Privacy Protection, Ethics and Regulations for Autonomous Systems summary

S. Altini, V. Mygdalis, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr

Version 2.1

Artificial Intelligence &
Information Analysis Lab

Privacy Protection, Ethics and Regulations for Autonomous Systems

- **Data Security**
- Privacy Protection
- Moral Machine
- Safety and Regulations
- Dual use

Data Security

Data security is the process of protecting sensitive information from unauthorized third-party access or malicious attacks and exploitation of data. It is set up to ensure the integrity of the data, including all of the different practices used to secure data from misuse (encryption, access restrictions, etc).

Data Security in Autonomous Cars

Data security in ACs ensures the integrity of data collected from surrounding environment, ranging from basic navigation to owner/passenger information.

- Data collected by autonomous vehicles raise privacy concerns.
- There is no complete legislation, especially for levels 4 to 5, and preventive measures for data security per se against misuses, to counter vulnerabilities exploited by attackers.

Data Protection issues in Autonomous Cars

Public perceive ACs as privacy infringing machines

- ***Data stored within ACs***: enabling access only on people with authentication/authorized use; encrypted data to customize comfort, safety, entertainment settings.
- ***Data stored in ground infrastructure***: a combination of sensors (radar, LiDAR, computer vision, sonar, GPS) captures continuous data about vehicle's operation and its surroundings;
- ***Data transmitted over the air***: Wi-Fi/GPS transmitted data are unencrypted and unauthenticated (civilian GPS signals); used for navigation purposes (route information, speed, real-time traffic data); data protection with authentication and encryption mechanisms needed.

Data Security in Autonomous Cars

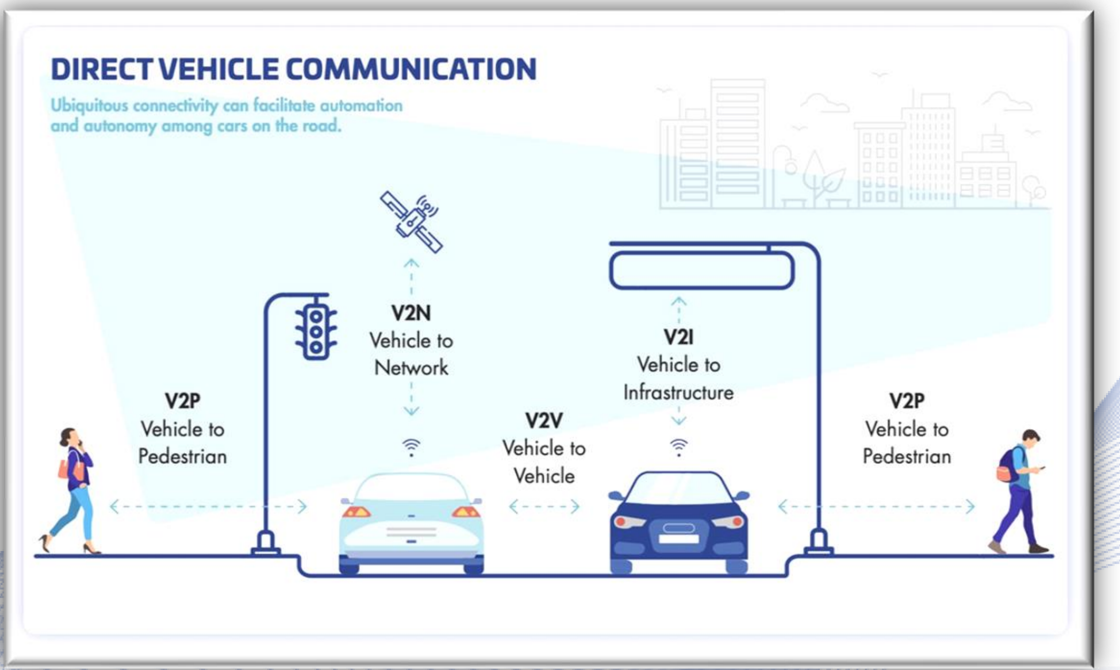
Attacker's objectives:

- Communication disruption;
- Jamming on components;
- Manipulation on software;
- Vehicle hacking.

Data Security in Autonomous Cars

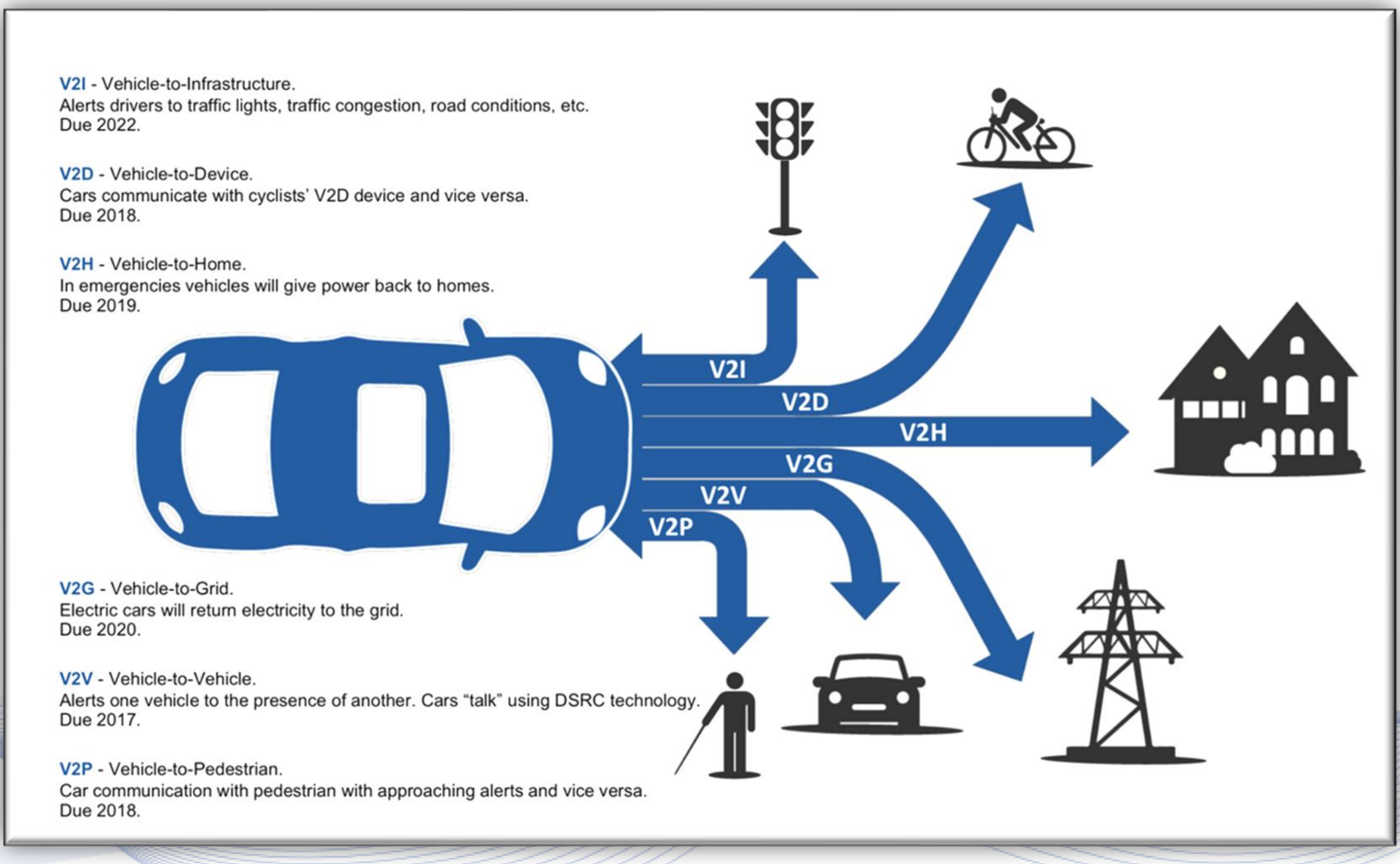
Potential vulnerable components and behavior

- ***Communication***: depending on specific target communication type (V2V, V2I, V2X);


DIRECT VEHICLE COMMUNICATION
Ubiquitous connectivity can facilitate automation and autonomy among cars on the road.
V2P
Vehicle to Pedestrian
V2N
Vehicle to Network
V2V
Vehicle to Vehicle
V2I
Vehicle to Infrastructure
V2P
Vehicle to Pedestrian

Direct vehicle communication [THA2020] (https://www.thalesgroup.com/en/markets/digital-identity-and-security/iot/industries/automotive/use-cases/v2x)

Data Protection issues in Autonomous Cars



V2I - Vehicle-to-Infrastructure.
Alerts drivers to traffic lights, traffic congestion, road conditions, etc.
Due 2022.

