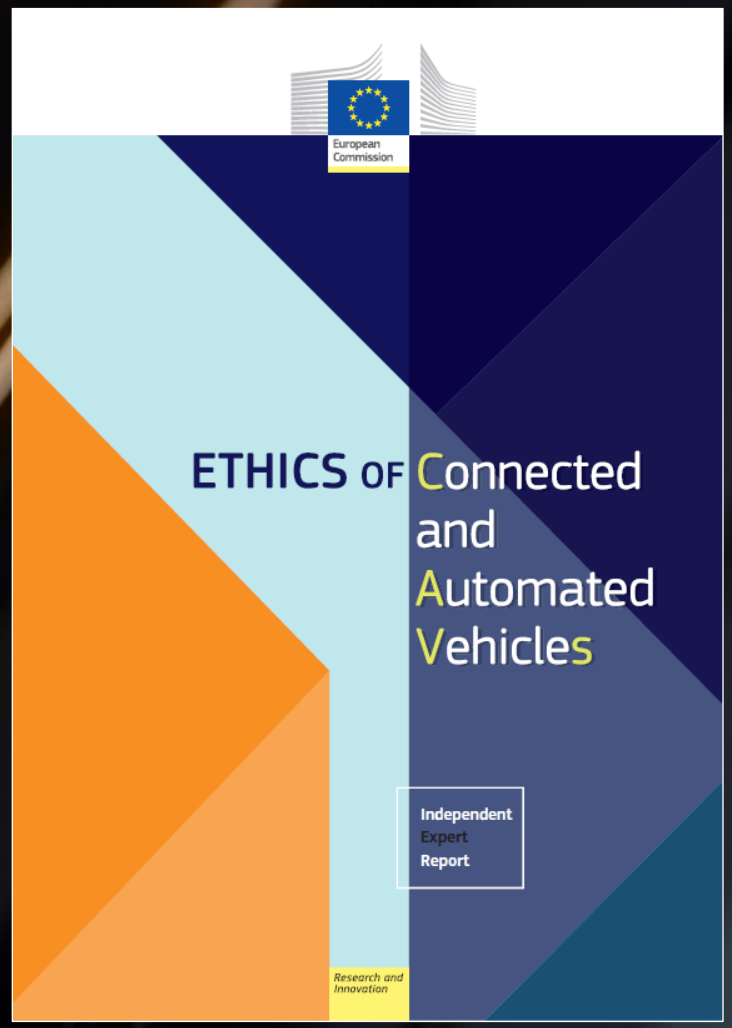
Engineering Ethics by Design: The Safety Imperative

Dr. Paula Palade
AI Ethics Senior Technical Specialist
JLR

Societal discussions influence the acceptance of new technology



Reference Frameworks

- German government Ethics Commission's 20 ethical rules on automated and connected driving – 2017
- **EU Report on Ethics of Connected and Automated Vehicles - Recommendations on road safety, privacy, fairness, explainability and responsibility – 2020**
- EGE statement on Artificial Intelligence – 2018
- AIHLEG Guidelines for Trustworthy AI – 2019
- Artificial Intelligence Act – April 2021
- IEEE Std 7000-2021 Standard Model Process for Addressing Ethical Concerns during System Design – Sept 2021
- Recommendation on the Ethics of Artificial Intelligence – UNESCO, Nov 2021
- ITU-T Focus Group on AI for Autonomous and Assisted Driving (FG-AI4AD
- **ISO/DIS 39003 Road traffic safety management systems – Guidance on safety ethical considerations for autonomous vehicles – August 2023**
- UK Law Commission Report 26 Jan 2022

Dr. Paula Palade – member of the Independent Expert Group that produced the EU Report in 2020 and of the ISO/DIS 39003 Guidance on safety ethical considerations for autonomous vehicles

European Commission

ETHICS of Connected and Automated Vehicles

Connected and Automated Vehicles (CAVs) have the potential to make transport:

SAFER GREENER MORE ACCESSIBLE

But new technologies do not just happen:
they are imagined by people and developed with purpose.
EU values need to be built-in at their core to ensure

- ☑ ETHICAL USE
- ☑ POSITIVE IMPACT
- ☑ ACCEPTANCE
- ☑ TRUST

Read the report and recommendations here
https://europa.eu/!VV67my

To tackle ethical challenges raised by CAVs, the European Commission formed an **INDEPENDENT EXPERT GROUP** to explore some important questions:

- How safe should CAVs be?
  Are pedestrians and cyclists more at risk with CAVs in traffic?
- Do you need to understand the technology behind it?
  What kind of data will a CAV share?
- Can the decisions of a CAV be trusted?
  Who is responsible for its behaviour?

