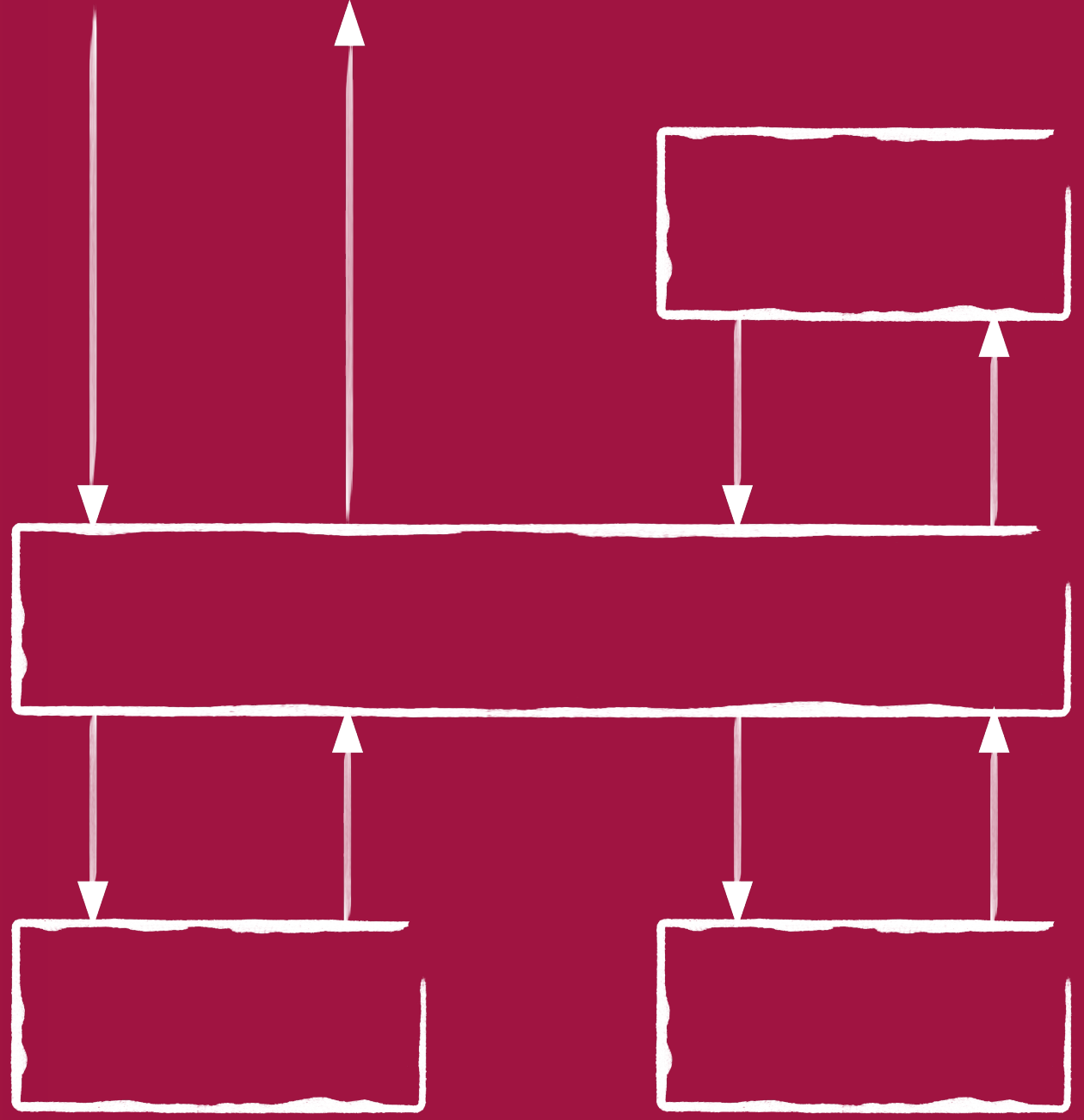
Value by-Design

STPA AS A TOOL FOR VALUE-BASED ENGINEERING



Science Finds, Industry Applies, Man Conforms?