V2D - Vehicle-to-Device.
Cars communicate with cyclists' V2D device and vice versa.
Due 2018.

V2H - Vehicle-to-Home.
In emergencies vehicles will give power back to homes.
Due 2019.

V2G - Vehicle-to-Grid.
Electric cars will return electricity to the grid.
Due 2020.

V2V - Vehicle-to-Vehicle.
Alerts one vehicle to the presence of another. Cars "talk" using DSRC technology.
Due 2017.

V2P - Vehicle-to-Pedestrian.
Car communication with pedestrian with approaching alerts and vice versa.
Due 2018.

Vehicle-to-X Communication and Use Cases [ALA2018]

Data Security in Autonomous Cars

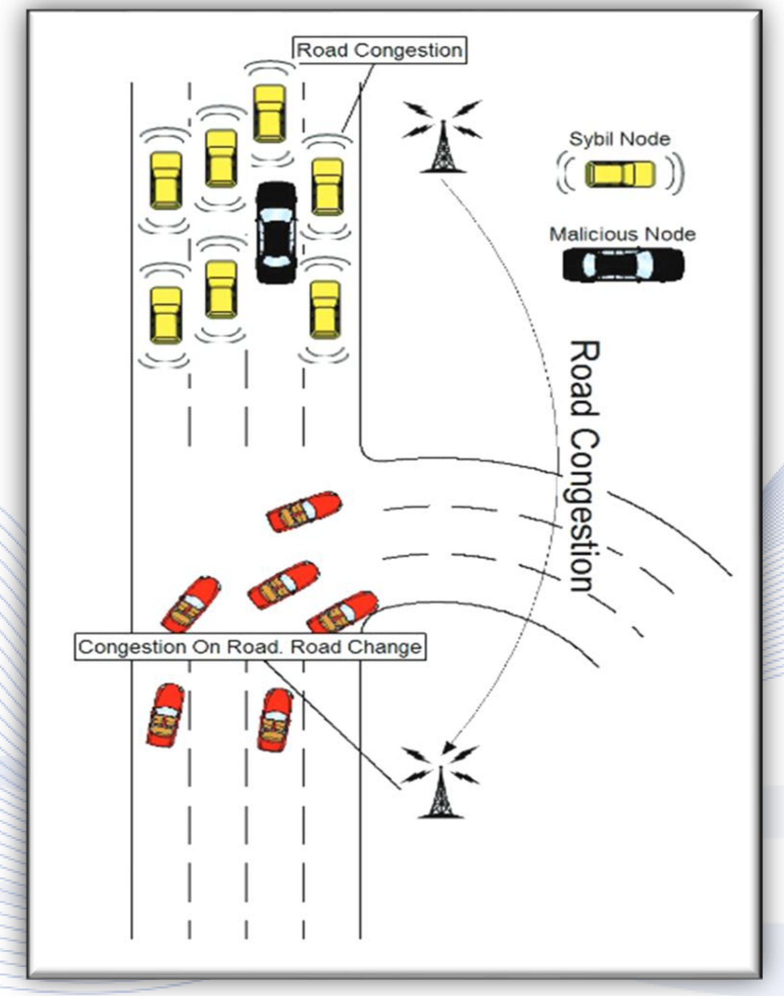
Types of Cyber Security Attacks on ACs:

- ***Attacks on authentication***: Sybil attack; GPS spoofing attack; Wormhole attack; Timing attack; Replication attack;
- ***Attacks on routing***: Eavesdropping attack; DoS attack; Misrouting attack; Flooding; Jamming attack;

Data Security in Autonomous Cars

Attacks on authentication

- ***Sybil attack***: a malicious vehicle creates fake identities (multiple vehicles at different locations) to gain the trust of legitimate ACs; use valid credentials to authenticate the Sybil vehicles; disruption/falsification of the whole network infrastructure.



Scenario for Sybil attack effect [RAB2015]

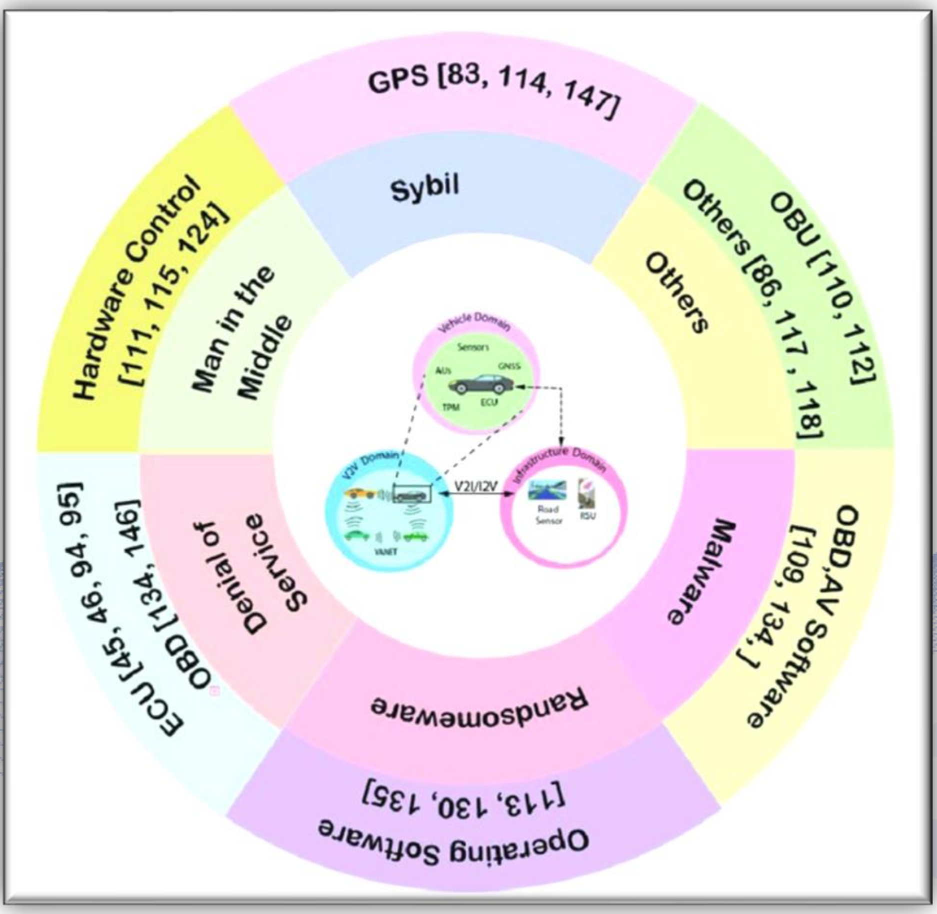
Data Security in Autonomous Cars

Types of Cyber Security attacks on ACs

Classification of the attacks happened on self-driving cars.

Middle layer: types of attacks;

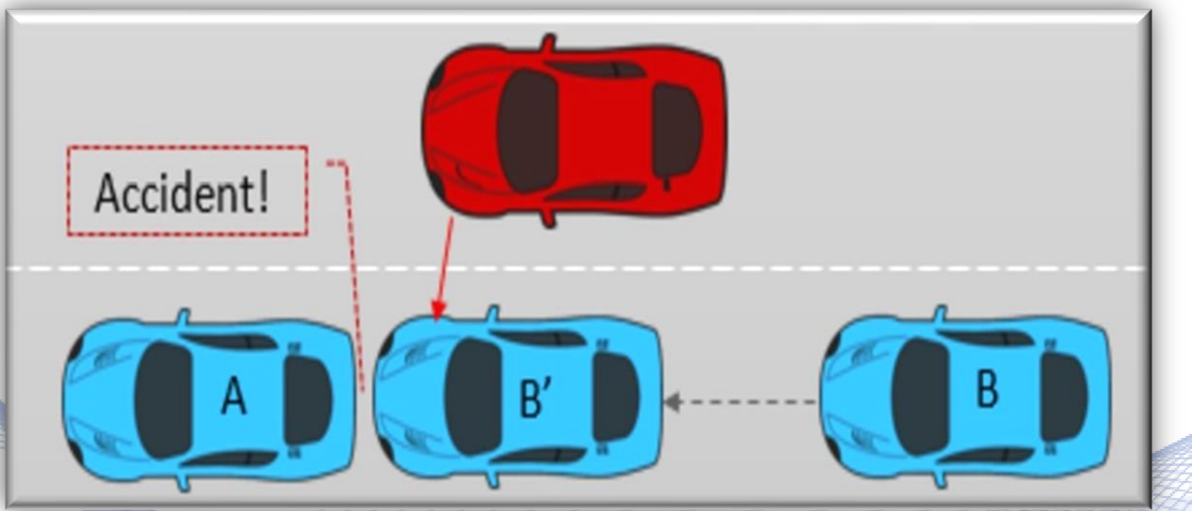
Out layer: components of a AV affected by relevant attacks. [CHO2020]



