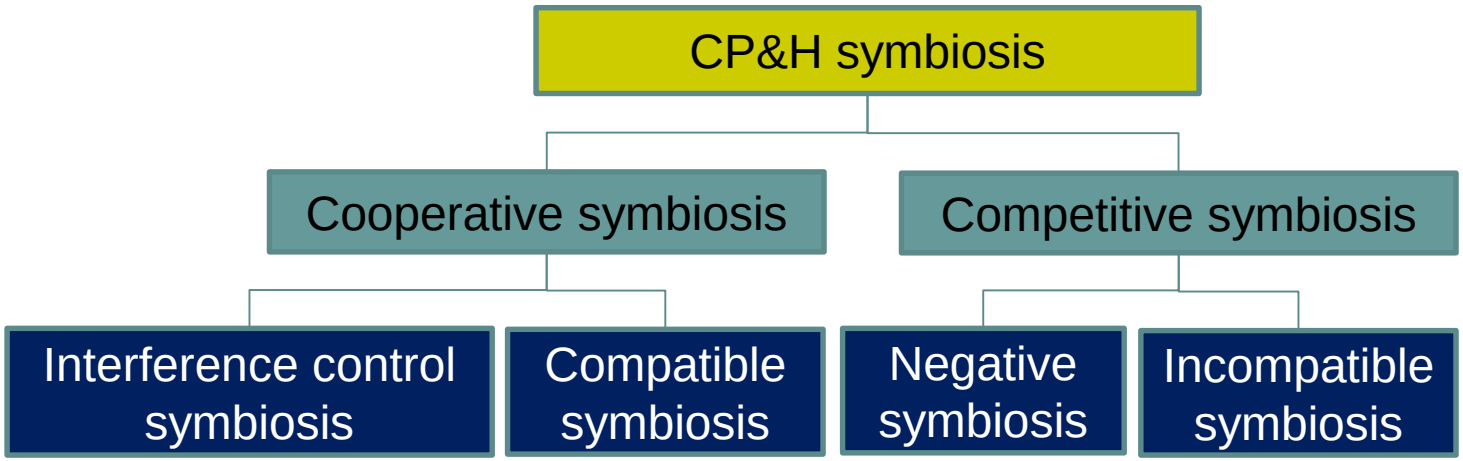
Responsible Cyber-Physical & Human Systems: dream or reality?

Frédéric VANDERHAEGEN
Université Polytechnique Hauts-de-France - INSA Hauts-de-France
Valenciennes, France

Introduction

What is Cyber-Physcal & Human System (CPHS) ?

➔ Inspired from human engineering view



Introduction

What is CPHS?

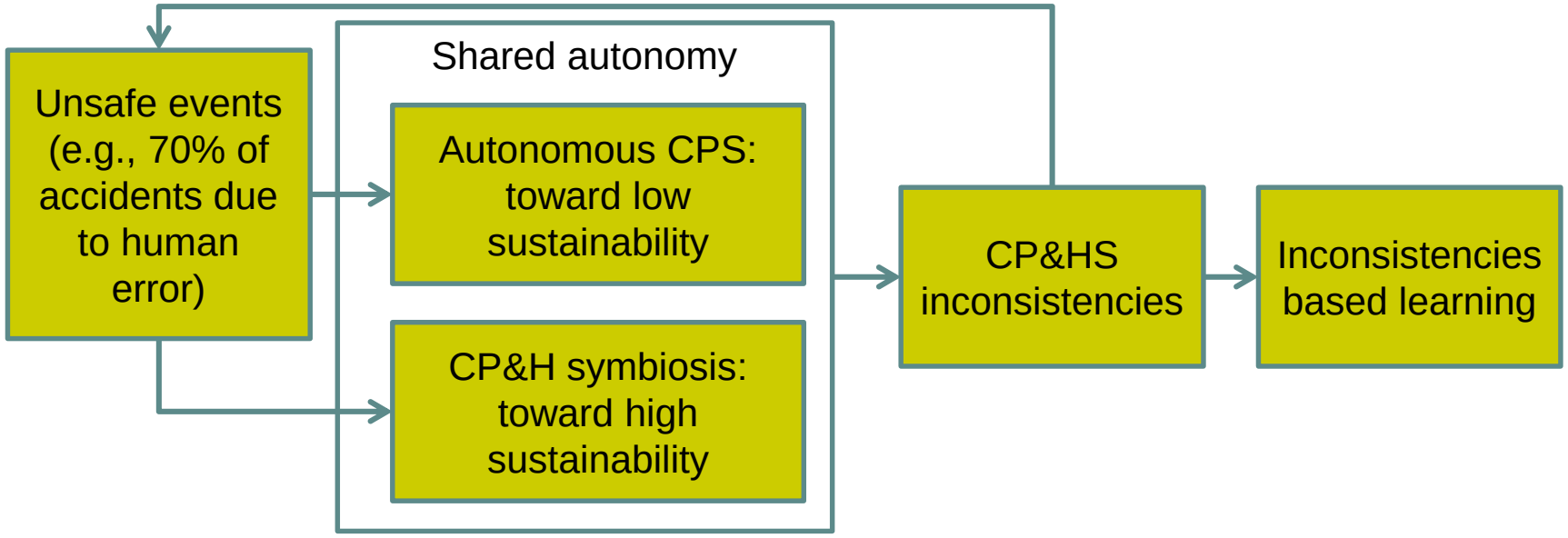
➔ Inspired from ecological / biological view

Paradigm		Principle
Cooperative symbiosis	Mutualism	Interaction that brings mutual benefits to each decision-maker without harming anyone
	Commensalism	Interaction where a decision-maker takes benefits from others without impacting them
Competitive symbiosis	Parasitism	Dependence interaction between decision-makers, one depends on others and inflicts negative impacts on them

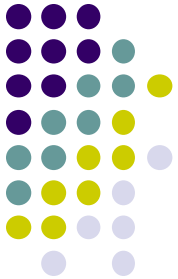
Introduction

What is CPHS?

➔ Inspired from sustainability view (seminar of GIS GRAISyHM / regional project TIFSID – 12-13 may 2026)
➔Concept of Responsible CPHS (RCPHS)



Content



- Responsible CPHS (RCPHS) inspired by Responsible Artificial Intelligence (RAI)
- RCPHS inspired by shared autonomy levels
- Examples of possible hazardous RCPHS
- Conclusion: toward "all-inclusive" RCPHS

RCPHS inspired by RAI

A German view: (VDE, 2022). VCIO based description of systems for AI trustworthiness characterization. VDE SPEC 90012 V1.0 report (2022).

➔ Toward an IEEE standard with contribution of Verband Deutscher Elektrotechniker, IRT SystemX, Positive AI)

A French view: (CE, 2019). Intelligence artificielle et action publique : construire la confiance, servir la performance. Rapport d'étude, Conseil d'Etat, https://www.vie-publique.fr/files/rapport/pdf/288900.pdf (2022).