Motto of the Chicago World Fair 1933



But science has found: Technology shall be built for humans!

Today’s goal is to develop systems to:

+ Avoid human error (being a symptom of flawed systems)
+ Consider human needs and values

Value-based Engineering

Humanistic ideas in the digital age

Current debates on how humanistic ideas can be preserved in the digital age:

+ Digital Humanism [1]
+ Value-Based Engineering (VBE) [2]

Goals of Value-based Engineering

+ Integrating individual and social-ethical values with technical requirements
+ Ensuring traceability between values and technical solutions
+ Targeting organizations, projects, clients, and suppliers
+ IEEE 7000-2021: A standard to support

„Do you know a method that can help here?"

Value-based Engineering

+ Design of systems with explicit consideration of individual and societal ethical values
+ Traceability between values and technical solutions

STPA

+ Prevention of losses
+ A loss involves something of value to stakeholders
+ Every STPA result is traceable to one or more losses



IEEE 7000-2021

A standard to support value-based engineering

The following process steps are defined in IEEE 7000 [3]:

1. Creation of the system concept from the user's perspective and the system context
2. Identification and prioritization of ethical values
3. Definition of ethical requirements
4. Ethical risk-based design

STPA is discussed here as a suitable methodology for conducting these processes.

Also available as ISO/IEC/IEEE 24748-7000:2022

IEEE 7000-2021

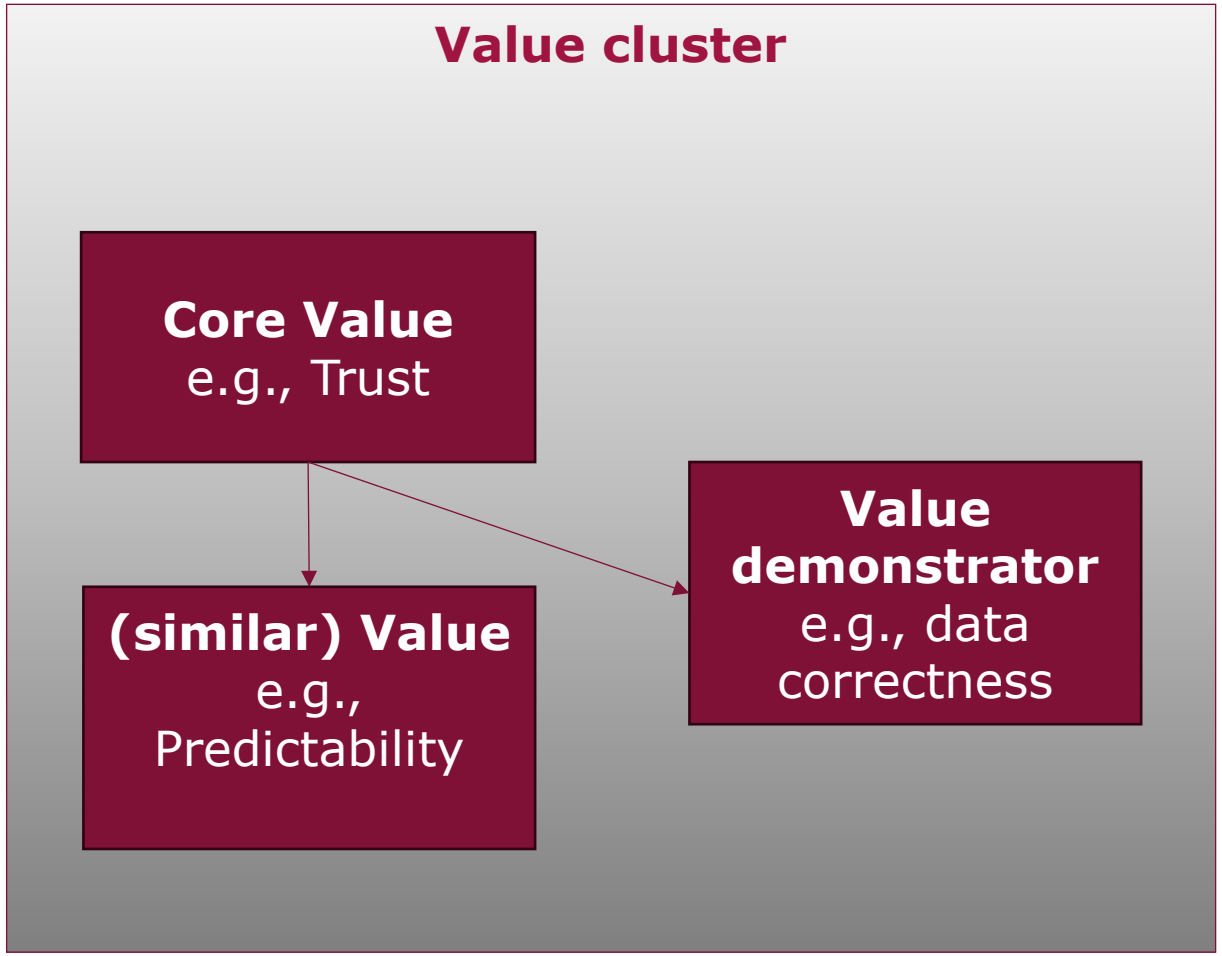
Additional information relevant for the presentation