Artificial Intelligence & Information Analysis Lab

Data Security in Autonomous Cars

Attacks on authentication



An example of Timing attack: attacker is obligated to communicate positional information of A, when B change the lane; but the attacker adds a time delay to the information and delivers the information only when B changes its position to B, leading to an accident [ELR2020]

Data Security in Autonomous Cars

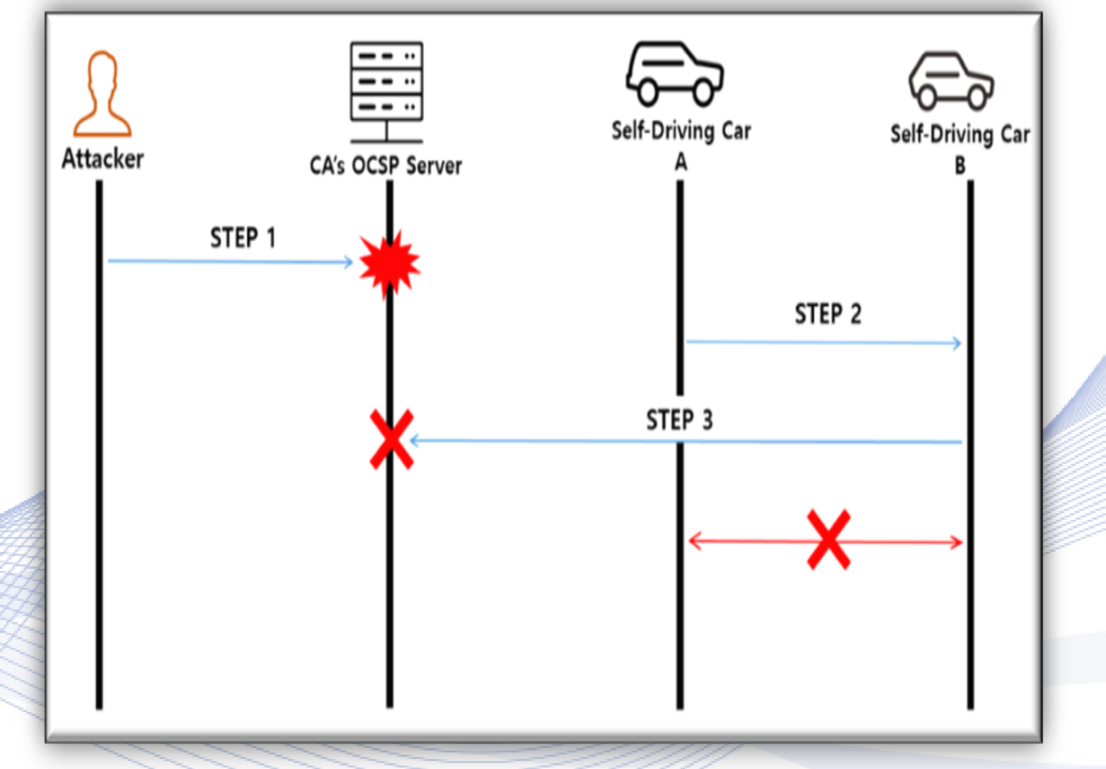
Attacks on routing

- ***Eavesdropping attack***: affect crucial information, e.g., exploitation of location;
- ***Misrouting attack***: divert traffic by adding virtual nodes;
- ***Flooding attack***: hinders communication path with messages, destructing service requests for RSI/AV.

Data Security in Autonomous Cars

Attacks on routing

- ***Distributed Denial of Service*** (***DOS***) ***attack targets***: getting access in network resources/services via malicious code/spear phishing; making services disabled; impossible to verify signatures included data; making authentication impossible at V2V communication/overall self-driving car environments; road traffic paralysis.



Example of Distributed Denial of Service (D-DoS) attack [LIM2017]

Drone Data Security

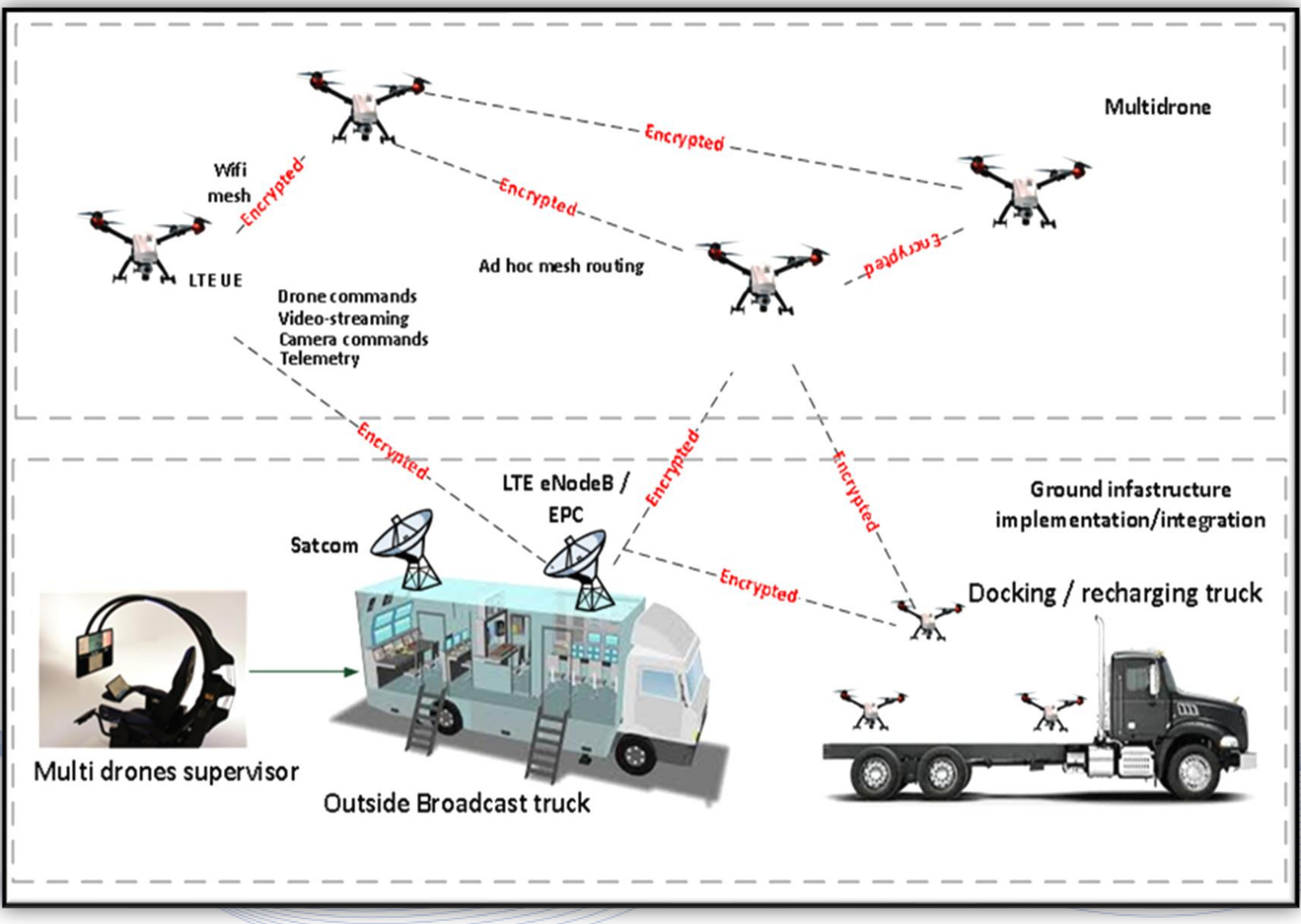
Classification of drone communications:

- ***Drone-To-Drone (D2D)***: Peer-to-Peer (P2P) communication expose system in various P2P vulnerabilities and attacks (D-DoS, sybil attacks). Machine Learning optimizes the wireless communication system, but yet not reached the standardized status.
- ***Drone-To-Ground station (D2GS)***: communication used protocols (Bluetooth/Wi-Fi)/public and unsecure/using single factor authentication. Vulnerable functioning to active (man-in-the-middle) and passive (eavesdropping) attacks.