The European Commission view: (AI-HLEG, 2019). Ethics guidelines for trustworthy AI. European Commission report, https://www.europarl.europa.eu/cmsdata/196377/AI HLEG_Ethics Guidelines for Trust-worthy AI.pdf

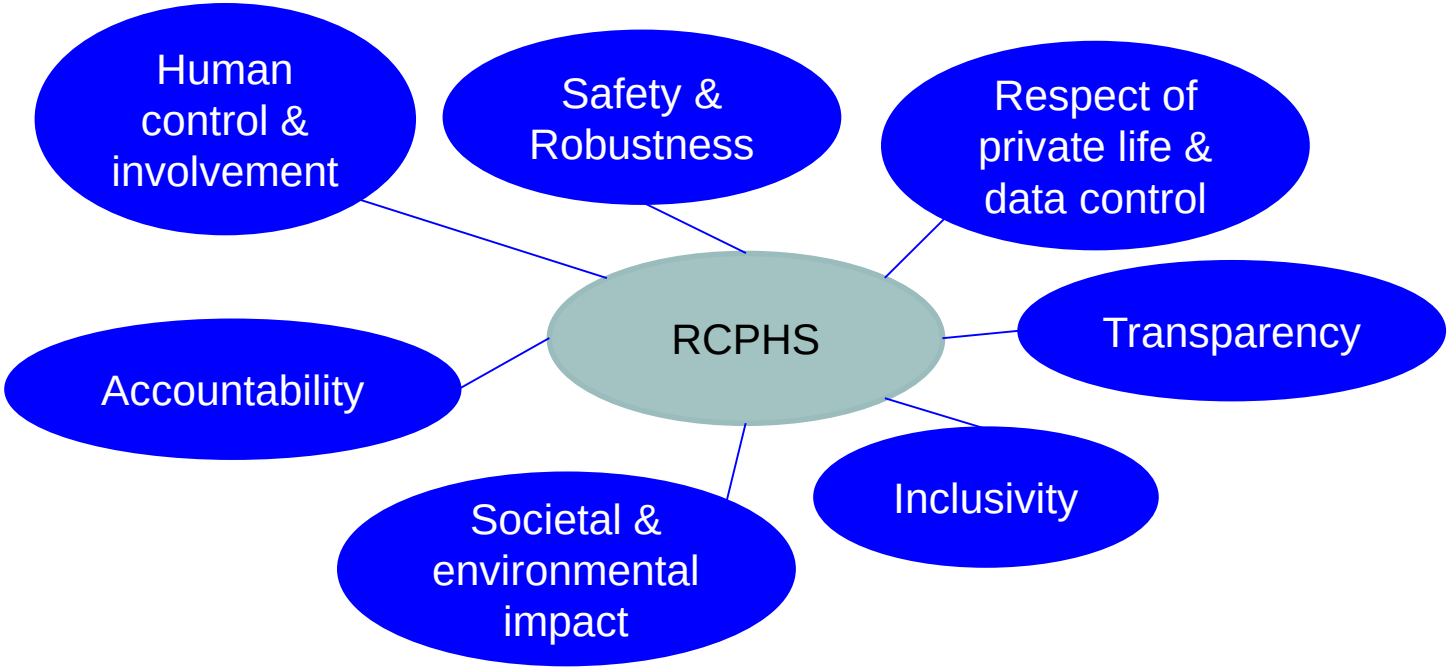
An American view: (NIST-AI, 2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). The National Institute of Standards and Technology, NIST report NIST AI 100-1, https://doi.org/10.6028/NIST.AI.100-1.

RCPHS inspired by RAI

VDE, 2022	CE, 2022
Transparency Accountability Privacy Fairness Reliability	Transparency Security Strategic autonomy Equity and non-discrimination Performance Human primacy Environmental sustainability

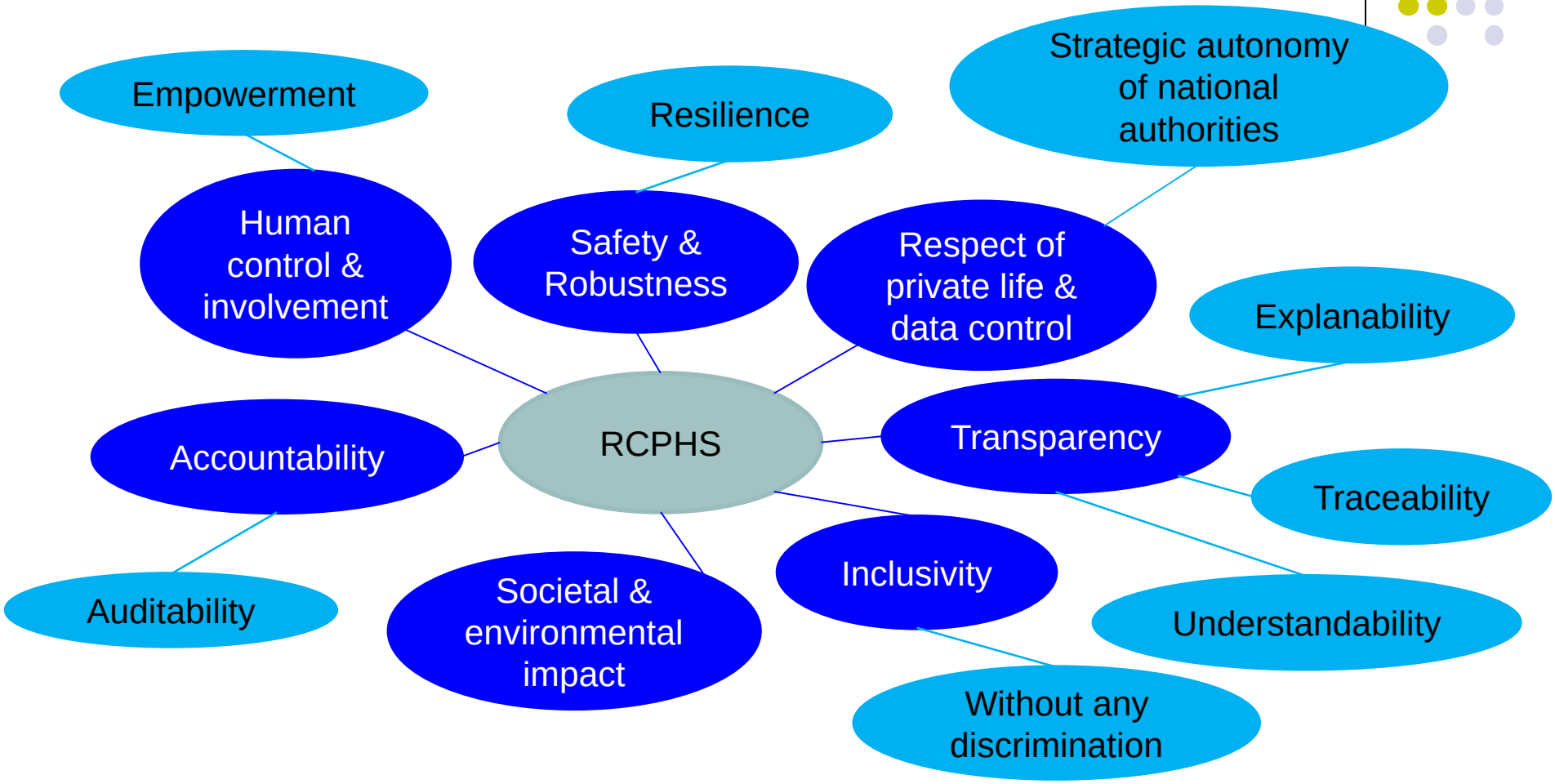
AI-HLEG, 2019	NIST, 2023
Transparency Accountability Privacy and data governance Diversity, non-discrimination and fairness Technical robustness and safety Human agency and oversight Sustainable societal and environmental wellbeing	Transparency and accountability Interpretability and explainability Privacy Fairness Validity and reliability Safety Security and resilience