20 RECOMMENDATIONS

are now available to support researchers, policymakers, manufacturers and deployers in the safe and responsible transition towards CAVs, with focus on:

ROAD SAFETY

DATA, ARTIFICIAL INTELLIGENCE AND ALGORITHMS

RESPONSIBILITY

Read the report and recommendations here
https://europa.eu/!VV67my

Publications Office of the European Union

Research and Innovation

Print: ISBN 978-92-76-18442-3, doi:10.2777/57288, KI-03-20-296-EN-C
WEB: ISBN 978-92-76-18441-6, doi:10.2777/402331, KI-03-20-296-EN-N

To tackle ethical issues, the Commission formed in 2019 an independent Expert Group to advise on specific ethical issues raised by driverless mobility. The Expert Group focused on three themes:

ROAD SAFETY, RISK, DILEMMAS:

- Safety benefits and improvements of CAVs should comply with basic ethical and legal principles: they should be publicly **demonstrable, monitored** and **updated** through **solid** and **shared scientific research,** and continuously adjusted to the needs of all road users.

DATA AND ALGORITHM ETHICS: PRIVACY, FAIRNESS, EXPLAINABILITY:

- Artificial Intelligence (AI) and automated systems used in CAVs should be **explainable** and **transparent** to empower users and to protect their data.
- This should be reflected through **rules and regulations** that take into account the fast-changing nature of CAV technologies (especially AI and big data) and favour **inclusive deliberation** at all levels.

RESPONSIBILITY:

- Responsibilities should be **clearly attributed** and **shared,** going beyond blame and compensation in case of a collision. No single person or system can be held solely accountable.
- From inception to use, best practices promoting **ethical responsibility** must be fostered and shared. This way, humans can remain **accountable to users,** instead of complex systems.

Research and Innovation

20 RECOMMENDATIONS are available to support researchers, policymakers, manufacturers and deployers in the safe and responsible transition towards CAVs.

Safety 6

1. Ensure that CAVs reduce physical harm to persons.
2. Prevent unsafe use by inherently safe design.
3. Define clear standards for responsible open road testing.
4. Consider revision of traffic rules to promote safety of CAVs and investigate exceptions to non-compliance with existing rules by CAVs.
5. Redress inequalities in vulnerability among road users.
6. Manage dilemmas by principles of risk distribution and shared ethical principles.

Transparency

Data and algorithm ethics 9

7. Safeguard informational privacy and informed consent.
8. Enable user choice, seek informed consent options and develop related best practice industry standards.
9. Develop measures to foster protection of individuals at group level.
10. Develop transparency strategies to inform users and pedestrians about data collection and associated rights.

