AUTONOMOUS DRIVING
& RISK ETHICS

Chapter 4 — Technology, Safety & Ethical Challenges

Definition & Levels | **Why Autonomous?** | **Use Cases** | **Human-Machine Interface** | **Ethics & Risk**

AGENDA

INTRODUCTION & HISTORY

A 140-year journey: from horseless carriages to robot taxis

1886

First Automobile

Benz Patent-Motorwagen — the world's first true petrol-powered car.

1970s–80s

Passive Safety

Seatbelts, crumple zones, airbags — protecting occupants after a crash.

1990s–2000s

Active Safety

ABS (1978), ESP (1995), ACC — systems that prevent accidents.

2010s

Driver Assistance

Lane keeping, emergency braking, blind-spot detection — cars get smarter.

Today

Autonomous Driving

Level 3–4 pilots enter production vehicles; Level 5 under research.

DEFINITION OF AUTONOMY

What does "autonomous" actually mean in the context of driving?

Autonomy = self-determination within an overarching (moral) law.

In autonomous driving: the human sets this "moral law" by programming the vehicle's behaviour — the car then learns to make driving decisions within that framework.

Machine decides

The vehicle independently assesses situations and acts — without real-time human input.

Human programs

Developers and ethicists define rules, priorities and edge-case behaviours in advance.

Learns over time

Modern AVs use machine learning to improve from millions of kilometres of driving data.

Bounded by law

All autonomous decisions remain constrained by traffic law and programmed ethical guidelines.

SAE AUTOMATION LEVELS (0–5)

L0

Driver Only

Driver: 100% | System: 0%

No automation. Driver fully controls vehicle at all times.

L1

Assisted

Driver: ~90% | System: 10%

System assists with either steering OR acceleration/braking.

L2

Partial Auto

Driver: ~60% | System: 40%

System controls both steering & speed; driver must supervise.

L3

Conditional

Driver: ~30% | System: 70%

System drives; driver may disengage but must be ready to take over.

L4

High Automation

Driver: ~5% | System: 95%

Fully automated in defined scenarios. No takeover required.

L5

Fully Driverless

Driver: 0% | System: 100%

No human driver needed. Operates everywhere, under all conditions.

WHY AUTONOMOUS DRIVING? — THE SAFETY CASE

Road safety is the primary — and overwhelming — motivation for developing autonomous vehicles.

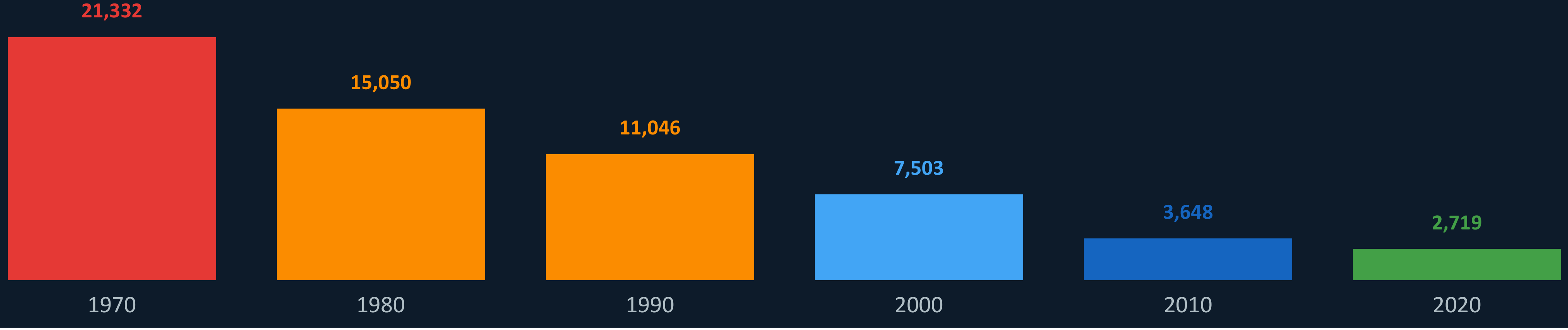
2.2M+
Accidents per year in Germany (2020)

94.4%
Caused entirely by human error

~21,000
Road deaths in Germany in 1970

~2,700
Road deaths in Germany in 2020

Deaths declining — technology is working



FACTORS DRIVING THE SAFETY IMPROVEMENT

A combination of legislation, passive safety features, and active digital systems has reduced road deaths by ~87% since 1970.

