Conference:
Human-Centered Digitalization: How to Develop Next Generation of Humans and Robots for a Secure, Harmonic and Prosperous Future of Europe and Japan?"
Graz, 20 & 21 September 2019

Workshop: Ethics in the digitalized Era: Western and Eastern Context

A EUROPEAN APPROACH FOR TRUSTWORTHY TECHNOLOGIES

Hristina Veljanova, MA <hristina.veljanova@uni-graz.at>
Norah Neuhuber, MSc <norah.neuhuber@v2c2.at>

Das Land Steiermark
Wissenschaft und Forschung, Gesundheit und Pflegemanagement

ZUKUNFTSFONDS STEIERMARK

UNI GRAZ

virtual vehicle

"Trust in digitalization using the example of systems for (partially) autonomous driving and driver assistance" → SAE Level 3

VERDI

Zukunftsfonds Styria/ Government of Styria

Funding

April 2019 – March 2021

Period

University of Graz & Virtual Vehicle

Project partners

AVL

Subcontractor

Ethics Law Sociology Psychology

Disciplines

1) VERDI Criteria Catalogue for trustworthy driver assisted systems &
2) Regulatory recommendations
3) Methods for trust assessment

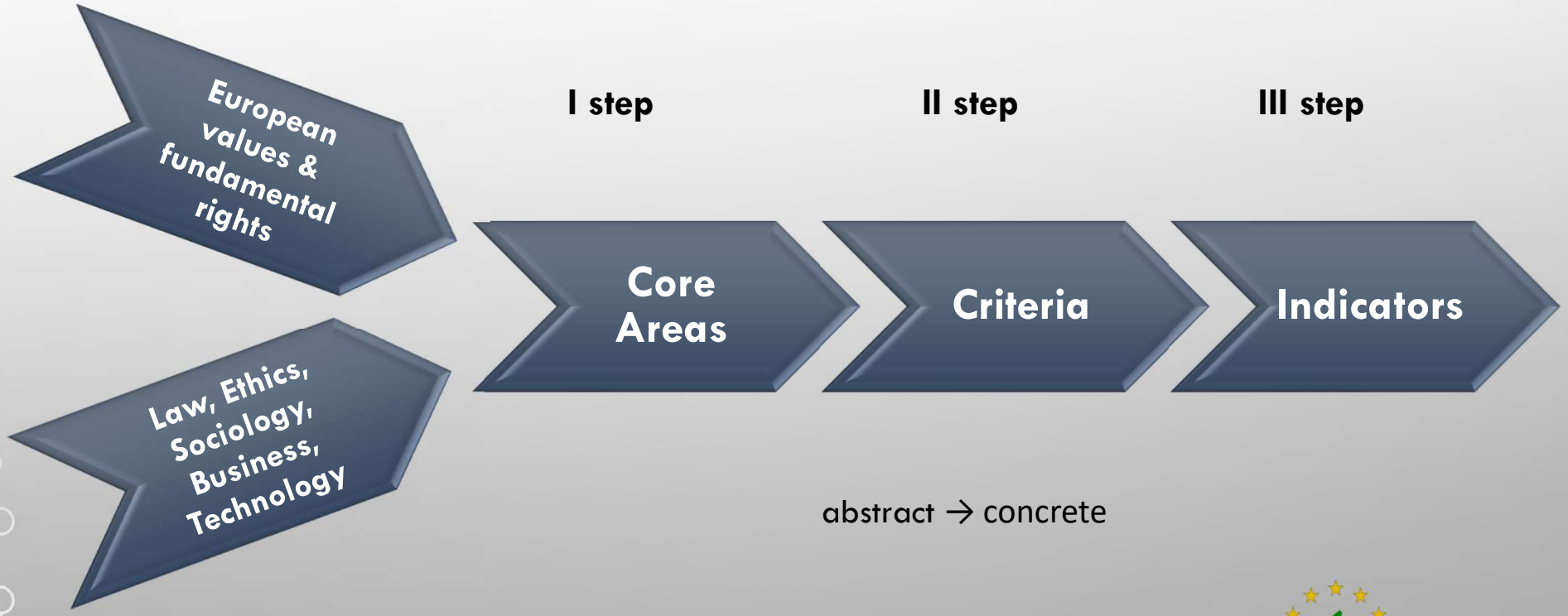
Project outcome

TRUESSEC.eu

SCOTT
Secure Connected Trustable Things

TRUESSEC.eu

APPROACH: TRUESSEC.EU CRITERIA CATALOGUE FOR TRUSTWORTHY ICT PRODUCTS AND SERVICES



TRUESSEC.eu

TRUESSEC.EU CORE AREAS OF TRUSTWORTHINESS

1) Transparency
2) Privacy
3) Anti-discrimination
4) Autonomy
5) Respect
6) Protection

TRUESSEC.eu

Example: Core Area Transparency

Ethics	Law	Sociology	Business	Technology	TRUESSEC.eu Core Areas
Transparency relates to two aspects: o providing clear and sufficient information about the products and services o providing information to users regarding activities with their personal data.	**Transparency** as in information duties laid down in the GDPR, the Directive on consumer rights or the e-commerce Directive.	