Transparency

Data and algorithm ethics 9

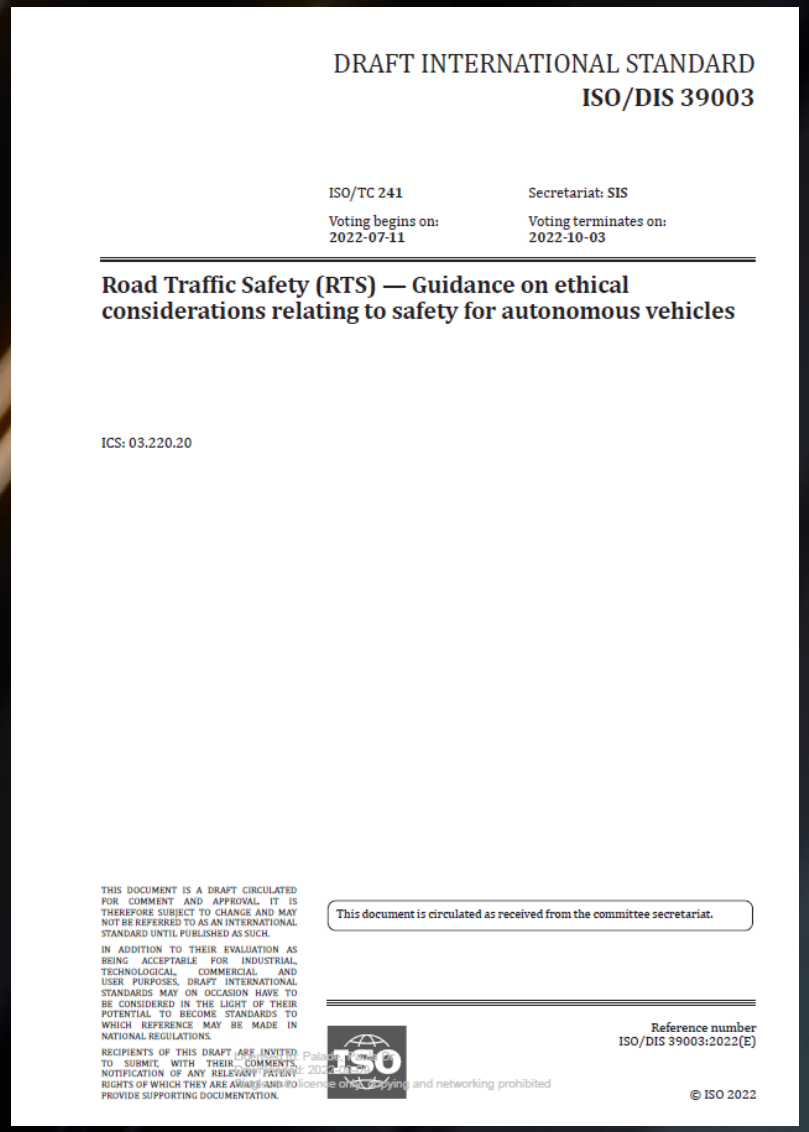
11. Prevent discriminatory differential service provision.
12. Audit CAV algorithms.
13. Identify and protect CAV relevant high-value datasets as public and open infrastructural resources.
14. Reduce opacity in algorithmic decisions.
15. Promote data, algorithmic, AI literacy and public participation.

Responsibility 5

16. Identify the obligations of different agents involved in CAVs.
17. Promote a culture of responsibility with respect to the obligations associated with CAVs.
18. Ensure accountability for the behaviour of CAVs (duty to explain).
19. Promote a fair system for the attribution of moral and legal culpability for the behaviour of CAVs.
20. Create fair and effective mechanisms for granting compensation to victims of crashes or other accidents involving CAVs.

Standards

ISO/DIS 39003



DRAFT INTERNATIONAL STANDARD
ISO/DIS 39003

ISO/TC 241 | Secretariat: SIS

Voting begins on: 2022-07-11 | Voting terminates on: 2022-10-03

Road Traffic Safety (RTS) — Guidance on ethical considerations relating to safety for autonomous vehicles

ICS: 03.220.20

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENT AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.

This document is circulated as received from the committee secretariat.

Reference number
ISO/DIS 39003:2022(E)