📋 Legal Regulations

- Seatbelt law (1976)
- Blood-alcohol limit 0.5‰ (2001)
- 100 km/h on country roads
- Mobile phone ban while driving (2001)
- Emergency call system (Euro NCAP)

🛡️ Passive Safety

- Front airbag (1980)
- Side airbag (1995)
- Curtain/head airbag (1997)
- Crumple zones & rigid passenger cell
- Child seat obligation (1993)

⚡ Active/Digital Safety

- ABS — Anti-lock Braking (1978)
- ESP — Electronic Stability (1995)
- Emergency Braking Assistant (AEB)
- Lane Departure Warning
- Adaptive Cruise Control (ACC)

→ ***Next step: remove human error entirely through full automation***

USE CASES — FOUR SCENARIOS

Autonomous driving is not one thing — it covers a spectrum of real-world applications.

1 Highway Autopilot

SAE Level 3 — Availability Driver

Operates exclusively on motorways / expressways. Driver is still present and legally responsible but may take hands off wheel and feet off pedals. Must be able to re-engage when prompted.

→ **Driver can take back control**

2 Full Automation (Extended)

SAE Level 4 — Approved Zones Only

Works beyond motorways in any pre-approved geographic zone. System handles all scenarios. Excluded zones: construction sites, areas with high pedestrian density, or unmapped roads.

→ **Driver can take back control**

3 Autonomous Valet Parking

SAE Level 4 — Parking Facility

Passengers exit the vehicle; the car parks itself without any human inside. Can relocate on command and return to a chosen pick-up spot — like a hotel valet, but fully robotic.

→ **No driver inside the vehicle**

4 Vehicle-on-Demand

SAE Level 5 — All Scenarios

Fully autonomous in every environment — with passengers, cargo or empty. Passengers can only input a destination or trigger the Safe-Exit protocol; they cannot take over driving.

→ **Passengers cannot drive**

USE CASES — TWO CATEGORIES COMPARED

The four use cases fall into two fundamentally different categories of human involvement.

Category A — Human in the Loop

Use Cases 1 & 2
Highway Autopilot + Full Automation

✓ Driver remains physically present
✓ Can take back control at any time
✓ Must intervene in critical situations
✓ Legal responsibility remains with driver

Suitable for: commuters, highway journeys, partially mapped areas

VS

Category B — Fully Autonomous

Use Cases 3 & 4
Valet Parking + Vehicle-on-Demand

✓ No driver needed
✓ Passengers are purely passengers
✓ Only input: destination / Safe-Exit
✓ Vehicle decides everything else

Suitable for: robo-taxis, freight, mobility for elderly/disabled

VIDEO — How Do Self-Driving Cars Work?

"How do Self-Driving Cars Work?"

https://www.youtube.com/watch?v=PRg5RNU_JLk

LEARNING SYSTEMS IN AUTONOMOUS VEHICLES

How AI and Machine Learning power autonomous driving decisions

Perception

Cameras, LiDAR, radar and GPS sensors feed real-time data about the vehicle's environment.

→

Processing

Neural networks classify objects (pedestrians, signs, lanes) and estimate their trajectory.

→

Planning

Path planning algorithms calculate the optimal route, speed and lane changes.

→

Actuation

Commands are sent to brakes, steering and throttle to execute the planned manoeuvre.

Machine Learning — Supervised Learning Approach

The AV's neural network is trained on millions of annotated scenarios — where both inputs (sensor data) and expected outputs (safe actions) are known. Parameters (weights & thresholds) are iteratively adjusted until the system correctly handles virtually every driving scenario. This is supervised machine learning — the same approach used in image recognition (e.g. classifying a cat).

CRITICISM & LIMITATIONS

Autonomous vehicles face significant technical, legal, and societal obstacles before mass deployment.