Only a minority **reads privacy statements** (less than a fifth) in general while about 4 out of 10 of internet users **read the terms and conditions** on online platforms; Over 90% **want to be informed** if their data ever was lost or stolen; About half of internet users **consider themselves not well informed** about the risks of cybercrime; those who feel well-informed are more likely to adapt their security behavior (e.g. changing passwords)	**Transparency** includes a wide range of business processes which range from being clear about terms of use of the online service, through to publishing transparency reports about the passing on of user data to 3rd parties, such as law enforcement. Transparency of service and use of personal data is increasingly being perceived by business as a competitive advantage. Finally, transparency may also refer to the ability to switch services from one provider to another. Therefore initiatives such as 'open banking' (under the Second Payment Services Directive) are examples of market transparency.	**Transparency** (in data protection) is defined as the property that all personal data processing can be **understood** (intelligible and meaningful) at any time by end-users (i.e., before, during, and after processing takes place). This in accordance with the second aspect described in the Ethics column as well as the legal transparency. As for 'clear and sufficient information about products and services', in the technical domain there is a concept named 'Service Level Agreement', SLA for short. SLAs describe technical specification of the service/product being used. You may think e.g. on a service availability, uptime, etc.	➔ **Transparency** The ICT product or service is provided in line with information duties regarding personal data processing and the product/service itself.

TRUESSEC.eu

TRUESSEC.EU CRITERIA OF TRUSTWORTHINESS

- Information
- User-friendly consent
- Enhanced control mechanisms
- Privacy commitment
- Unlinkability
- Transparent processing of personal data
- Anti-discrimination
- Cyber security
- Product safety
- Statement of legal compliance
- Appropriate dispute resolution
- Protection of minors