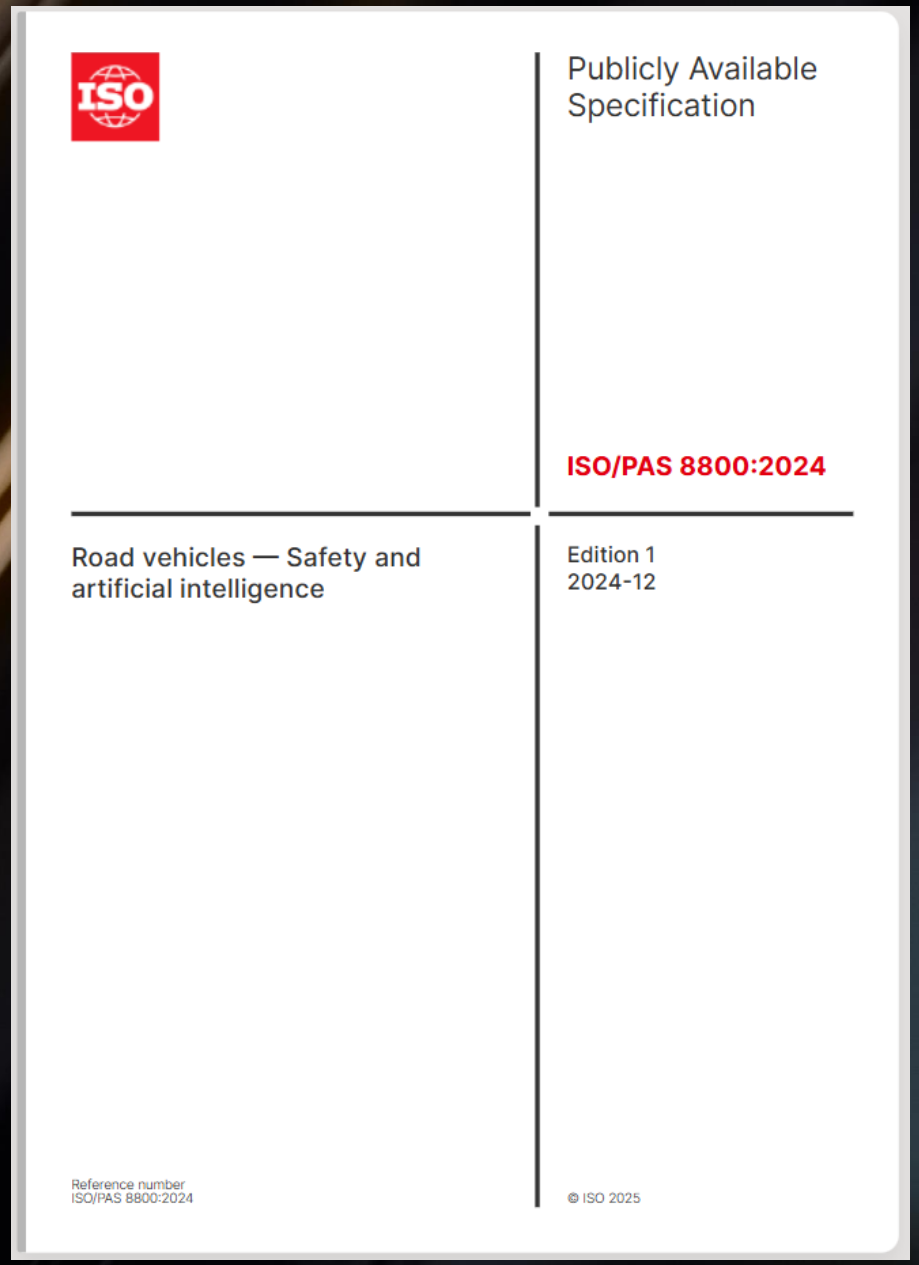
Road Traffic Safety (RTS) — Guidance on ethical considerations relating to safety for autonomous vehicles

- ✓ Applicable to all Level 5 autonomous vehicles SAE J3016
- ⊘ Is not a management system standard or technical standard
- ⊘ It does not set requirements for the outcomes of ethical decisions, nor does it offer guidance on methodology
- ⊘ It does not offer the technical precision to prescribe the required controls
- ✓ It offer a set of “protocol guidelines” that a vehicle manufacturer could choose to self-certify against to assure that the desired necessary ethical considerations were addressed during design and effectively controlled
- ✓ Provides manufacturers and distributers of the vehicles a mechanism to enable them to give **formal declaration of compliance** to an International Standard
- ✓ Give **assurance** to purchasers, end-users and society as a whole, that the vehicles’ design has considered and addressed the ethical issues identified within the standard

Formal declaration of compliance to an ISO that the vehicles’ design has considered and addressed the ethical issues identified within the standard.