Drone Data Security

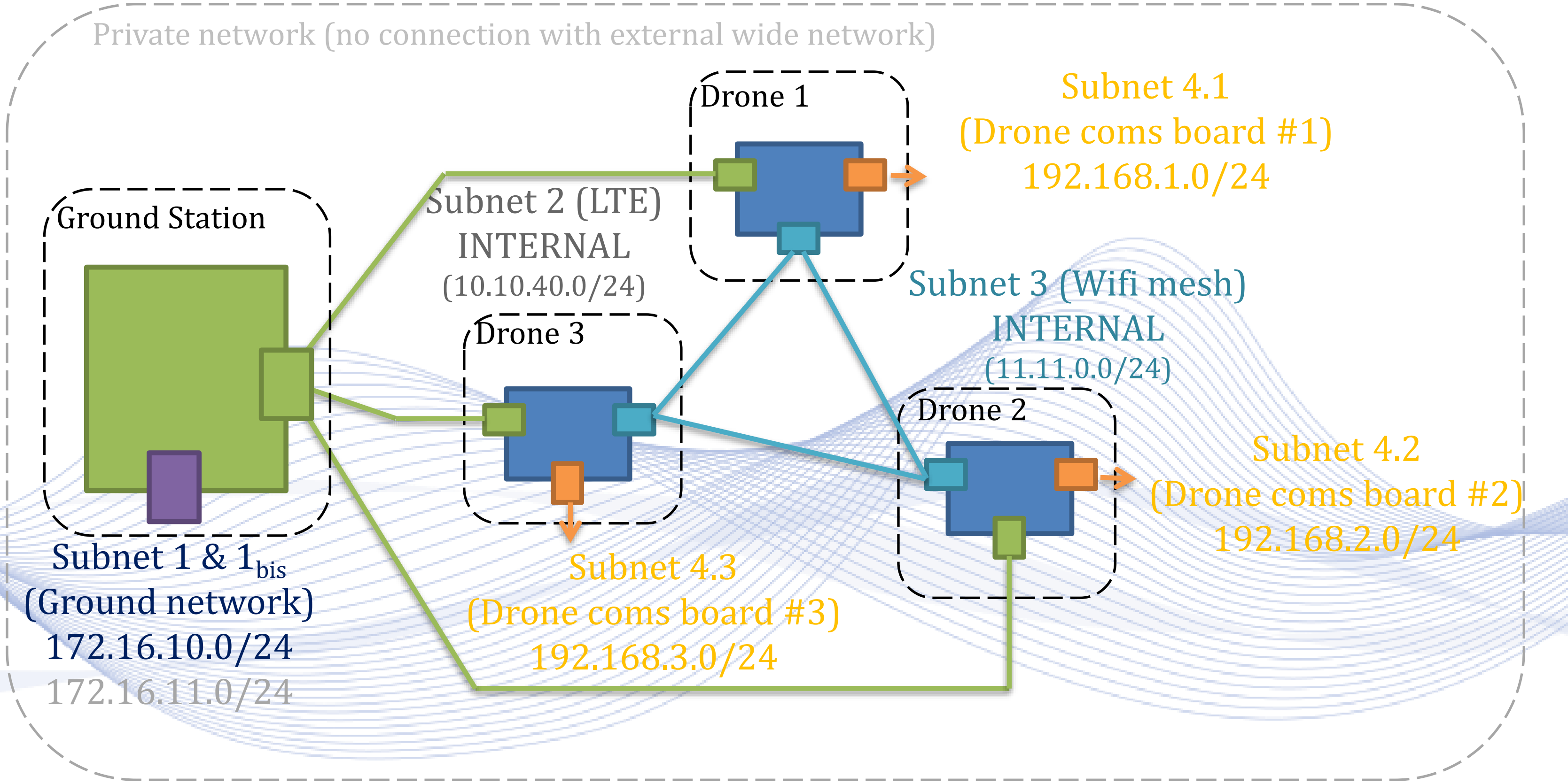
Classification of drone communications:

- ***Drone-To-Network (D2N)***: through this type, the given option is to choose the network, using the required security level, including cellular communications that, also, need to be secured.
- ***Drone-To-Satellite (D2S)***: communication used to send real-time coordinates via GPS. Satellite communications are considered as secure/safe, exhibit substantial cost/maintenance requirements.


Multidrone
Encrypted
Wifi mesh
Encrypted
Encrypted
LTE UE
Ad hoc mesh routing
Encrypted
Drone commands
Video-streaming
Camera commands
Telemetry
Encrypted
LTE eNodeB / EPC
Encrypted
Encrypted
Ground infastructure implementation/integration
Satcom
Encrypted
Docking / recharging truck
Multi drones supervisor
Outside Broadcast truck

Drone Communications

Objective: to provide secured and resilient transparent IP access to drones and ground station (using both LTE and Wifi).



Privacy Protection, Ethics and Regulations for Autonomous Systems

- Data Security
- **Privacy Protection**
- Moral Machine
- Safety and Regulations
- Dual use

Privacy and Data Protection

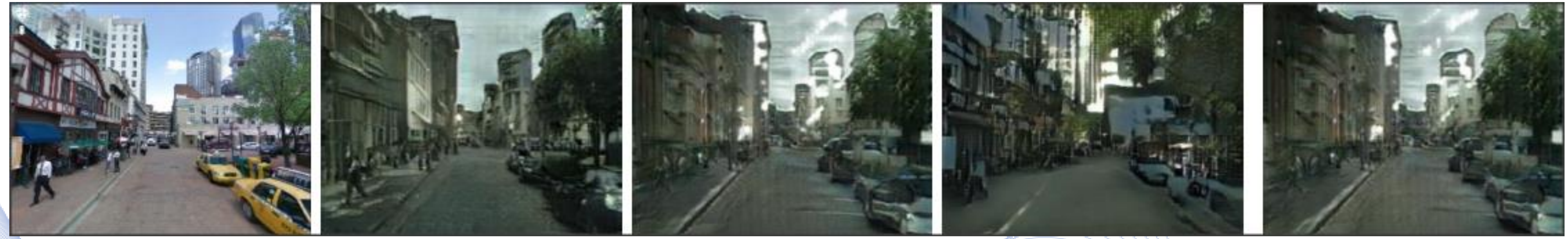
- ***GDPR***: data protection legislation within EU, replaces Directive 95/46/EC, complementing accountability requirement (article 5, paragraph 2).
- ***Special data***: “accessibility, exchange, re-use” of data related to static road data (article 4), dynamic road data (article 5), traffic data (article 6), according to EU Delegated Regulation 2015/962, have to be granted.
- ***Anonymization*:** data could be anonymized, ensuring that cannot be re-identified.

Data Protection issues in Autonomous Cars

***Usage data collected to be considered**:*

- ***Geolocation data***: collected data should be compliant with the following:
  - ***Adequate configuration of access frequency/detail level on geolocation data***, e.g. no access of weather application to AC's geolocation per second, even with owners' consent;

Privacy Protection, ethics and regulatory issues



ground truth pix2pix pix2pix+pri UNIT+pri ADGAN

Visual quality comparison of generated images for Google Street View. Left to right: ground truth of input, pix2pix result, pix2pix+pri result, UNIT+pri result and ADGAN result [XIO2019].

Data Protection issues in Drones

- ***Drone privacy breach issues***: trespassing/flights above private property are forbidden. Distinction between:
  - actors, spectators, crowd;
  - public events, private events.
- ***Data protection issues for AV shooting***:
  - broadcasting;
  - developing experimental databases.
- ***Use of data de-identification algorithms***, during a shooting.

Privacy Protection, Ethical and Regulatory issues

Ethics for Drones

- ***Privacy***: entrance/view of drones in private spaces; issues concerning over privacy in public settings, e.g., recording capabilities;
- ***Safety***: reckless/dangerous use of drones, especially in high-crowded areas (beaches, events);
- ***Enforceability***: official possibility for imposing regulations in drones;

Privacy Protection, Ethical and Regulatory issues

Ethics for Drones

- ***Crime***: used to thievery/break-in, infringement and trespassing;
- ***Nuisance***: used to harass/disrupt of individuals in public setting;
- ***Professional/private use***: whether regulation should be differentiated for professional and recreational purposes.

Privacy protection, Ethical and Regulatory issues

Technical issues

- No-filming zones;
- Face de-identification;
- Protection of private spaces.