TRUESSEC.eu

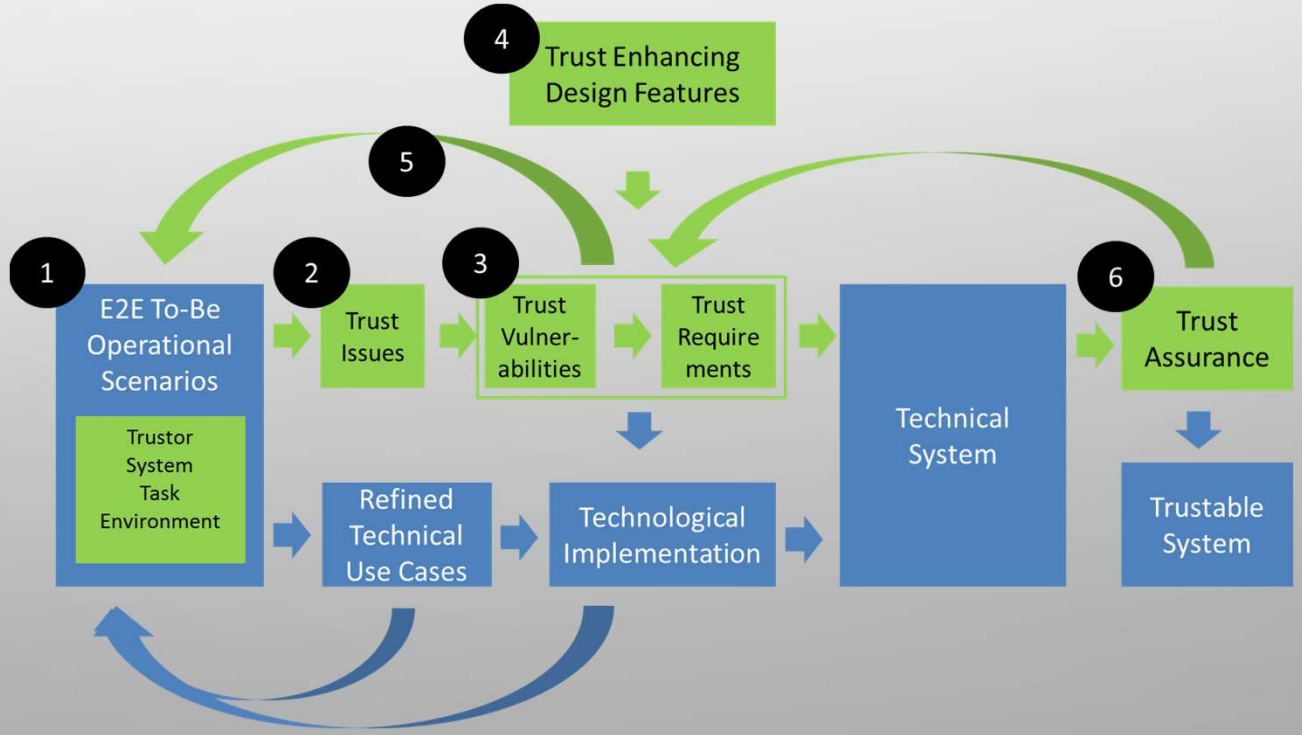
TRUSTWORTHINESS ENHANCER	CRITERION	INDICATORS
Transparency Privacy Anti-discrimination Autonomy Respect Protection	**Information**	i. Information is provided: a. in a user-friendly manner • in a plain language (understandable to lay persons) • as long as necessary and as short as possible (e.g. in a form of one pager) b. relevant to the context c. clearly visible and easy to locate d. in a structured machine-readable format. ii. Information is provided free of charge.

SCOTT
Secure Connected Trustable Things

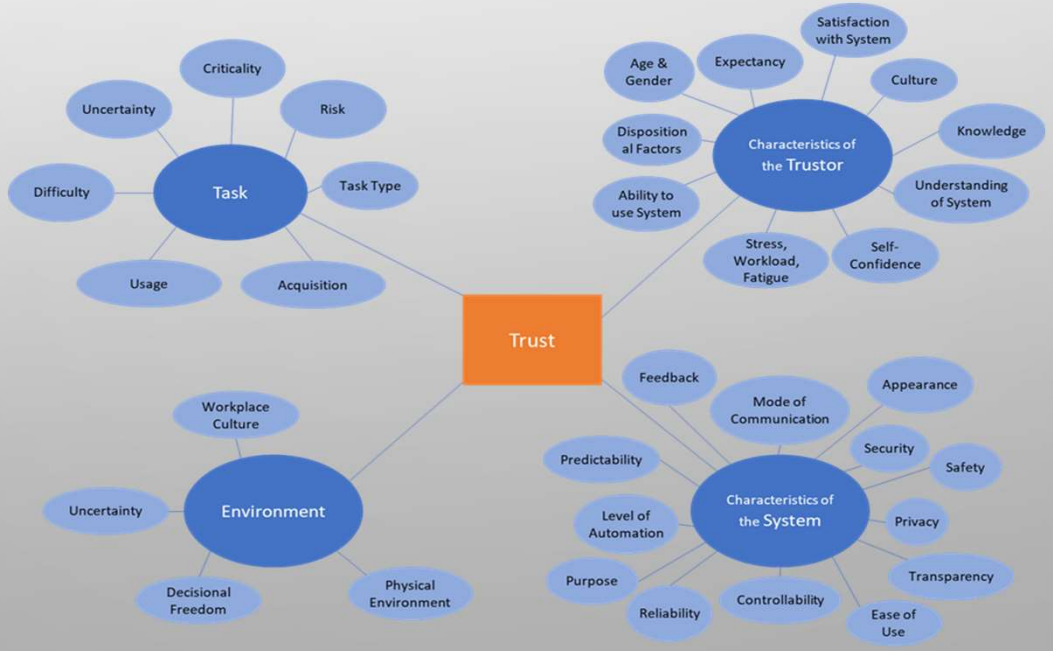
- Building trusted systems is a complex issue of technical and non-technical factors