Value clusters: Similar values grouped together, represented by a core value.

So-called "value demonstrators" can be defined:

+ Manifestation of values
+ Examples of how values can be supported or harmed
+ E.g., "data accuracy" is a value demonstrator, crucial for values such as "trust" and "predictability"

Values must be ordered by priority for necessary trade-off decisions



Comparison of definitions

An excerpt of the most relevant definitions for this presentation

An STPA loss is the loss or damage of positive values or the occurrence of negative values

The standard defines "value harms": often intangible effects on values and ethics of stakeholders

+ "Value harm": undermining of a value, not necessarily a complete loss of value
+ STPA "hazard": a system state that, (only) under worst-case conditions, leads to a "loss"
+ The terms "value harm" and the STPA "hazard" each represent a state potentially leading to a loss

Ethical value requirements (EVR, organizational or technical requirements catering to identified values) can be seen as or derived from STPA constraints.

IEEE 7000	STPA
Harm	Loss
	Harm
Value harm	Hazard
Hazard	
EVR	Constraint

PRACTICAL EXAMPLE

STPA analysis in context of VBE

EVmate – a (fictional) app for EV drivers

An app for route planning and energy payment for electric vehicles

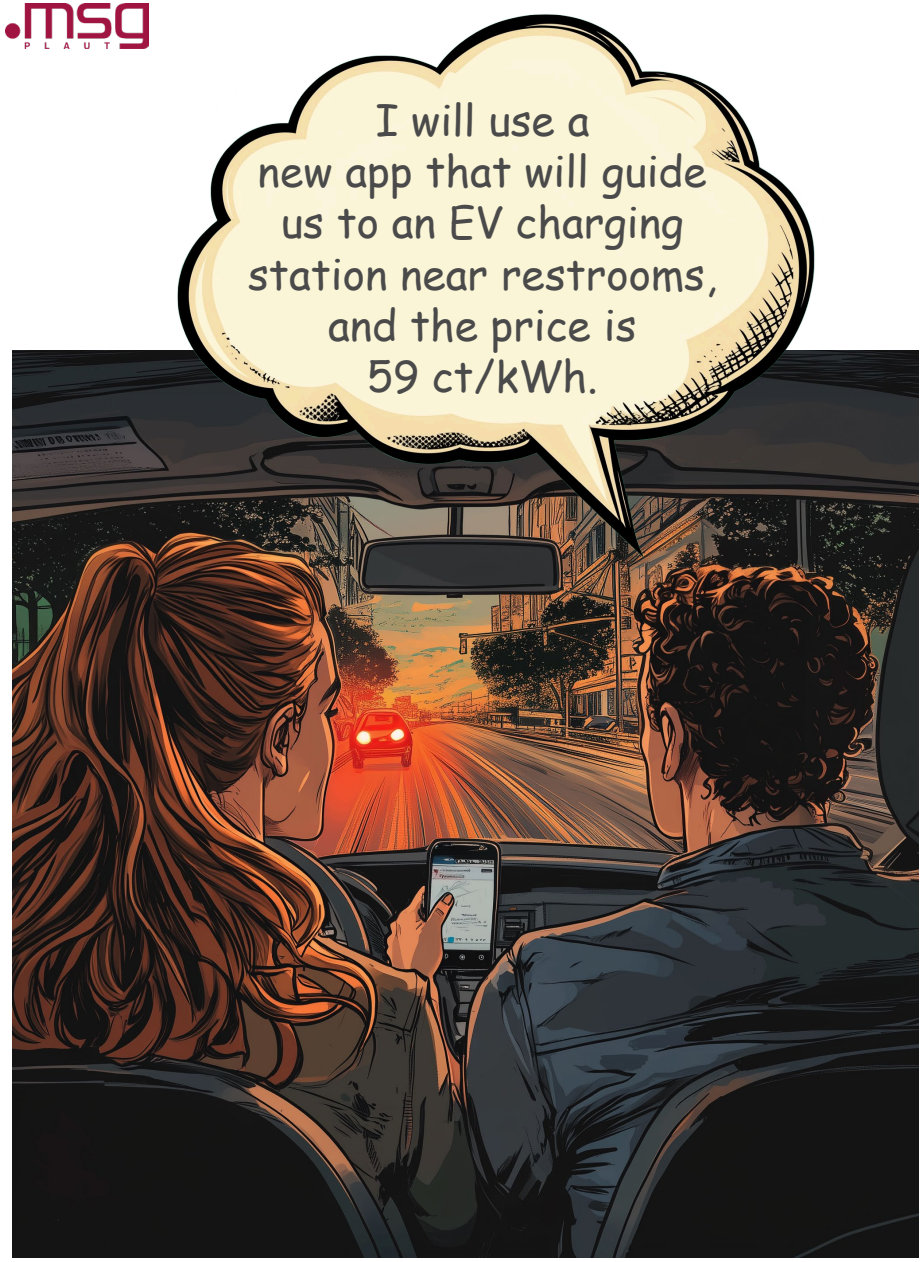
The app is intended to offer the following functions and services:

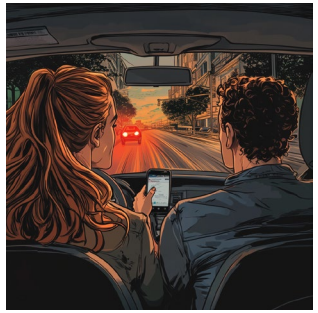
+ Payment for the energy charged at various charging station providers via the app

+ Price per kWh at a certain charging station

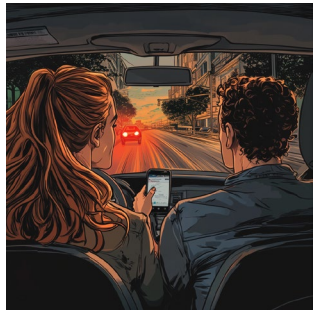
+ Information about infrastructure (restrooms, restaurants, stores, …) near a charging station

+ Route planning with AI-based optimization, such as optimizing on
  - overall travel time
  - charging price
  - infrastructure near a charging station, etc.