No-filming zones

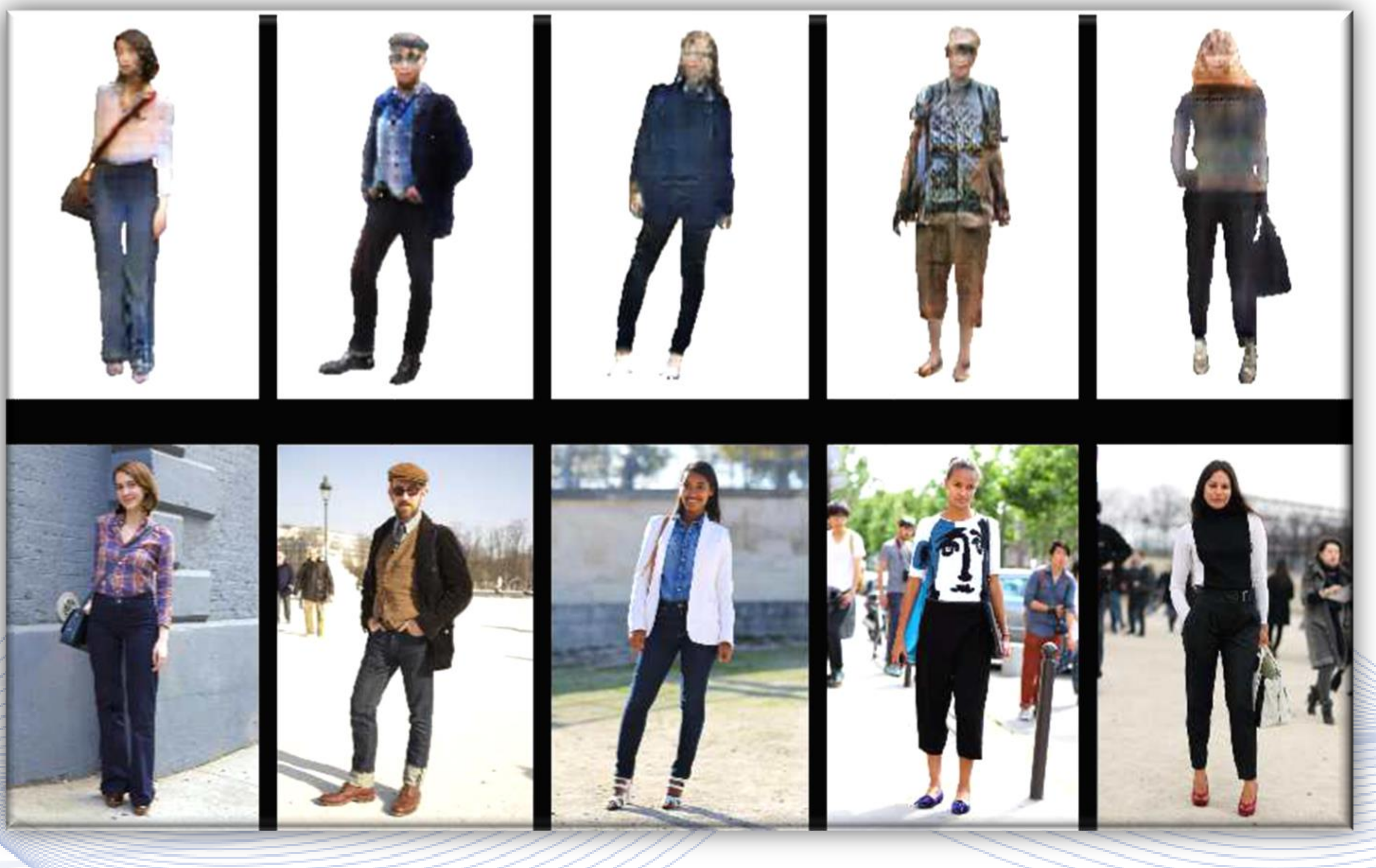
- ***UAV shooting*** over private areas is not allowed;
- ***No-filming zones*** must be automatically taken into account during mission planning and replanning;
- ***Private spaces*** (e.g., home gardens) can be geofenced. Geofencing information can be used to blur the image of such private spaces.

De-identification methods

- Face detection obfuscation.
- Face de-identification:
  - Hindering face recognition or verification.
- Human body de-identification.
- Car plate de-identification.

Example: Google Maps

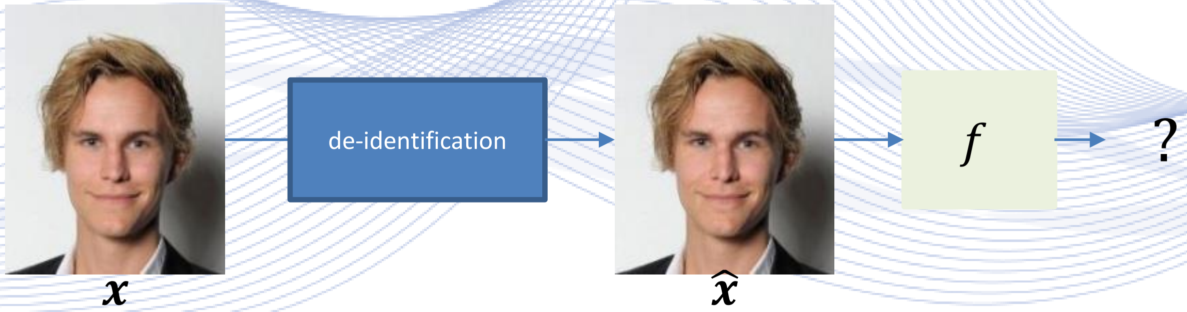
Body de-identification



[BRK2017] K. Brkic, I. Sikiric, T. Hrkac, Z. Kalafatic, “I Know That Person: Generative Full Body and Face De-Identification of People in Images”, in proc. *CVPR*, 2017.

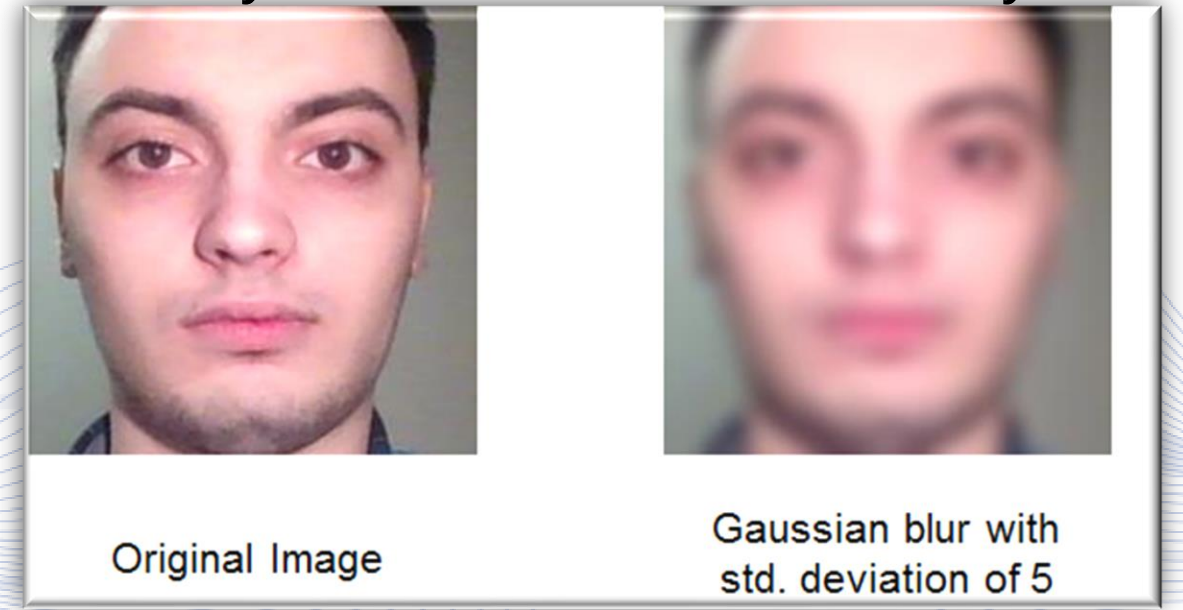
Face De-Identification

- Face recognition systems f take a facial image $\mathbf{x}$ as input and predict its corresponding identity y, $f(\mathbf{x}) \rightarrow y$.
- Therefore, ***face de-identification*** methods aim to alter the original facial image $\boldsymbol{x}$ and produce a de-identified image $\hat{\mathbf{x}}$ that can no longer be identified by face recognition systems, $f(\mathbf{x}) \rightarrow ?$.