- There are plenty theories and knowledge about trust
  - How do we translate this into system design?
  - How can we create generalizable lessons learned?

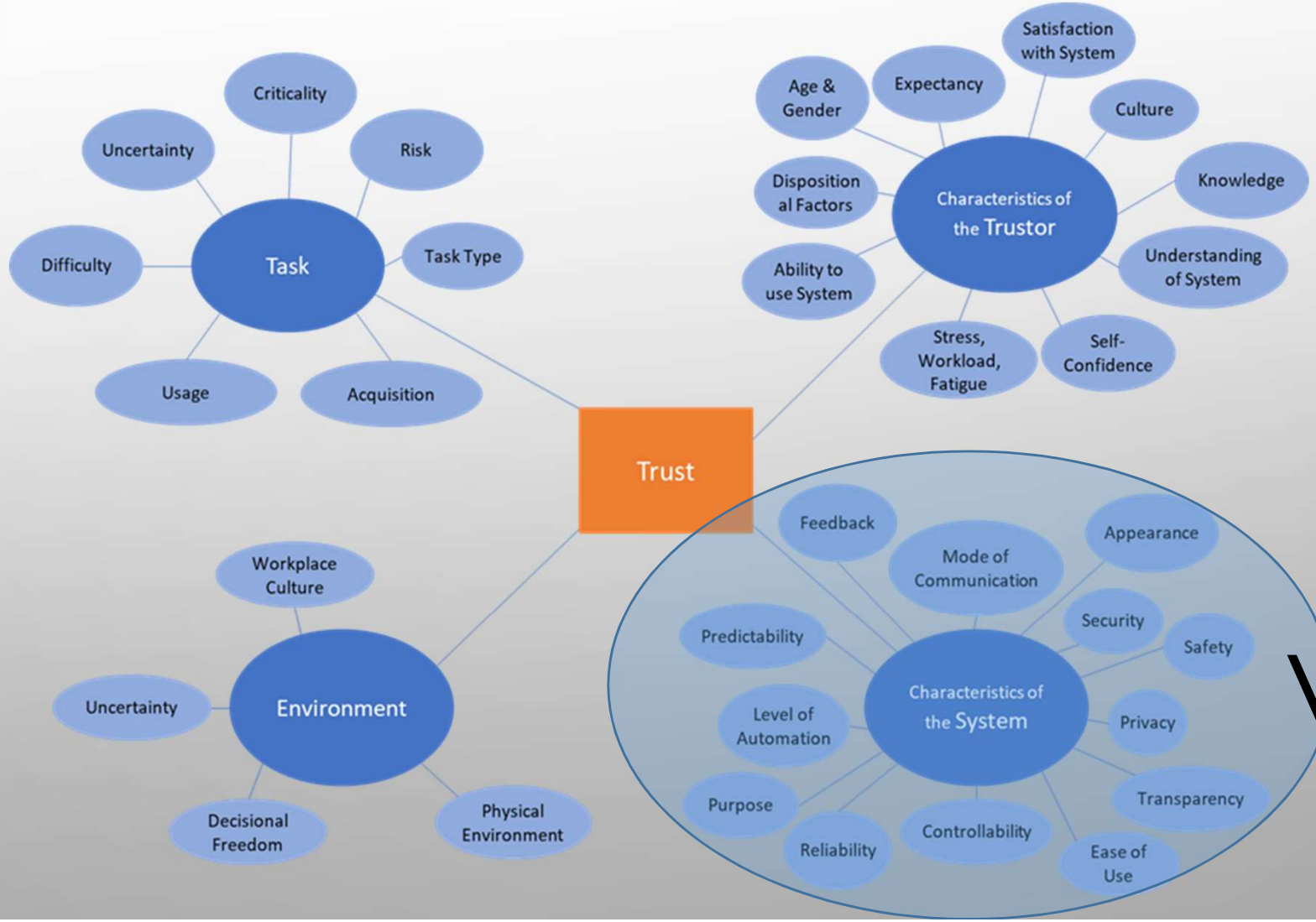
- The central concept in our framework are „trust issues“ → specific, contextualized concerns that a system does not meet the trustors goals