RCPHS inspired by RAI

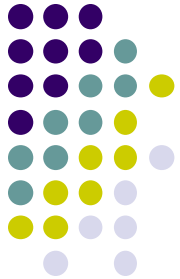


➔ Goal: make RCPHS-based decisions trustworthy

RCPHS inspired by RAI



RCPHS inspired by shared autonomy levels



Human-In-The-Loop ambiguity

Autonomous CPS ➔ no human intervention!

Autonomous CPS ➔ knowledge about human factors

Autonomous CPS ➔ assistance to human behavior
➔ human supported by CPS

Autonomous CPS ➔ learning from human behavior
➔ CPS supported by human

RCPHS inspired by shared autonomy levels

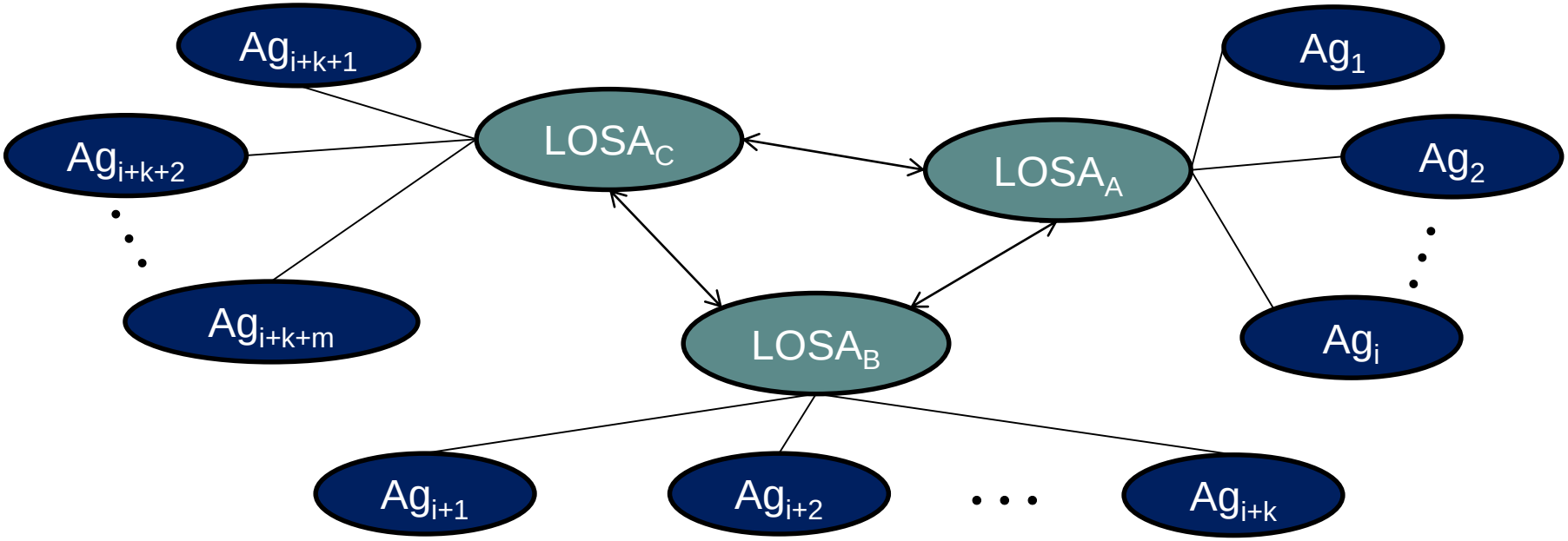
Level Of Computer/Robot/Vehicle/Train Autonomy (LOCA, LORA, LOVA, LOTA)

Shared autonomy principle	LOCA (Sheridan 1992; Inagaki, 2006)	LORA (Beer et al., 2014)	LOVA (SAE 2016; Inagaki, Sheridan, 2019)	LOTA (Powell et al., 2016; Brandenburger et al., 2021)
Humans control the process entirely: Human-Entirely-In-The-Loop	Level 1	Level 0	Level 0	Level 0
Group 1 of intermediate levels: Human-In-The-Loop	From Level 2 to Level 5	From Level 1 to Level 4	Level 1 and Level 2	Level 1 and Level 2
Group 2 of intermediate levels: Human-Back-To-The-Loop	From Level 6 to Level 9	From Level 5 to Level 8	Level 3 and Level 4	Level 3
Human-Out-Of-The-Loop	Level 10	Level 9	Level 5	Level 4

Adapted from:
Vanderhaegen, F. (2021). Heuristic-based method for conflict discovery of shared control between humans and autonomous systems - A driving automation case study. Robotics and Autonomous Systems, 146, 103867.