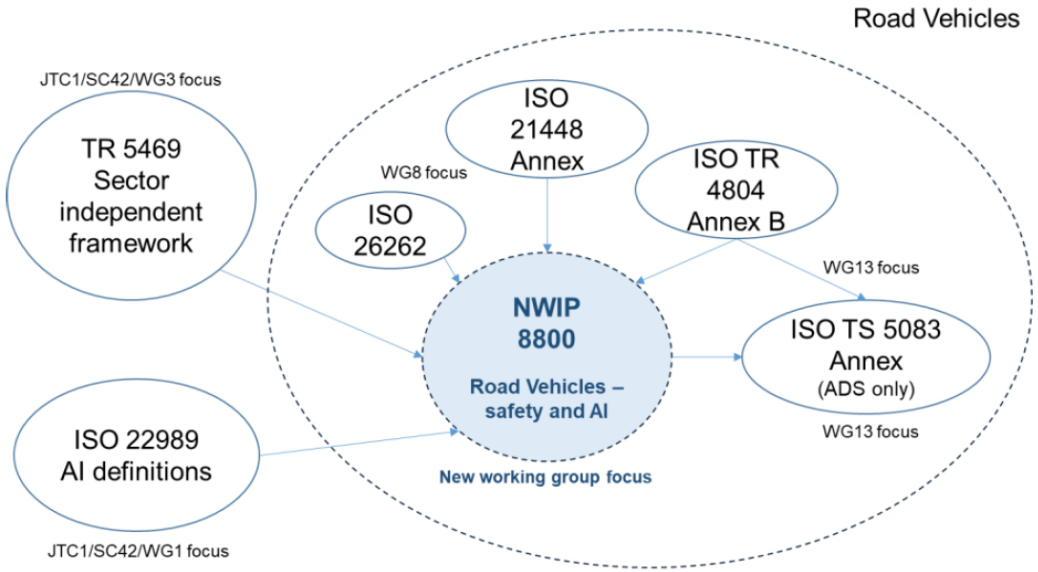
Ethical Assessment

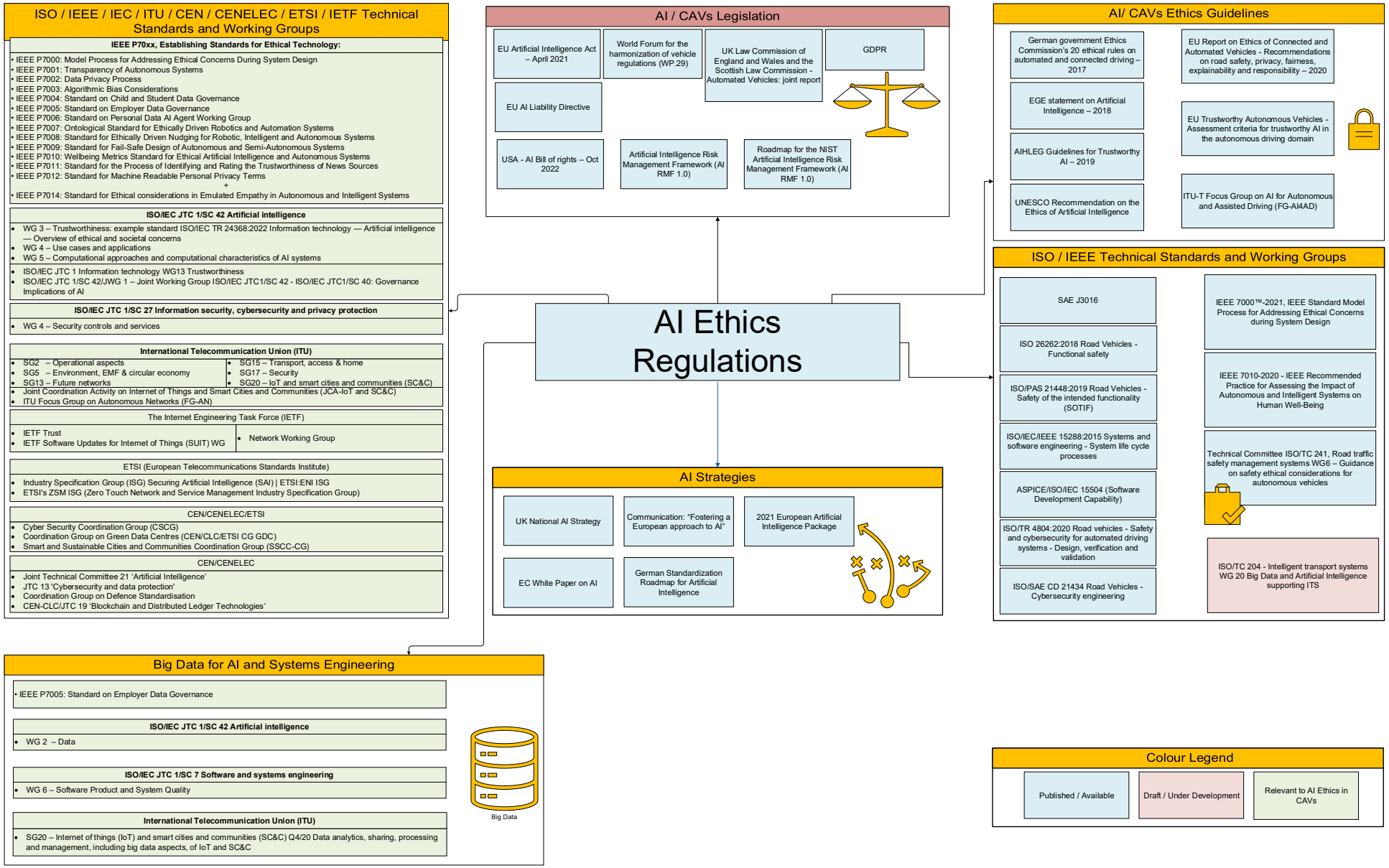
- Higher Organisational Level
- Development Organisation Level
- Specific Development and Implementation Processes
- Ethical Validation and Verification Activities