- Trust entities are the main actors which form a trust system: trustor, system (trustee), task & environment
- Trust entities have specific characteristics, which influence the development of trust

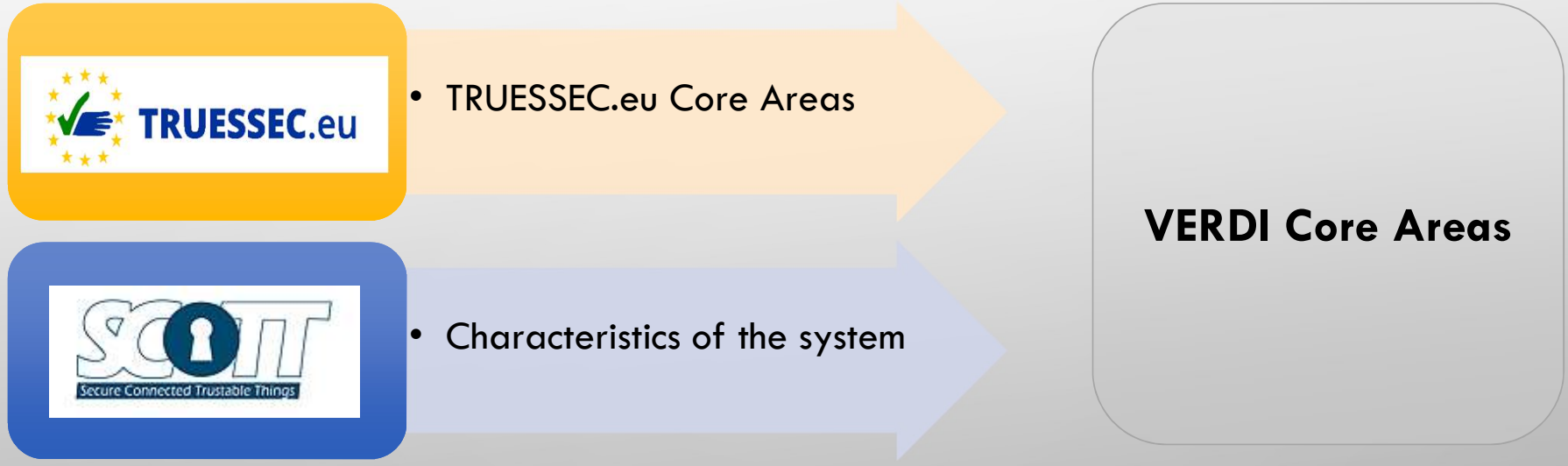