RCPHS inspired by shared autonomy levels

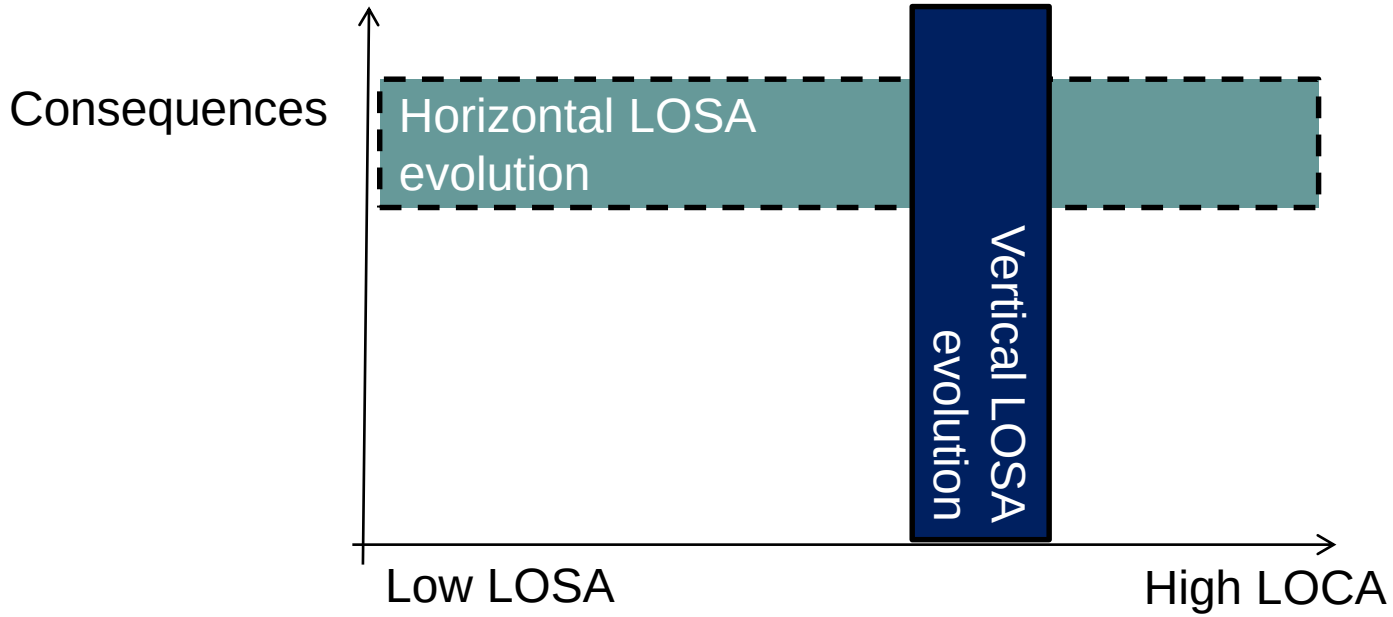
multi-Level Of Shared Autonomy (LOSA) based RCPHS



Adapted from:
Vanderhaegen, F. (2021). Heuristic-based method for conflict discovery of shared control between humans and autonomous systems - A driving automation case study. Robotics and Autonomous Systems, 146, 103867.

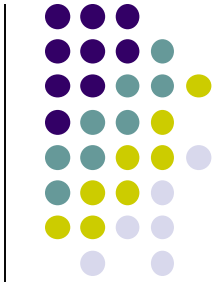
RCPHS inspired by shared autonomy levels

LOSA evolution ?



Adapted from:
Vanderhaegen, F. (2021). Heuristic-based method for conflict discovery of shared control between humans and autonomous systems - A driving automation case study. Robotics and Autonomous Systems, 146, 103867.

RCPHS inspired by shared autonomy levels



➔ Human needs for transparency

1. situation awareness ➔ explainability, traceability, auditability
2. sensemaking ➔ understandability

➔ Other features for RCPHS interaction

1. adaptability
2. usability
3. acceptability
4. learnability
5. memorability
6. educability

…

➔ Resilience

1. successful control of unprecedented scenarios
2. control of weak signal based RCPHS inconsistencies
3. learning process of these events

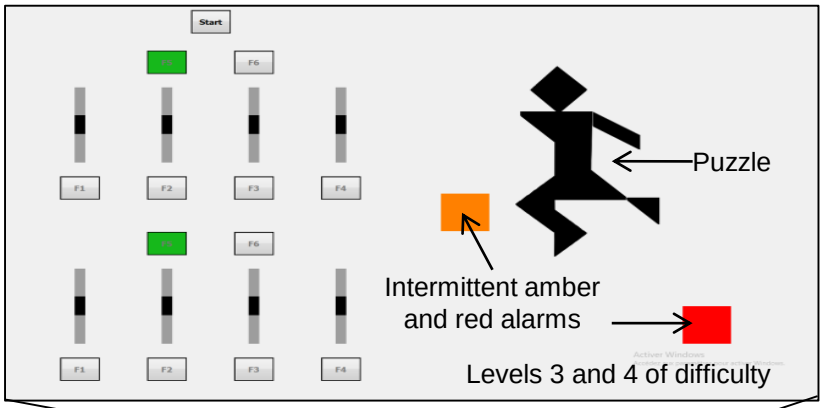
Examples of possible hazardous RCPHS

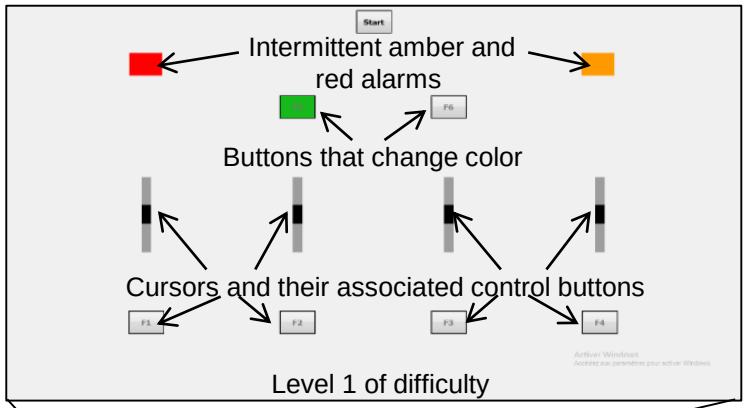
Shared autonomy principle	RCPHS example	Benefit (High probability of occurrence)	Inconsistency (Low probability of occurrence)
Group 1	Head-Up Display (HUD)	Reduction of workload by grouping information on a same field of vision	Degradation of perception ability when events occur outside the field of vision
	Auditive and visual alarms	Facilitation of danger detection and human reactivity	Degradation of perception ability when alarms occurrence is synchronized with heartbeat

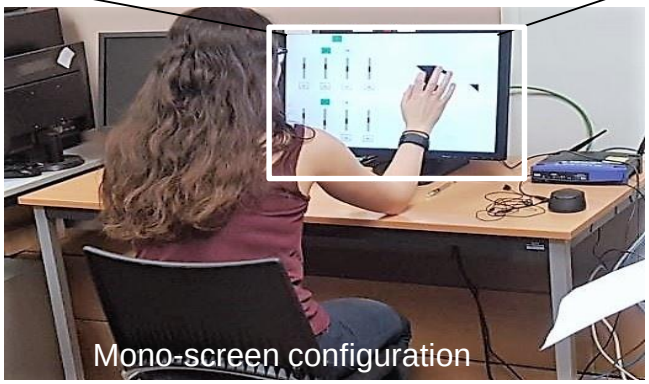
Adapted from:
Vanderhaegen, F. (2021). Heuristic-based method for conflict discovery of shared control between humans and autonomous systems - A driving automation case study. Robotics and Autonomous Systems, 146, 103867.