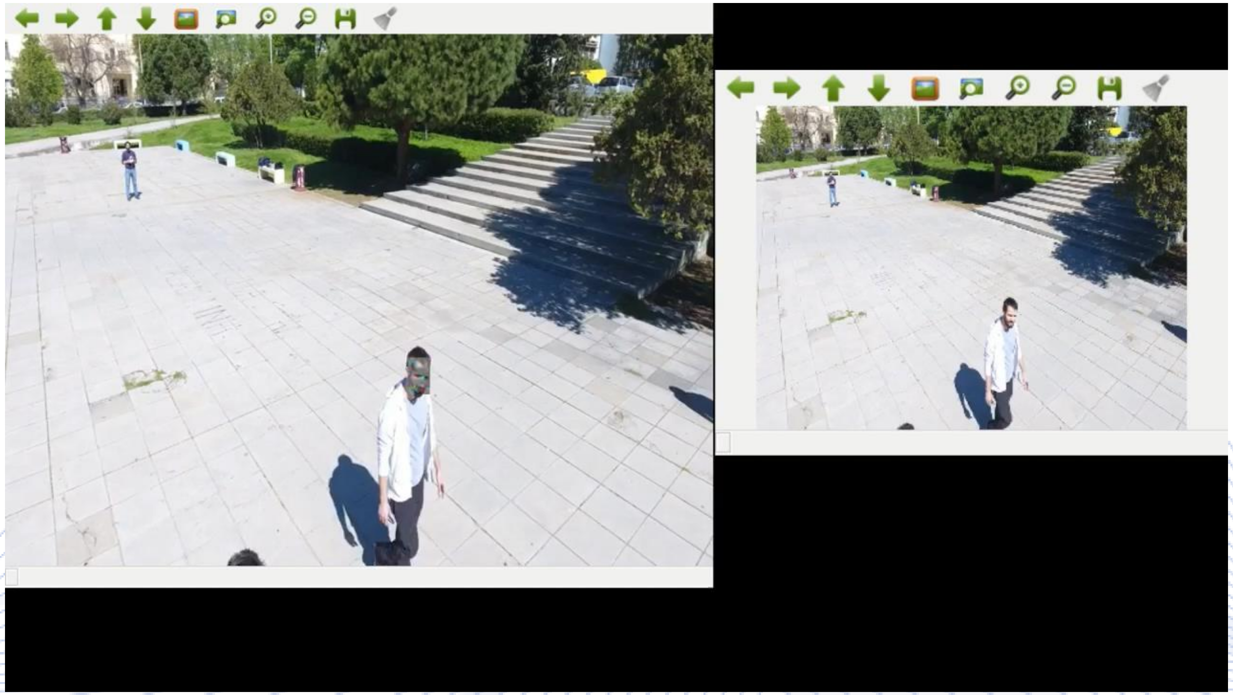
Naïve approaches

- Naïve de-identification refer to applying additive noise (e.g., Gaussian, impulse) to the (detected) input facial image region, until the system fails to detect/classify the face.



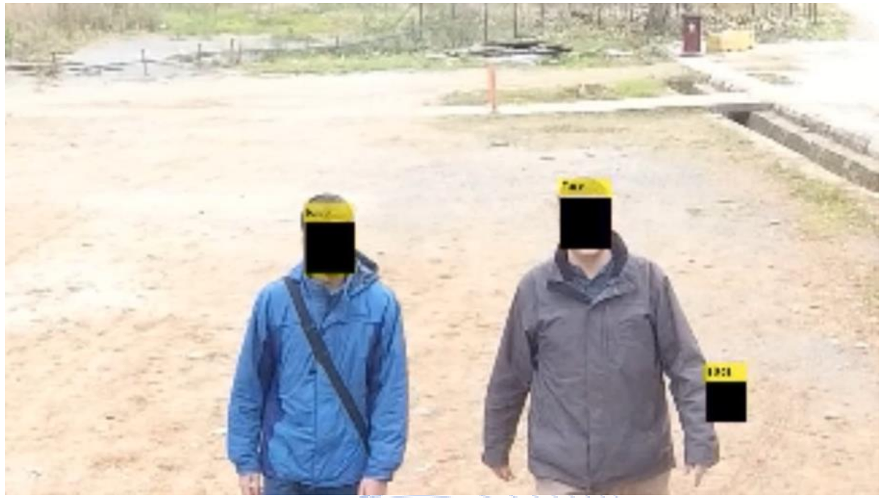
Facial images after Gaussian filtering [CHR2018]

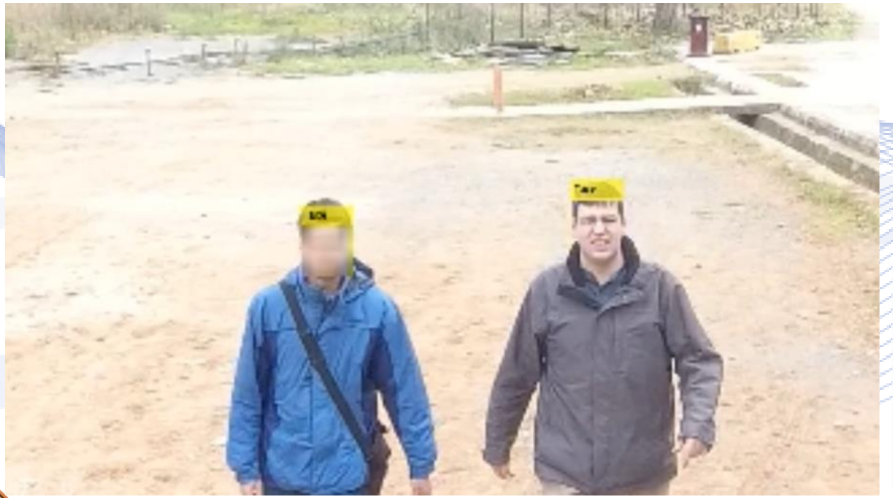
Naïve approaches

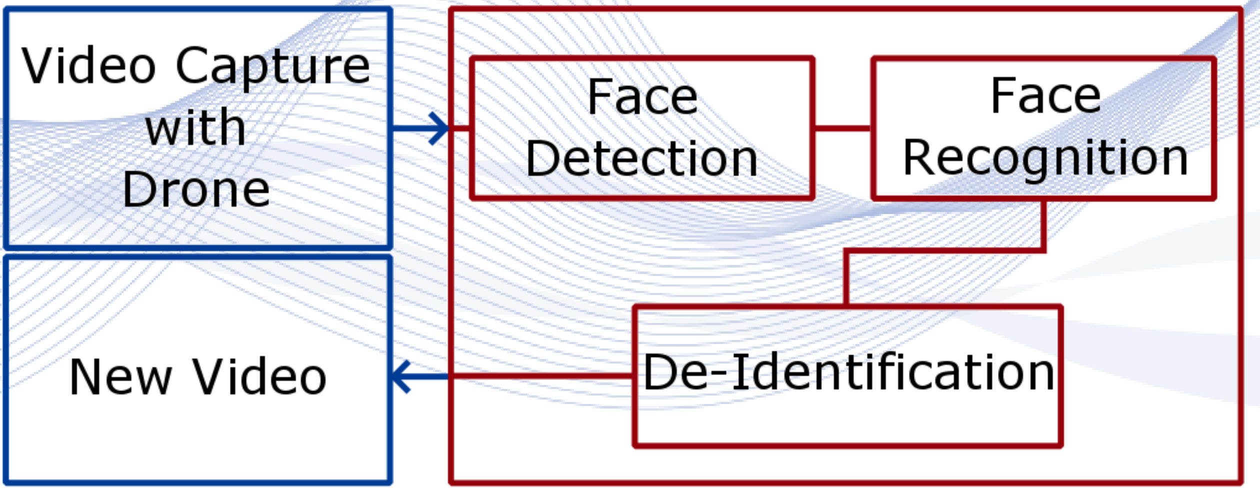


SVD-DID video de-identification.