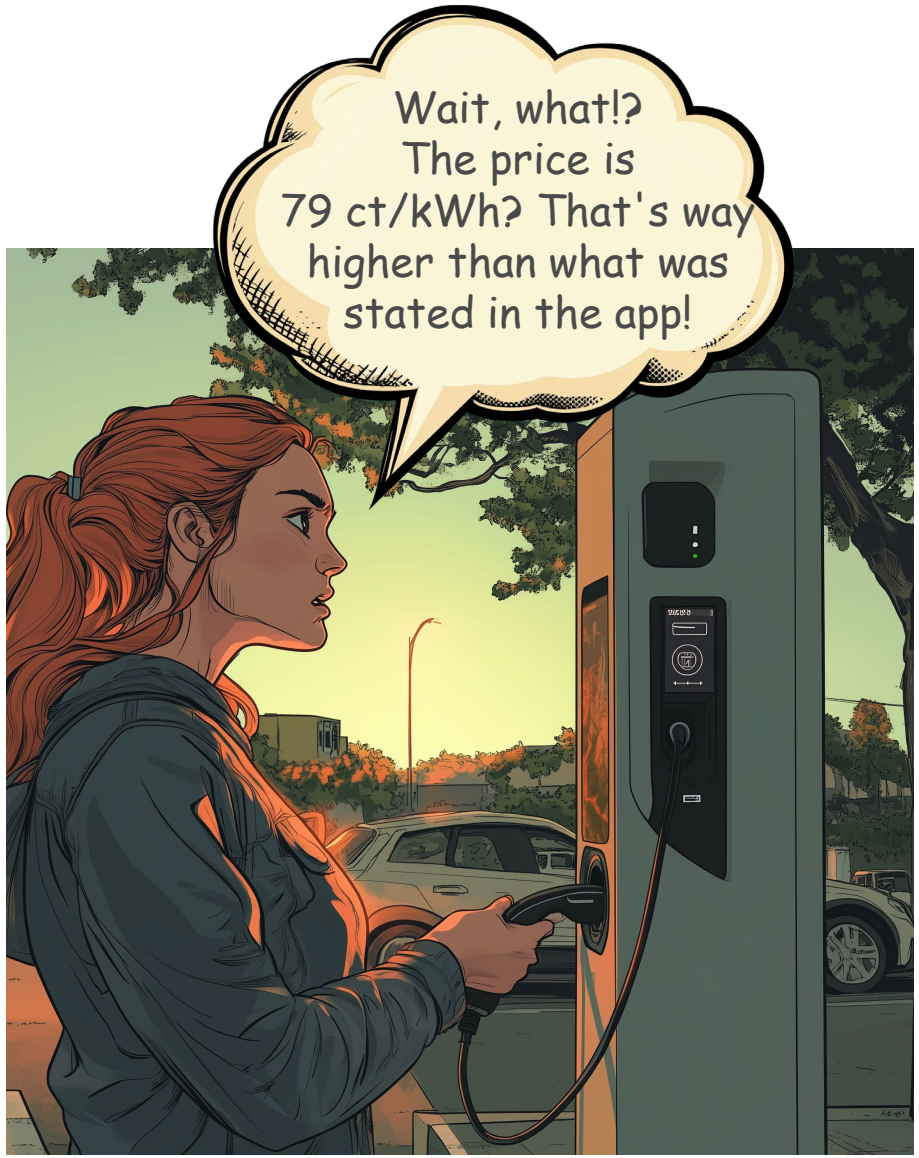

I will use a new app that will guide us to an EV charging station near restrooms, and the price is 59 ct/kWh.

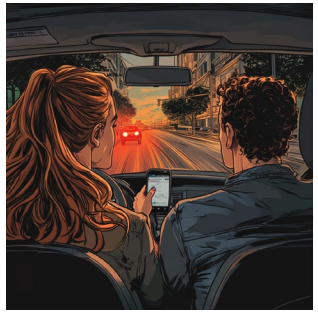



But there aren't any restrooms here...