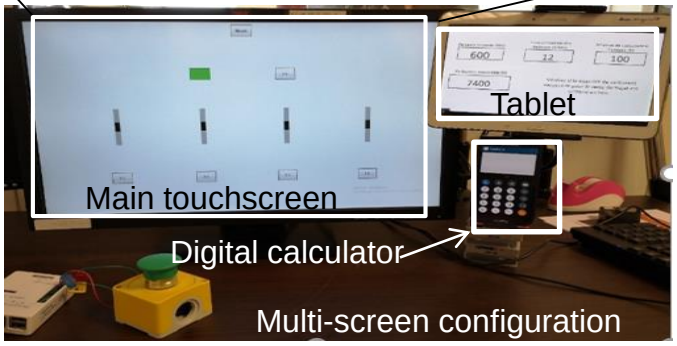
Examples of possible hazardous RCPHS

Experiments done with the NASA MultiAttribute Task Battery (MATB-II)



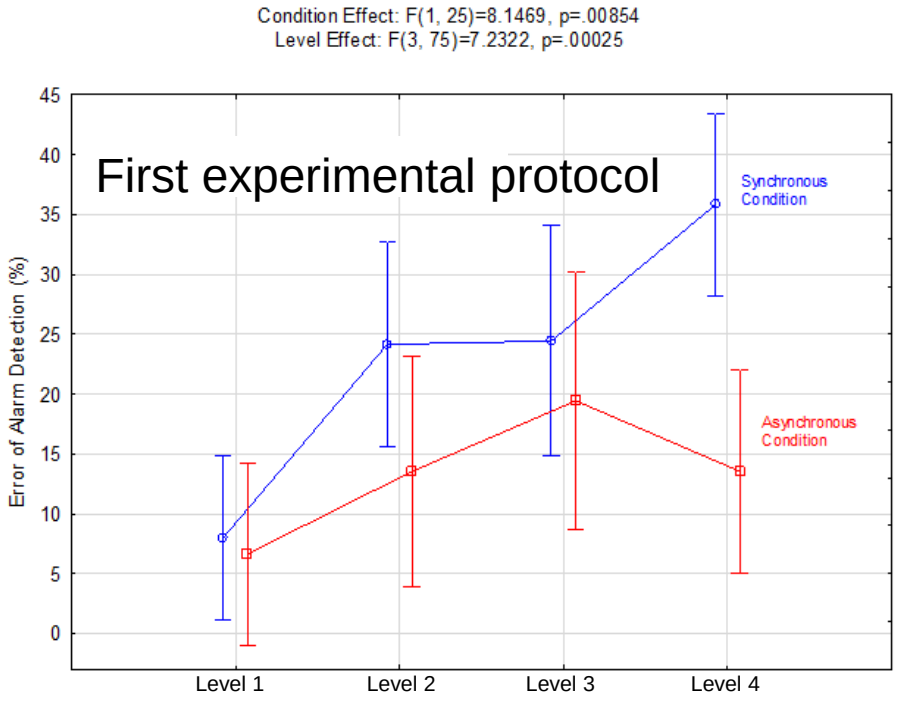


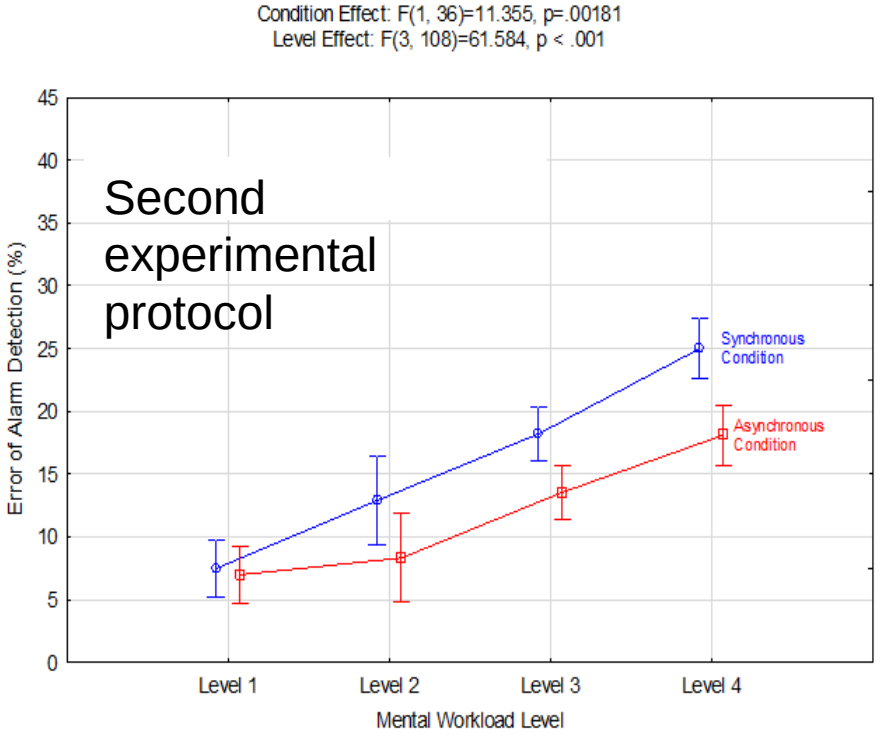




Source:
Vanderhaegen, F., Wolff, M., Mollard, R. (2023). A heartbeat-based study of attention in the detection of digital alarms from focused and distributed supervisory control systems. Cognition Technology & Work 25, 119–134.

Examples of possible hazardous RCPHS





Source:
Vanderhaegen, F., Wolff, M., Mollard, R. (2023). A heartbeat-based study of attention in the detection of digital alarms from focused and distributed supervisory control systems. Cognition Technology & Work 25, 119–134.

Examples of possible hazardous RCPHS

Shared autonomy principle	RCPHS example	Benefit (High probability of occurrence)	Inconsistency (Low probability of occurrence)	References
Group 2	Cruise Control (CC)	Reduction of workload and increasing of safety by avoiding overspeed	Increasing of reaction time, reduction of sustainable attention and inter-distances	Dufour, 2014; Lazarre, 2021
			Possible out-of-control in case of aquaplaning	NSC, 2009; Tan, 2019
			Possible overconsumption of fuel downhill and uphill	Johnston, 2018
			Possible collision in case of emergency if button "-" is used as braking system	Vanderhaegen, 2016

Adapted from:
Vanderhaegen, F. (2021). Heuristic-based method for conflict discovery of shared control between humans and autonomous systems - A driving automation case study. Robotics and Autonomous Systems, 146, 103867.