ISO/PAS 8800
Road vehicles — Safety and artificial intelligence

- Industry-specific guidance on safety-related AI/ML functions
- Define suitable safety principles, methods and evidence fulfilling objectives with ISO 26262 (functional safety) and ISO 21448 (safety of the intended functionality)
- Harmonize concepts already described in Annexes of ISO/TR 4804 and ISO 21448* Build upon generic guidance from ISO/EC TR 5469* Scope is road vehicles, not restricted to automated driving functions or specific ML techniques




ISO / IEEE / IEC / ITU / CEN / CENELEC / ETSI / IETF Technical Standards and Working Groups
IEEE P70xx, Establishing Standards for Ethical Technology:
• IEEE P7000: Model Process for Addressing Ethical Concerns During System Design
• IEEE P7001: Transparency of Autonomous Systems
• IEEE P7002: Data Privacy Process
• IEEE P7003: Algorithmic Bias Considerations
• IEEE P7004: Standard on Child and Student Data Governance
• IEEE P7005: Standard on Employer Data Governance
• IEEE P7006: Standard on Personal Data AI Agent Working Group
• IEEE P7007: Ontological Standard for Ethically Driven Robotics and Automation Systems
• IEEE P7008: Standard for Ethically Driven Nudging for Robotic, Intelligent and Autonomous Systems
• IEEE P7009: Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems
• IEEE P7010: Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems
• IEEE P7011: Standard for the Process of Identifying and Rating the Trustworthiness of News Sources
• IEEE P7012: Standard for Machine Readable Personal Privacy Terms
+
• IEEE P7014: Standard for Ethical considerations in Emulated Empathy in Autonomous and Intelligent Systems
ISO/IEC JTC 1/SC 42 Artificial intelligence
• WG 3 – Trustworthiness: example standard ISO/IEC TR 24368:2022 Information technology — Artificial intelligence — Overview of ethical and societal concerns
• WG 4 – Use cases and applications
• WG 5 – Computational approaches and computational characteristics of AI systems
• ISO/IEC JTC 1 Information technology WG13 Trustworthiness
• ISO/IEC JTC 1/SC 42/JWG 1 – Joint Working Group ISO/IEC JTC1/SC 42 - ISO/IEC JTC1/SC 40: Governance Implications of AI
ISO/IEC JTC 1/SC 27 Information security, cybersecurity and privacy protection
• WG 4 – Security controls and services
International Telecommunication Union (ITU)
• SG2 – Operational aspects
• SG5 – Environment, EMF & circular economy
• SG13 – Future networks
• SG15 – Transport, access & home
• SG17 – Security
• SG20 – IoT and smart cities and communities (SC&C)
• Joint Coordination Activity on Internet of Things and Smart Cities and Communities (JCA-IoT and SC&C)
• ITU Focus Group on Autonomous Networks (FG-AN)
The Internet Engineering Task Force (IETF)
• IETF Trust
• IETF Software Updates for Internet of Things (SUIT) WG
• Network Working Group
ETSI (European Telecommunications Standards Institute)
• Industry Specification Group (ISG) Securing Artificial Intelligence (SAI) | ETSI:ENI ISG
• ETSI's ZSM ISG (Zero Touch Network and Service Management Industry Specification Group)
CEN/CENELEC/ETSI
• Cyber Security Coordination Group (CSCG)
• Coordination Group on Green Data Centres (CEN/CLC/ETSI CG GDC)
• Smart and Sustainable Cities and Communities Coordination Group (SSCC-CG)
CEN/CENELEC
• Joint Technical Committee 21 'Artificial Intelligence'
• JTC 13 'Cybersecurity and data protection'