TECHNICAL

Sensor Limitations

LiDAR and cameras struggle in heavy rain, snow, fog, or direct sunlight — edge cases that humans handle intuitively.

TECHNICAL

Infrastructure Dependency

High-definition maps and V2X communication networks are required — but not yet globally available.

LEGAL

Unclear Liability

Who is responsible in an accident — manufacturer, software developer, vehicle owner, or road authority?

LEGAL

Regulatory Gaps

Most jurisdictions still lack clear legal frameworks for fully driverless vehicles on public roads.

SOCIETAL

Public Acceptance

Many people distrust autonomous systems — especially after high-profile accidents involving AV test vehicles.

SOCIETAL

Job Displacement

Professional drivers (taxi, truck, delivery) face structural unemployment as automation scales.

HUMAN–MACHINE INTERFACE (HMI)

How the driver's role transforms — and what this means for design and responsibility.

L0–L1

Active Driver

- Hands on wheel
- Feet on pedals
- Eyes on road
- Full legal duty

L2–L3

Supervising

- Monitor system
- Ready to intervene
- May look away briefly
- Legal duty remains

L3–L4

Passenger

- Can relax or work
- Must re-engage on request
- Can override if needed
- Shared responsibility

L5

Pure Passenger

- No wheel / pedals
- Set destination only
- Trigger Safe-Exit only
- No legal driving duty

← More Human Involvement *More Automation →*

HMI — KEY DESIGN CHALLENGES

As automation increases, designing the interface between human and machine becomes safety-critical.

Vigilance Decay

When systems handle routine driving, humans lose situational awareness and react more slowly when intervention is needed — the "automation complacency" problem.

Handover Problem

Transitioning control back to the human (L3 → human) in an emergency requires the driver to re-engage within seconds — a fundamental HMI design challenge.

Trust Calibration

Users must trust the system enough to disengage — but not so much that they stop monitoring. Achieving correct trust calibration requires sophisticated UI feedback.

Safe-Exit Protocol

If the system fails at L4/L5 and no driver is present, a Safe-Exit manoeuvre must safely bring the vehicle to a halt — the ultimate HMI safety net.

ETHICS & RISK — OVERVIEW

Autonomous systems raise fundamental ethical and legal questions that society has not yet resolved.

⚖️ Accountability Gap

Algorithms make life-or-death decisions. Who is legally and morally responsible — developer, manufacturer, or owner?

🔍 Transparency

AI decisions emerge from opaque correlations. Even creators cannot always explain why a specific decision was made.

🚋 Trolley Problem

In unavoidable accidents: protect the passenger or minimize total casualties? These are not purely technical questions.

🔒 Data Privacy

AVs continuously log position, speed, routes and behaviour. Who owns this data — and who can access it?

📢 Algorithmic Bias

AI trained on skewed data may perform differently for different demographic groups, raising fairness and discrimination concerns.

📋 Regulatory Need

Law must keep pace with technology. Adaptive, international regulatory frameworks are essential — and currently lagging.

VIDEO — The Ethical Dilemma of Self-Driving Cars

"The Ethical Dilemma of Self-Driving Cars"

youtube.com/watch?v=ixIoDYVfKA0

THE TROLLEY PROBLEM — PROGRAMMING ETHICS

When an accident is unavoidable, what should an autonomous vehicle prioritise?

> *Scenario: A vehicle's brakes fail. It can swerve left (hitting one elderly pedestrian) or continue straight (hitting three children). What should it do?*

Utilitarian View

Minimize total harm — the vehicle should always choose the option resulting in fewer casualties, regardless of who they are.

✓ Mathematically clear

✗ Treats lives as numbers

Rights-Based View

The vehicle must never actively harm an uninvolved party (pedestrians). It may only allow harm as a side effect of protecting its passengers.

✓ Respects individual rights

✗ May lead to more total harm

German Ethics Commission (2017)

Personal features (age, gender) must NEVER be used. Protecting the greatest number of lives is preferred, but programmers may not target specific individuals.

✓ Legal/ethical baseline

✗ Still leaves edge cases unresolved

Conclusion: There is no perfect answer — but societies must agree on a framework before AVs are widely deployed.