Examples of possible hazardous RCPHS

Example of CC use:

Rule base model of the CC use
RA1: (To deactivate the CC → To brake)
RA2: (To increase the speed setpoint of the CC → To push the button "+")
RA3: (To decrease the speed setpoint of the CC→ To push the button "-")

Rule base model of the CC behaviour:
RB1: ($Speed_{real}$ < $Speed_{setpoint}$ → To accelerate the engine
RB2: ($Speed_{real}$ > $Speed_{setpoint}$ → To brake the engine

Rule base model of the manual car speed:
RC1: (To increase the speed → To push on the gas pedal)
RC2: (To decrease the speed → To release the gas pedal)

Rule based affordances:
RD1: (To push the button "+" → To increase the speed)
RD2: (To push the button "-" → To decrease the speed)

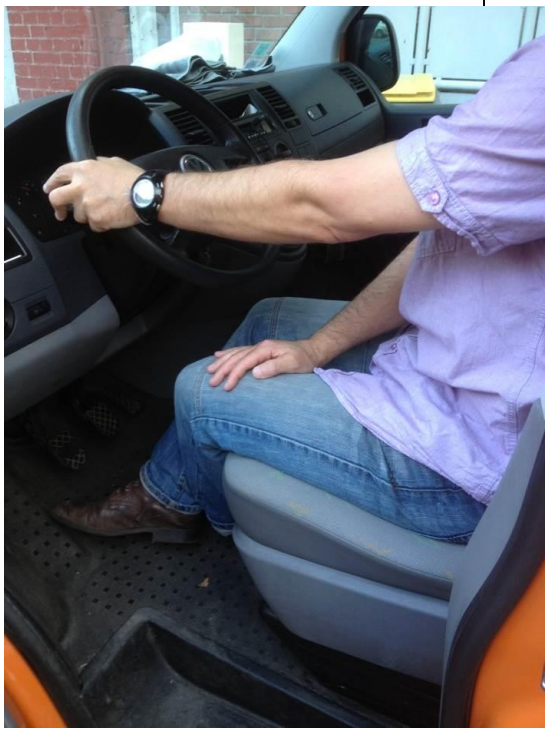
	Agree	Not agree	No opinion
RA & RB & RC	15	2	3
RD1	15	3	2
RD2	15	4	1

Adapted from:
Vanderhaegen, F. (2016). A rule-based support system for dissonance discovery and control applied to car driving. Expert Systems With Applications 65, 361-371.

Examples of possible hazardous RCPHS







Source:
Vanderhaegen, F. (2016). A rule-based support system for dissonance discovery and control applied to car driving. Expert Systems With Applications 65, 361-371.

Examples of possible hazardous RCPHS

Example of CC use:

Etude de 2013 du Prof. André Dufour (Centre d'Investigations Neurocognitives et Neurophysiologiques, Univ. Strasbourg) sur 90 sujets testés sur simulateur

Usage de régulateurs de vitesse:

- Augmentation des temps de réponse
- Réduction des distances de sécurité
- Augmentation de l'hypovigilance







Examples of possible hazardous RCPHS

Shared autonomy principle	RCPHS example	Benefit (High probability of occurrence)	Inconsistency (Low probability of occurrence)	References
Group 2	Adaptive Cruise Control (ACC)	Reduction of workload and increasing of safety by avoiding rear-end collision	Automated stop in a dangerous area	Vanderhaegen, 2021

Sources:
Vanderhaegen, F. (2021). Heuristic-based method for conflict discovery of shared control between humans and autonomous systems - A driving automation case study. Robotics and Autonomous Systems, 146, 103867.

Examples of possible hazardous RCPHS

Example of an ACC use





Adapted from:
Vanderhaegen, F. (2021). Heuristic-based method for conflict discovery of shared control between humans and autonomous systems - A driving automation case study. Robotics and Autonomous Systems, 146, 103867.

Examples of possible hazardous RCPHS

Shared autonomy principle	RCPHS example	Benefit (High probability of occurrence)	Inconsistency (Low probability of occurrence)	References
Group 2	Lane Keeping Assist	Reduction of workload and increasing of safety by avoiding involuntary deviations from a road lane	Possible increasing of fatigue and frustration due to regular resisitance on wheel by the driver againt LKA action	Lorentz Automotive, 2025
			Possible collision when overtaking	Vanderhaegen, 2021

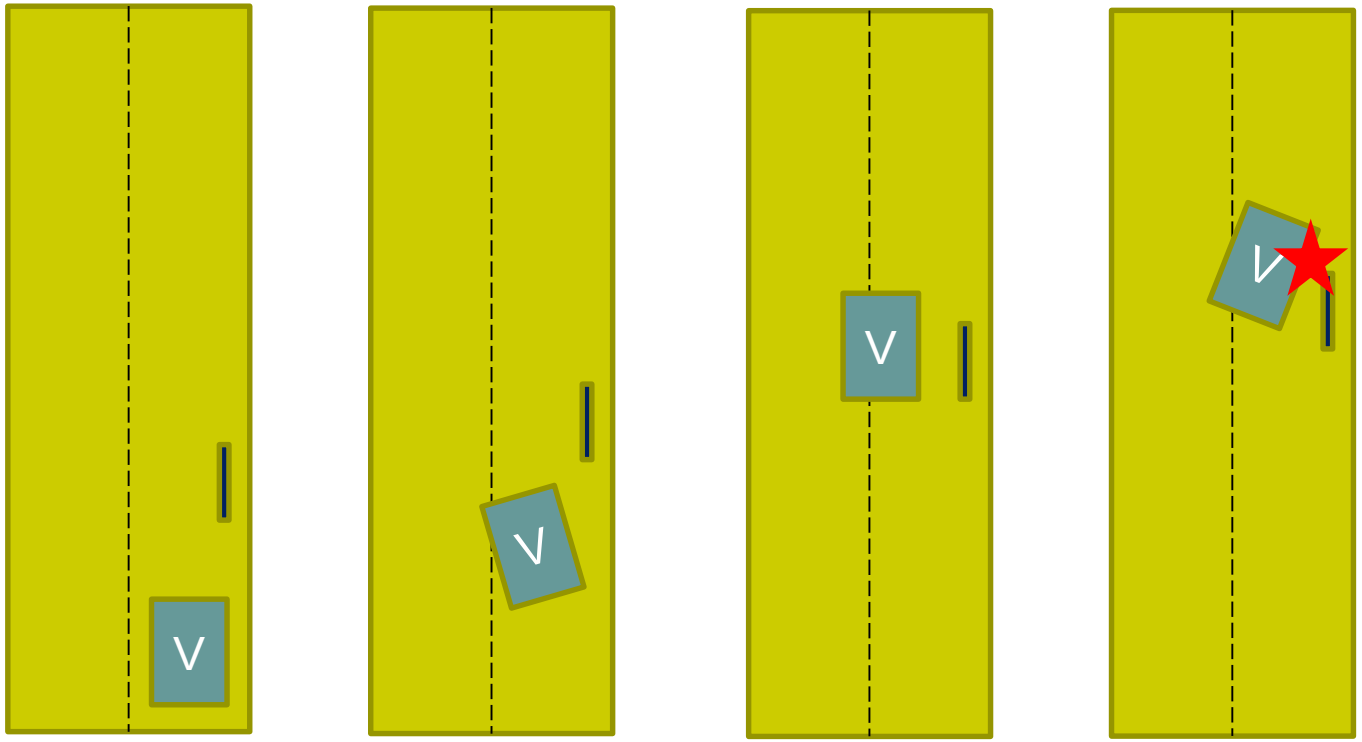
Source:

Lorentz Automotive (2025). The hidden dangers of Lane Keeping Assist systems. https://lorentzautoinc.com/blog/the-hidden-dangers-of-lane-keeping-assist-systems/3

Vanderhaegen, F. (2021). Heuristic-based method for conflict discovery of shared control between humans and autonomous systems - A driving automation case study. Robotics and Autonomous Systems, 146, 103867.