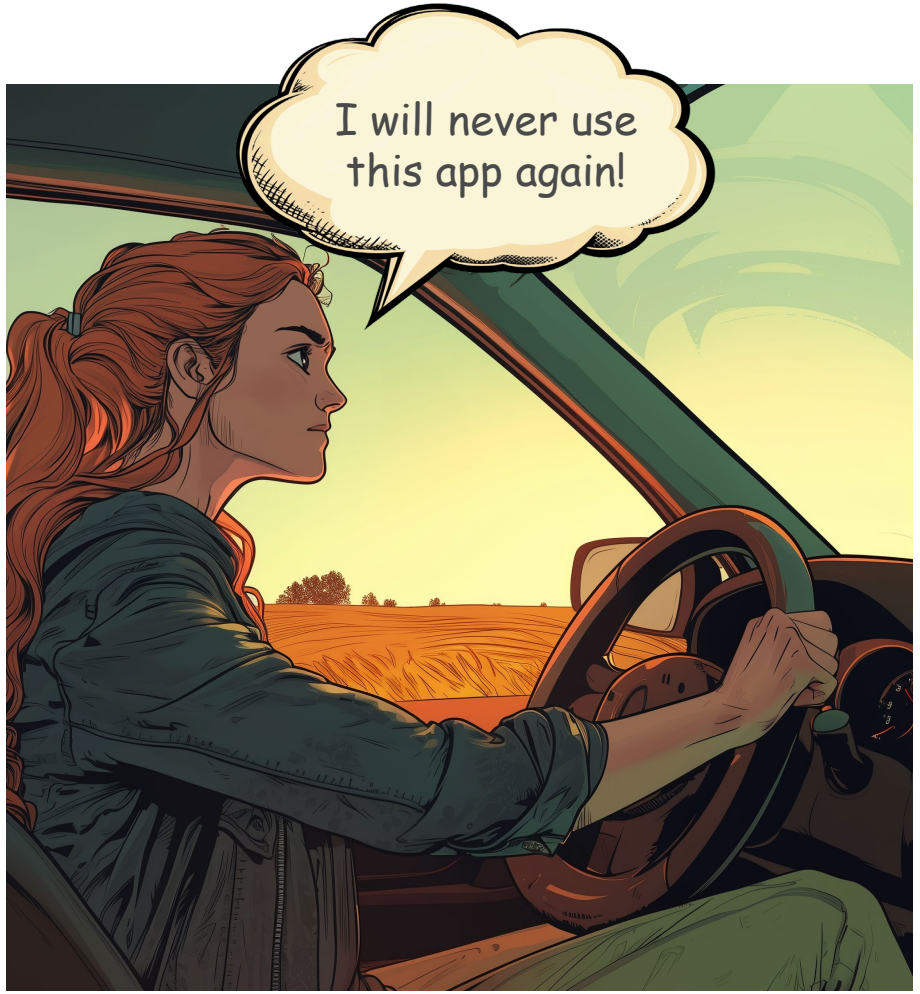



Wait, what!?
The price is
79 ct/kWh? That's way
higher than what was
stated in the app!

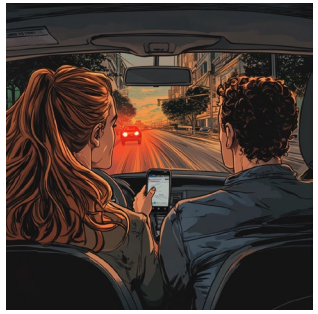







I will never use this app again!