• Coordination Group on Defence Standardisation
• CEN-CLC/JTC 19 'Blockchain and Distributed Ledger Technologies'
AI / CAVs Legislation
EU Artificial Intelligence Act – April 2021
World Forum for the harmonization of vehicle regulations (WP.29)
UK Law Commission of England and Wales and the Scottish Law Commission - Automated Vehicles: joint report
GDPR
EU AI Liability Directive
USA - AI Bill of rights – Oct 2022
Artificial Intelligence Risk Management Framework (AI RMF 1.0)
Roadmap for the NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0)
AI Ethics Regulations
AI Strategies
UK National AI Strategy
Communication: "Fostering a European approach to AI"
2021 European Artificial Intelligence Package
EC White Paper on AI
German Standardization Roadmap for Artificial Intelligence
AI/ CAVs Ethics Guidelines
German government Ethics Commission's 20 ethical rules on automated and connected driving – 2017
EU Report on Ethics of Connected and Automated Vehicles - Recommendations on road safety, privacy, fairness, explainability and responsibility – 2020
EGE statement on Artificial Intelligence – 2018
EU Trustworthy Autonomous Vehicles - Assessment criteria for trustworthy AI in the autonomous driving domain
AIHLEG Guidelines for Trustworthy AI – 2019
UNESCO Recommendation on the Ethics of Artificial Intelligence
ITU-T Focus Group on AI for Autonomous and Assisted Driving (FG-AI4AD)
ISO / IEEE Technical Standards and Working Groups
SAE J3016
IEEE 7000™-2021, IEEE Standard Model Process for Addressing Ethical Concerns during System Design
ISO 26262:2018 Road Vehicles - Functional safety
IEEE 7010-2020 - IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being
ISO/PAS 21448:2019 Road Vehicles - Safety of the intended functionality (SOTIF)
ISO/IEC/IEEE 15288:2015 Systems and software engineering - System life cycle processes
Technical Committee ISO/TC 241, Road traffic safety management systems WG6 – Guidance on safety ethical considerations for autonomous vehicles
ASPICE/ISO/IEC 15504 (Software Development Capability)
ISO/TR 4804:2020 Road vehicles - Safety and cybersecurity for automated driving systems - Design, verification and validation
ISO/SAE CD 21434 Road Vehicles - Cybersecurity engineering
ISO/TC 204 - Intelligent transport systems WG 20 Big Data and Artificial Intelligence supporting ITS
Big Data for AI and Systems Engineering
• IEEE P7005: Standard on Employer Data Governance
ISO/IEC JTC 1/SC 42 Artificial intelligence
• WG 2 – Data
ISO/IEC JTC 1/SC 7 Software and systems engineering
• WG 6 – Software Product and System Quality
International Telecommunication Union (ITU)
• SG20 – Internet of things (IoT) and smart cities and communities (SC&C) Q4/20 Data analytics, sharing, processing and management, including big data aspects, of IoT and SC&C
Big Data
Colour Legend
Published / Available
Draft / Under Development
Relevant to AI Ethics in CAVs

How can Ethics be integrated in Engineering? Ethics-by-Design

Creating a Safe AI Framework – Ethics By Design

Societal values

From Abstract concepts
&
Qualitative KPIs

INCOSE SYSTEMS engineering

To Engineering Design Framework
&
Quantitative KPIs

To Reimagined Quality
&
Societal Value Products

Ethics captured in System Design Requirements to guarantee product integrity and compliance

THANK YOU

Dr Paula Palade

AI Ethics & Standardisation Senior Technical Specialist

ppalade1@jaguarlandrover.com