MERGING TRUESSEC.EU AND SCOTT → VERDI

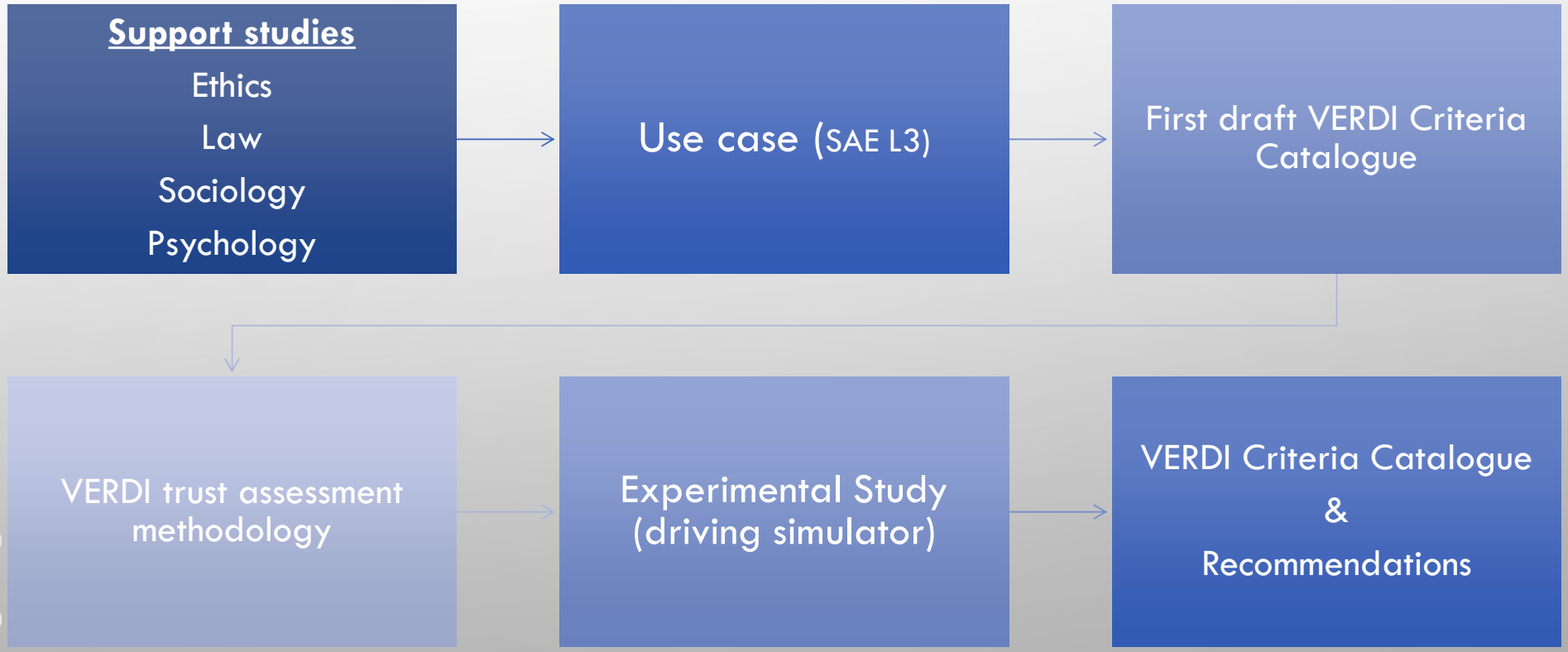


VERDI

MERGING TRUESSEC.EU AND SCOTT → VERDI



OUR APPROACH:
HOW DO WE DEFINE THE VERDI CORE AREAS AND THE VERDI CRITERIA CATALOGUE?