Face de-identification on drone videos







Video Capture with Drone

Face Detection

Face Recognition

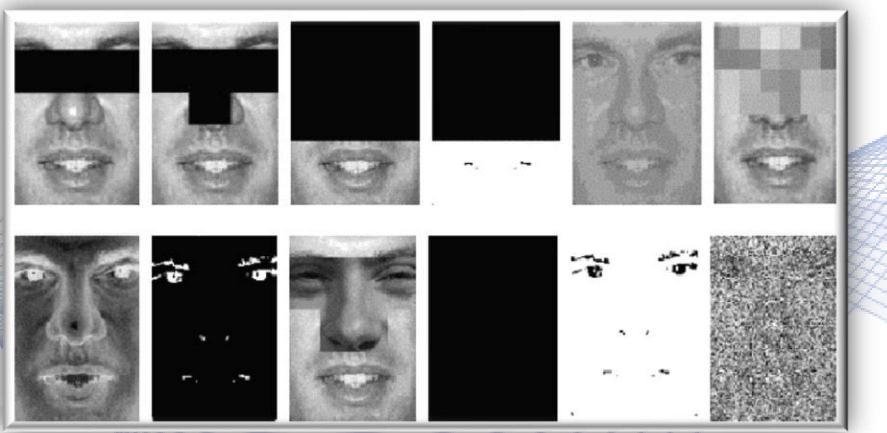
De-Identification

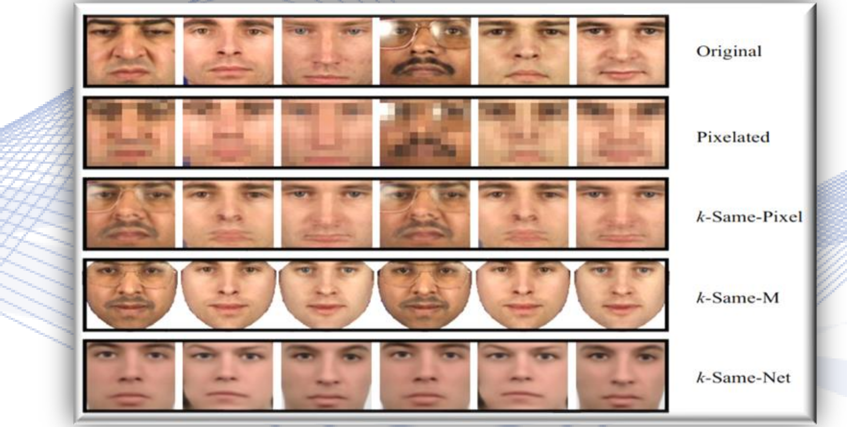
New Video

Adversarial Face De-Identification

Drawbacks of previous face de-identification methods.

- Previous face de-identification methods strongly alter original images.
- De-identified image should retain the original facial image unique characteristics (e.g. race, gender, age, expression, pose).





[NOU2020]

Adversarial attacks & Defenses

- ***Adversarial Attacks***
- May be employed for privacy protection

- ***Adversarial Defenses***
  - May be employed for content protection against adversarial attacks (e.g., copyright protection systems).

Privacy Protection, Ethics and Regulations for Autonomous Systems

- Data Security
- Privacy Protection
- **Moral Machine**
- Safety and Regulations
- Dual use

Privacy protection, Ethics and Regulations

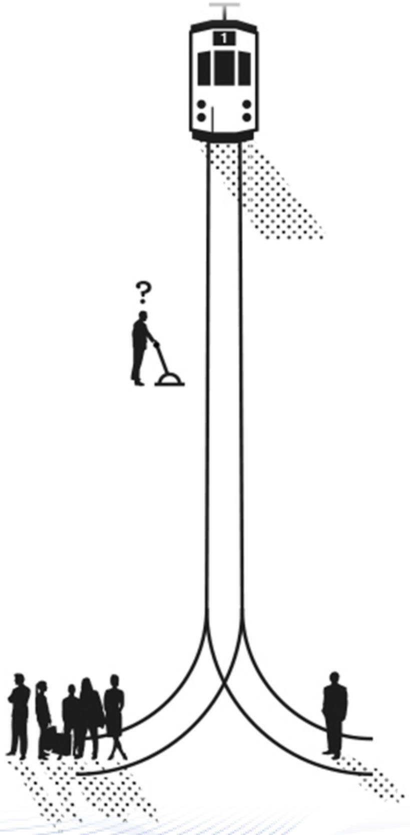
Ethical norms and moral values

Question: How "fully autonomous" cars are or whether they should be morally autonomous?

AVs as different agents than humans, need to be adjusted to ***a human oriented "original" virtue ethics***.

Privacy Protection, Ethics and Regulations

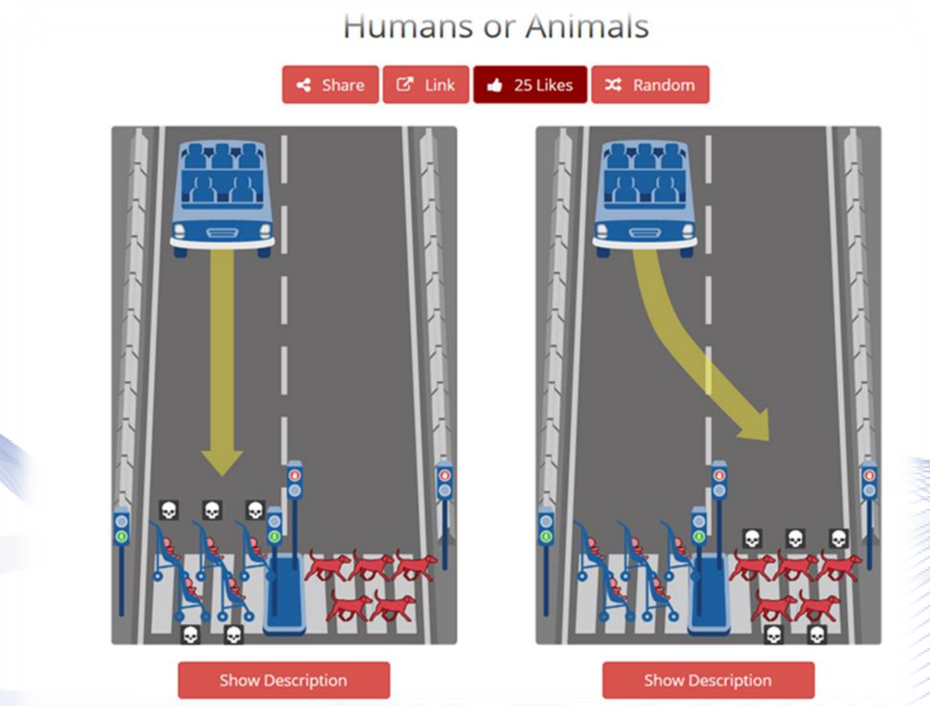
- ***Trolley problem*** describes "a moral dilemma that either way, harm to persons is unavoidable and there are good ethical reasons for one or the other behaviour".
- ***Trolley cases*** are "dramatic, stylized, black-and-white situations that have little resemblance to real-life extreme traffic situations".

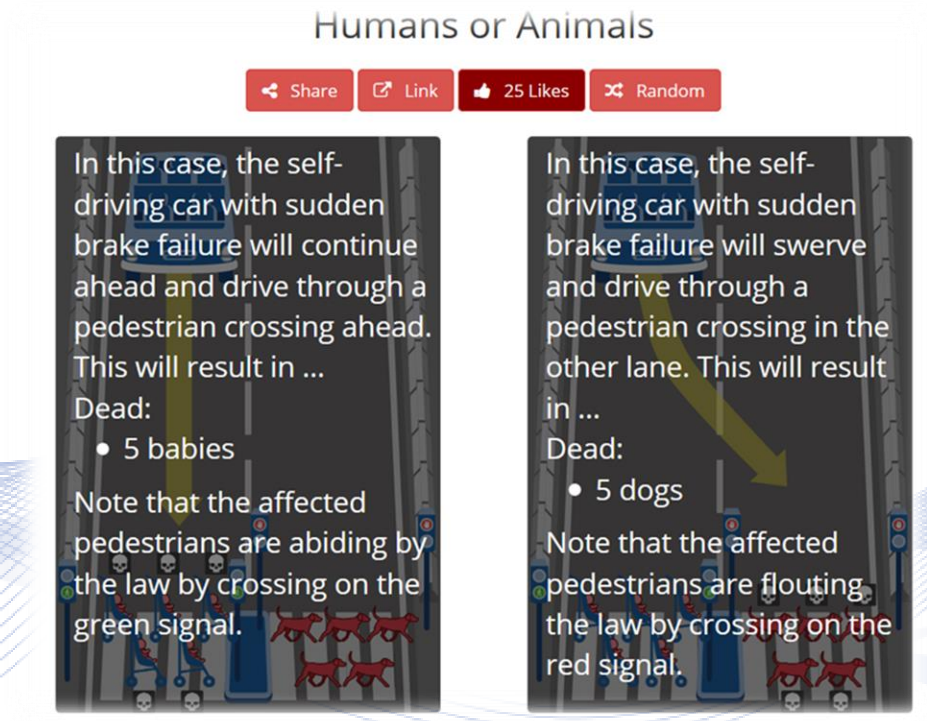


The Trolley Experiment [BEA2018]

Privacy Protection, Ethics and Regulations

“***Moral Machine***” - an online experimental platform: moral preferences in AVs moral dilemmas





In this case, the self-driving car with sudden brake failure will continue ahead and drive through a pedestrian crossing ahead. This will result in ...

Dead:

- 5 babies

Note that the affected pedestrians are abiding by the law by crossing on the green signal.

In this case, the self-driving car with sudden brake failure will swerve and drive through a pedestrian crossing in the other lane. This will result in ...