Relevant values for customer acceptance






Value	Description
Autonomy	Self-determination in goals and actions
Control	Decision-making authority, options for choice
Sustainability	Environmental protection, maintainability, renewability
Safety	No danger to life, property, or the environment
Transparency	Traceability, accessibility
Trust	Competence, integrity, predictability, reliability

Ethical Values Elicitation in IEEE 7000

How the standard can support conduction of STPA

Definition of roles for ethical value elicitation

+ Roles to conduct the process
+ Roles to be consulted during the process

Guidance to ask the right questions

+ The standard does not determine what is ethical or not
+ Ethical theories contribute questions to think about a set of values relevant for the organization
+ List of typical value (positive and negative) and questions to ask in the process

STPA analysis

An examplary appliation of STPA to fulfil IEEE 7000-2021

The following example will focus on loss

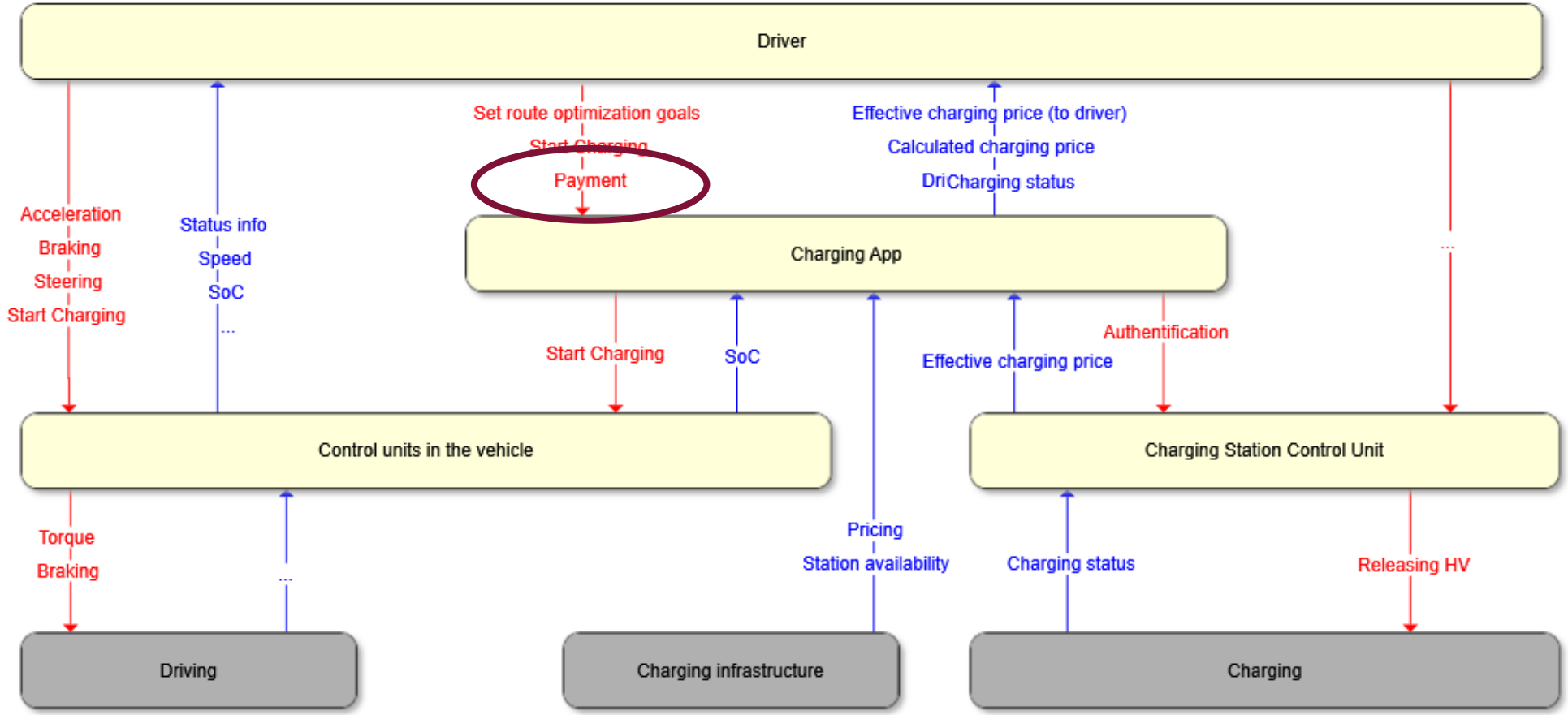
L-8 - Loss of Trust

+ Value cluster "trust" includes, e.g., the values of the app's predictability, reliability, and accuracy.
+ If the promised services are only offered unreliably, trust in the app diminishes