AI ETHICS PRINCIPLES FOR AUTONOMOUS DRIVING

Six core principles that must guide the design, deployment and regulation of autonomous AI systems.

Beneficence

AI must serve the common good — improving safety, mobility and quality of life for everyone.

Transparency

Decision logic must be auditable. Passengers, engineers and authorities deserve explainability.

Non-Maleficence

Systems must avoid harm. Technical safeguards and kill-switches must be built in from the start.

Autonomy

Human agency must be preserved. Where possible, humans should retain the ability to override the system.

Justice

No algorithmic discrimination. Diverse training data and independent audits are essential to ensure fairness.

Data Privacy

Personal movement data must be protected. Encryption, data minimisation and legal safeguards are non-negotiable.

REGULATION & GOVERNANCE CHALLENGES

Autonomous vehicles challenge existing legal structures — new frameworks are urgently needed.

Liability

- Who is liable when a Level 5 car injures someone?
- Vienna Convention (1968) requires driver in control — incompatible with L5
- Product liability law may shift responsibility to manufacturers

Certification

- How do you certify a system that learns and changes over time?
- Traditional type approval was designed for static, predictable systems
- New standards (ISO 26262, SOTIF) are emerging but not yet mature

Data & Privacy

- AVs generate terabytes of personal location data daily
- GDPR applies in Europe — but enforcement in vehicles is unclear
- Data sovereignty: who owns driving data — the user or the OEM?

International

- Regulations differ wildly: USA, EU, China, Japan, Germany
- Cross-border operation of AVs requires harmonised international law
- UN ECE WP.29 is developing global technical regulations

WEAK AI vs STRONG AI — A CRITICAL DISTINCTION

Where do today's autonomous vehicles sit — and where might the technology lead?

🤖 WEAK AI (Narrow AI)

Current state of all autonomous vehicles

• Operates within a specific, narrow domain
• Can surpass humans in that domain (e.g. reaction time)
• Cannot transfer skills to other tasks
• No genuine understanding or consciousness

Examples: Tesla Autopilot, Waymo Driver, Mercedes DRIVE PILOT

VS

🧠 STRONG AI (AGI)

Hypothetical — does not yet exist

• At least human-level intelligence across all domains
• Can learn and adapt to entirely new situations
• Would require entirely new ethical frameworks
• Questions of rights, personhood and accountability

Why it matters: ethical/legal debate must begin NOW — before it arrives.

→ *The transition from Weak to Strong AI represents a paradigm shift that society must proactively prepare for.*

VIDEO — Why Self-Driving Cars Have Stalled

"Why Self-Driving Cars Have Stalled"

https://www.youtube.com/watch?v=4sCK-a33Nkk&t=47s

OUTLOOK — THE ROAD AHEAD

Autonomous driving will arrive incrementally — with technology, law and ethics evolving in parallel.

Now — L2–L3 in production (Tesla, Mercedes)

2026–28 — L4 in geofenced urban areas (Waymo)

2030s — L4/L5 motorway automons scale

2040+ — Full L5 deployment requires legal clarity

🔬 Technology

- Robust all-weather sensors
- Redundant compute systems
- V2X communication infrastructure

📋 Regulation

- International liability framework
- Certification standards for learning AI
- Clear data sovereignty rules

🤝 Society

- Public education and trust-building
- Retraining programmes for displaced workers
- Inclusive design (elderly, disabled)

SUMMARY & CONCLUSION

1. Autonomous driving spans SAE Levels 0–5 — from zero automation to fully driverless. Today's vehicles are mostly at L2–L3.

2. The primary driver is safety: over 94% of accidents stem from human error. Autonomous systems could save thousands of lives annually.

3. Four key use cases — highway pilot, full automation, valet parking and vehicle-on-demand — each with different human control levels.

4. HMI design is critical: as automation grows, maintaining appropriate human vigilance and safe handover becomes a core engineering challenge.

5. Ethics is foundational — not optional. Accountability, transparency, fairness, privacy, and the trolley problem must all be addressed by design.

6. International regulation, public acceptance and social adaptation are equally important as the technology itself for successful deployment.

"Ethical frameworks for AI are not a brake on innovation — they are its foundation."