Example: Core Area Transparency

Ethics	Law	Sociology	Psychology	VERDI Core Areas SAE L3
Transparency in the context of SAE L3 relates to the following two aspects: o providing clear and sufficient information to the user about system functionality o providing information to users regarding activities with their personal data collected and processed in the context of SAE L3.	**Transparency** as in information duties laid down in the GDPR, the Directive on consumer rights or the e-commerce Directive.	90% of users are expressing at least some concern regarding the liability of self-driving vehicles. AVs need to **transparently** justify their actions and what data they are based upon, therefore, providing insight into the decision-making without overwhelming the user with unnecessary complexity. AVs need to clearly define and communicate their operational status and the abilities and restrictions of their ODD (Operational Driving Domain) to every actor involved (driver, passengers, other road users) within an easy to use interface. This also means to provide users with statistics about their own reliability.	**Transparency** can be one important feature for increasing trustworthiness but needs to be seen depending on the context of use. Transparency means the user understands the functionality of the system: what the system is doing, why is it doing that and is able to anticipate what the system is doing next. Transparency is often produced through feedback (amount, type of feedback, timing of feedback) and related to predictability, reliability and level of automation. Within Level 3 automation, transparency covers several aspects, e.g. transparency about current mode of operation and the specific meaning (Level 2 vs. Level 3 – what is the driver allowed to do? What means Level 2 and what means Level 3?). Transparency also addresses the general understanding of the system functionality and limitations and relates directly to the development of a "mental model" about the system, informing also, in as how far the driver can rely on the system in various contexts – addressing the topic of calibrated trust.	➔ **Transparency** The SAE L3 is provided in line with information duties regarding • the understanding of the system SAE L3 i.e. its functionality (What is it doing/can do?), limitations (What is it not doing/cannot do?) and anticipation (What will it do next?) • dealing with the (personal) data collected and processed for the purposes of L3. The above understanding of transparency is predominantly user-centered. Nevertheless, it can be argued that transparency can also be relevant for producers/manufacturers as well as for the society as a whole for the purpose of making system improvements.

Example: Core Area Privacy

Ethics	Law	Sociology	Psychology	VERDI Core Areas SAE L3
Privacy stands for the individual's claim to control the access to and use of one's personal information. The idea behind it is that people have the claim to determine who knows what about them thus preventing unjustified interferences by others.	**Privacy** as preserving Respect for private life (Art 7 CFR) and the Protection of personal data (Art 8 CFR). This includes Directive 95/46/EC as well as the GDPR and Directive 2002/58/EC.	By relying strongly on different kinds of data being collected for prediction of both road conditions and user behaviour, the issue of **Privacy** and use of data arises. SAE L3 needs to ensure the protection of this data and must justify any usage, storage and sharing beyond immediate necessity and must especially take data collected from other road users into account. Studies show that people are not worried about data sharing with surrounding vehicles, vehicle developers and roadway organizations, they are, however, quite reluctant to share data with insurance companies or tax authorities (69% think that personal information is not kept secure by public authorities). Yet, highly automated vehicles must assume responsibility towards the entire society, meaning that manufacturers provide data on errors and accidents to the public/authorities.	**Privacy** defined within SCOTT as the right of an individual to control personal data. Privacy can either be ensured by letting the trustor choose what information should be gathered, or by finding out what information the trustor is ready to give and only collect the most important and acceptable ones. Or increasing trust by a transparent way of informing the trustor which data are collected from him/her, since transparency and feedback are important considerations for trust too. Although, an adequate way of informing the user about privacy aspects needs to be chosen in order to not overwhelm the user with too much information in situations where s/he may not have time to process it or cannot understand and therefore really decide.	➔ **Privacy** The SAE L3 respects the protection of personal data and allows users to control their data. Even though at Level 3 we do not talk about a 'fully connected car', it is a fact that the system will still need to collect certain data to be able to function, hence 'functionality/necessary data'. Moreover, the system might also collect data that go beyond the functionality purpose such as for personalization. In both cases, for the system to be privacy-friendly the user should have the possibility to control that data to the extent possible.

VERDI Core Areas – First draft	
Transparency	The SAE L3 is provided in line with information duties regarding • the understanding of the system SAE L3 i.e. its functionality (What is it doing/can do?), limitations (What is it not doing/cannot do?) and anticipation (What will it do next?) • dealing with the (personal) data collected and processed for the purposes of L3.
Privacy	The SAE L3 respects the protection of personal data and allows users to control their data.
Anti-discrimination	The SAE L3 does not include any discriminative practices and biases based on parameters that are not relevant for the functioning of the system.
Autonomy	The SAE L3 gives users the opportunity to make decisions and respects those decisions. The system also respects other parties'/persons' rights and freedoms.
Respect	The SAE L3 is provided in accordance with the legitimate expectations of the users in relation to the system functionality and reliability.
Protection	The SAE L3 is designed to guarantee the utmost possible protection from harms to users and to the surrounding.

THANK YOU!

<hristina.veljanova@uni-graz.at>
<norah.neuhuber@v2c2.at>