One of the identified hazards is

H-5 - The app provides unreliable information and instructions [L-1, L-6, L-8]

STPA analysis – Control structure



STPA analysis – UCA and loss scenario

Further, UCA-1 is discussed:

UCA-1 - Driver provides a higher payment than the effective charging price communicated by the app [H-5].

One loss scenario that could lead to this UCA is LS-1:

LS-1 - Due to a flaw in the pricing algorithm of the charging app, the effective charging price communicated to the driver is incorrect. Consequently, the driver pays more than the amount specified by the app [UCA-1].

"Value demonstrators" can be derived from the loss scenarios

+ For example, LS-1 shows that the algorithm for price calculation algorithm is a "value demonstrator"
+ A faulty algorithm in the negative sense but a correct one in the positive sense

Traceability between loss and value demonstrator is a solid and explainable guidance for design decisions.

Conclusion and Outlook

…and a Disclaimer

Focus of this work was the proof-of-concept:

+ Some IEEE 7000 process steps were omitted because little novelty was expected within the scope of this work
+ Some process steps of IEEE 7000 and STPA were not conducted with the required personnel as needed

Planned future work

+ Conduct the processes with the required personnel
+ The process to fulfill IEEE 7000 using STPA should be more formalized
+ Conduction of omitted steps

Conclusion and Outlook

How STPA can support conduction of IEEE 7000

VBE practitioners recommend analyzing systems with high-value expectations using comprehensive risk-based system design methods [2]

+ STPA proves to be a suitable method in this context.

IEEE demands to address control and observability over system elements

+ STPA control structure supports this goal

Conclusion and Outlook

How STPA can support conduction of IEEE 7000

"Core values" and "value demonstrators" are sometimes overshadowed by apparent technical solutions

+ E.g., some claim Tesla’s "powerful battery" to be a "core value" [4]
+ In reality, the battery supports "value demonstrators" such as flexibility, time savings, and range, which are essential for the value of "mobility"

Clear traceability between "values," "value demonstrators," and technical solutions is essential

+ Helps to tweak the right parts of the system design
+ Subsequently, the benefits of certain "values" are increased and "value harms" reduced

Let‘s connect!



Florian Wagner

Senior Business Consultant Mobility Solutions
https://www.linkedin.com/in/florian-wagner-at
florian.wagner@msg-plaut.com



Florian Wagner

Lead Business Consultant Mobility Solutions
https://www.linkedin.com/in/andreas-kerschl-2827a1142/
andreas.kerschl@msg-plaut.com

Thank you!

msg Plaut Austria GmbH
Modecenterstraße 17, Unit 4 / 6.OG
1110 Vienna, Austria

office.at@msg-plaut.com
msg-plaut.com

References

[1] The Vienna Manifesto on Digital Humanism: https://caiml.org/dighum/dighum-manifesto/Vienna_Manifesto_on_Digital_Humanism_EN.pdf

[2] Spiekermann, Sarah. Value-Based Engineering: A Guide to Building Ethical Technology for Humanity, Berlin, Boston: De Gruyter, 2023. https://doi.org/10.1515/9783110793383

[3] IEEE 7000-2021: IEEE Standard Model Process for Addressing Ethical Concerns during System Design

[4] Osterwalder, A., & Pigneur, Y. (2010). Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers. Hoboken, New Jersey, USA: John Wiley & Sons.