Examples of possible hazardous RCPHS

Example of a LKA use



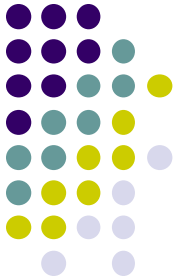
Source:.
Vanderhaegen, F. (2021). Heuristic-based method for conflict discovery of shared control between humans and autonomous systems - A driving automation case study. Robotics and Autonomous Systems, 146, 103867.

Examples of possible hazardous RCPHS

Shared autonomy principle	RCPHS example	Benefit (High probability of occurrence)	Inconsistency (Low probability of occurrence)	References
Group 2	Auto-Pilot	Reduction of workload or human errors and increasing of safety	Degradation of perception ability in case of system failure	Siddiqui, Merrill, 2023

Sources:
Siddiqui, F., & Merrill, J. B. (2023). 17 fatalities, 736 crashes: The shocking toll of Tesla's Autopilot. The Washington Post, July 10, 2023

Examples of possible hazardous RCPHS



Fatal use of driving auto-pilots

- ➔ Hands on the steering wheel not sufficient for maintaining human attention (Kravets, 2017; Olsen, 2018)
- ➔ No hands on the steering wheel for a few seconds before an accident (Olsen, 2018; Hawkins, 2019)
- ➔ Regular removal of hands from the steering wheel (BBC, 2019)

Source:
Vanderhaegen, F. (2021). Heuristic-based method for conflict discovery of shared control between humans and autonomous systems - A driving automation case study. Robotics and Autonomous Systems, 146, 103867.

Examples of possible hazardous RCPHS

Shared autonomy principle	RCPHS example	Benefit (High probability of occurrence)	Inconsistency (Low probability of occurrence)	References
Autonomous system	Automatic Emergency Braking (AEB)	Increasing system safety by avoiding head-on collision	Phantom braking phenomena can occur and generate accidents	Lilley, 2025

Source:
Lilley, Z. (2025). Reports of ‘phantom braking’ by cars on French roads raise safety concerns. The Connexion, August 5th 2025

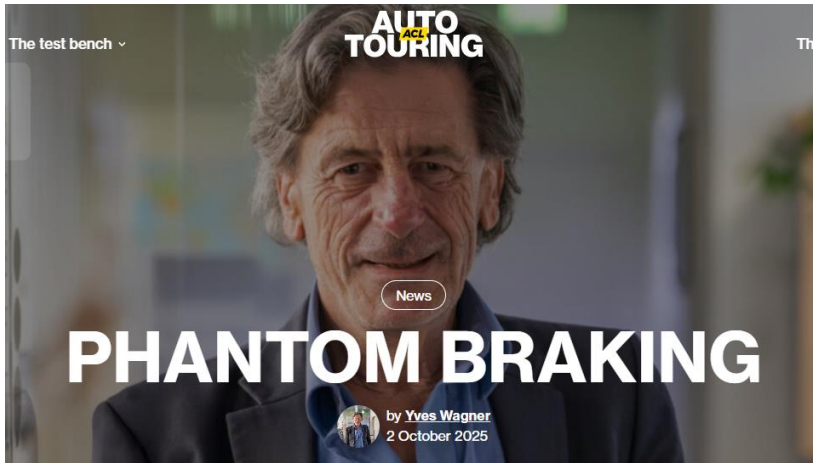
Examples of possible hazardous RCPHS


The Connexion
FRENCH NEWS IN ENGLISH SINCE 2002
Subscriber resources
French News | Visas & residency cards | EES | Money | Property | Practical | Your Questions | All categories

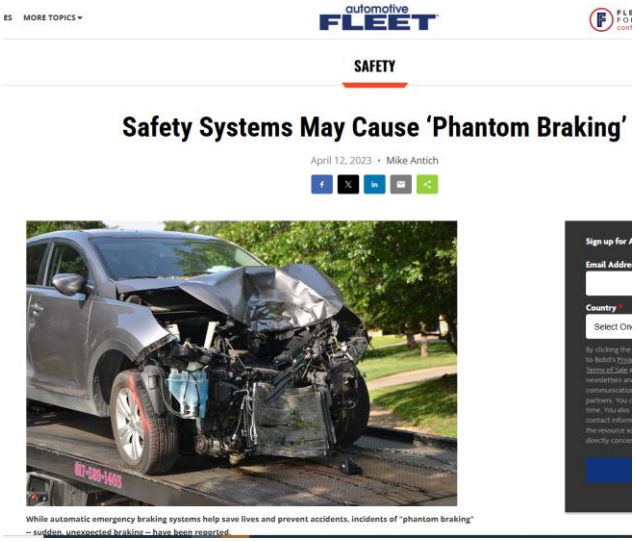
Reports of 'phantom braking' by cars on French roads raise safety concerns

Fear that potentially faulty braking assistance systems could lead to major crashes

Lilley, Z. (2025). Reports of 'phantom braking' by cars on French roads raise safety concerns. The Connexion, August 5th, 2025



Wagner, Y. (2025). Phantom braking. Autotouring, October 2nd , 2025


automotive FLEET
SAFETY
Safety Systems May Cause 'Phantom Braking'
April 12, 2023 • Mike Antich
While automatic emergency braking systems help save lives and prevent accidents, incidents of "phantom braking" -- sudden, unexpected braking -- have been reported.

Antich, M. (2023). Safety systems may cause 'Phantom Braking'. Automotive Fleet, April 12th, 2023

Examples of possible hazardous RCPHS

Increase of crashes:
- CC: +12%
- ACC: + 8%

Decrease of crashes:
- LKA: -19,1%
- AEB: -10,7%

Source:
de Winkel, K. N., Christoph,C. (2025). Rethinking Advanced Driver Assistance System taxonomies: A framework and inventory of real-world safety performance.Transportation Research Interdisciplinary Perspectives, 29, 101336.

Conclusion: toward “all inclusive” RCPHS

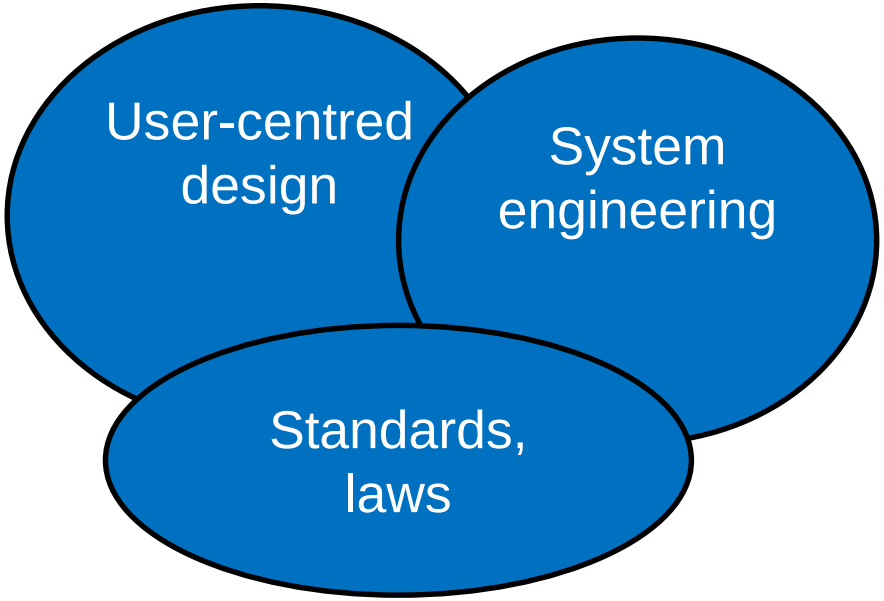
Inclusivity / Inclusiveness

Paradigm	Principle
RCPHS for inclusive education/pedagogy/learning	Education for all without any discrimination
RCPHS for inclusive design	Design for all
RCPHS for inclusive manufacturing	Product manufactured by all and used by all
RCPHS for inclusive robotics	Robots for human autonomy growth
RCPHS for inclusive mobility	Sustainable accessibility to transport for everybody and everywhere
All-inclusive RCPHS	All-in-one

Adapted from:
Vanderhaegen, F., Jimenez, V. (2022). Opportunities and threats of interactions between humans and Cyber-Physical Systems – Integration and inclusion approaches for CPHS. In: Cyber-Physical-Human Systems: Fundamentals and Applications (CPHS). A. Annaswamy, P. Khargonekar, F. Lamnabhi-Lagarrigue, S. Spurgeon (Eds). John Wiley & Sons / IEEE Press.

Conclusion: toward “all inclusive” RCPHS

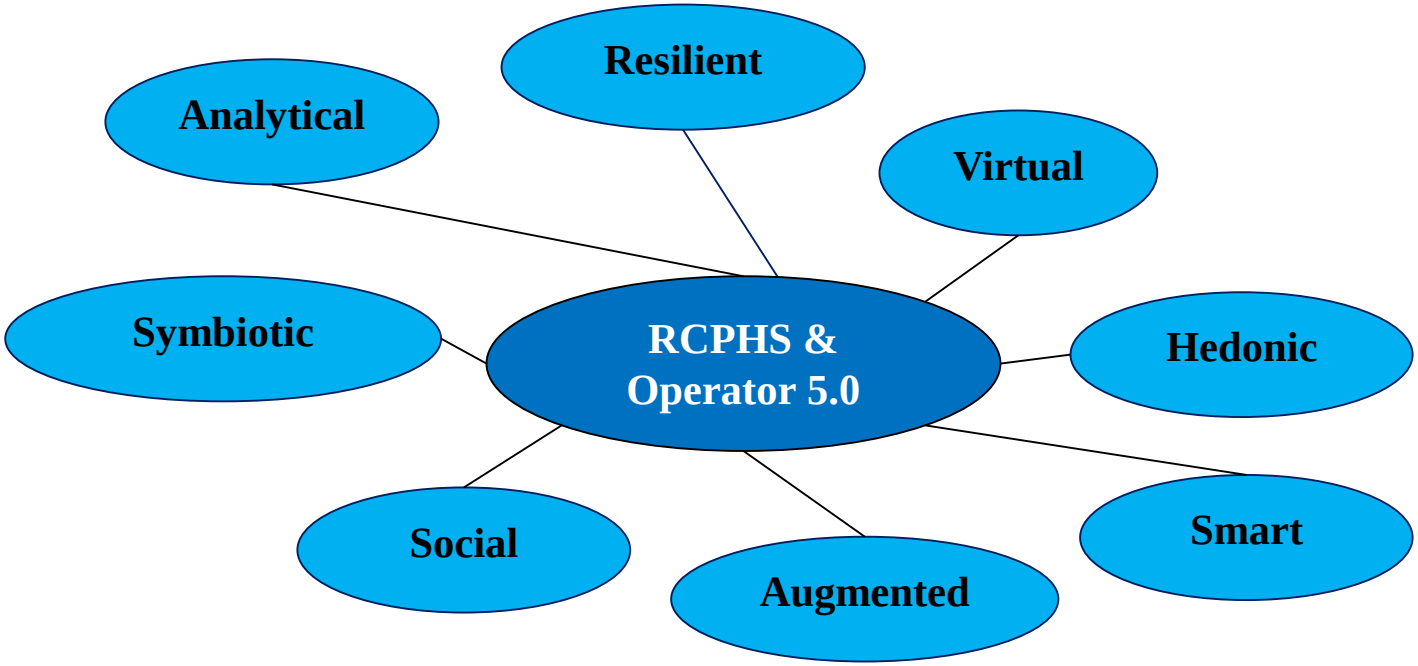
"I never lose. I either win or learn“ Nelson Mandela



Tailor-made symbiosis	All-in-one symbiosis
Integrative CPHS	Inclusive CPHS
Adapt all users to norms	Adapt norms to variability of users

Adapted from:
Vanderhaegen, F., Jimenez, V. (2022). Opportunities and threats of interactions between humans and Cyber-Physical Systems – Integration and inclusion approaches for CPHS. In: Cyber-Physical-Human Systems: Fundamentals and Applications (CPHS). A. Annaswamy, P. Khargonekar, F. Lamnabhi-Lagarrigue, S. Spurgeon (Eds). John Wiley & Sons / IEEE Press.

Conclusion: toward “all inclusive” RCPHS



Adapted from:

Romero, D., Stahre, J. (2021). Towards The Resilient Operator 5.0: The Future of Work in Smart Resilient Manufacturing Systems. Procedia CIRP, 104, 1089-1094.

Romero, D., Stahre, J., Wuest, T., Noran, O. (2016). Towards an Operator 4.0 typology: a human-centric perspective on the fourth industrial revolution technologies. International Conference on Computers & Industrial Engineering (CIE46), At Tianjin, China

Responsible Cyber-Physical & Human Systems: dream or reality?

Frédéric VANDERHAEGEN
Université Polytechnique Hauts-de-France - INSA Hauts-de-France
Valenciennes, France