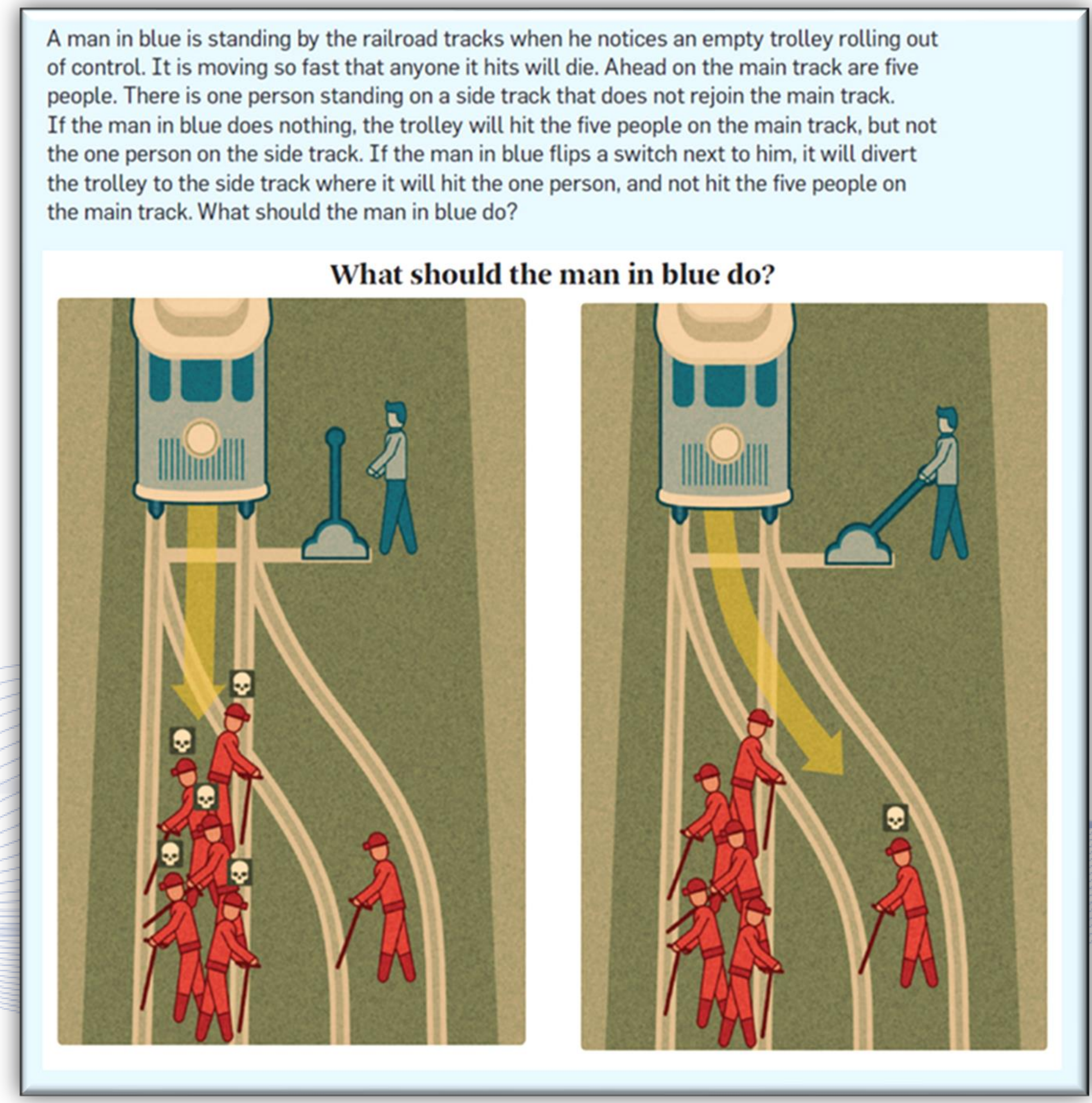
Dead:

- 5 dogs

Note that the affected pedestrians are flouting the law by crossing on the red signal.

Moral machine scenarios: https://www.moralmachine.net

Moral dilemmas [AWA2018]



A man in blue is standing by the railroad tracks when he notices an empty trolley rolling out of control. It is moving so fast that anyone it hits will die. Ahead on the main track are five people. There is one person standing on a side track that does not rejoin the main track. If the man in blue does nothing, the trolley will hit the five people on the main track, but not the one person on the side track. If the man in blue flips a switch next to him, it will divert the trolley to the side track where it will hit the one person, and not hit the five people on the main track. What should the man in blue do?

What should the man in blue do?

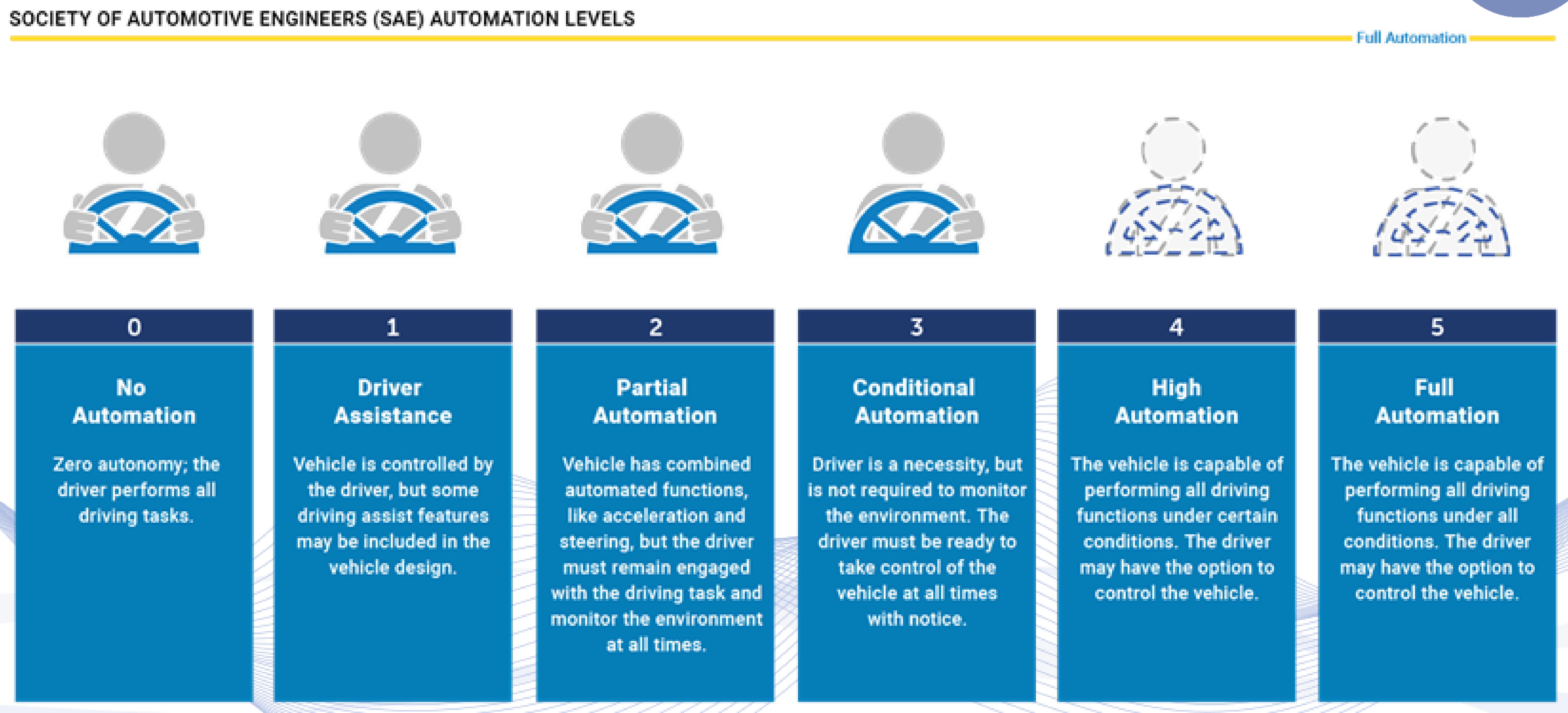
Privacy Protection, Ethics and Regulations for Autonomous Systems

- Data Security
- Privacy Protection
- Moral Machine
- **Safety and Regulations**
- Dual use

ACs levels Autonomy

The 6 Levels of Cars Autonomy

- ***Level 0***: No Driving Automation
- ***Level 1***: Driver Assistance
- ***Level 2***: Partial Driving Automation
- ***Level 3***: Conditional Driving Automation
- ***Level 4***: High Driving Automation
- ***Level 5***: Full Driving Automation



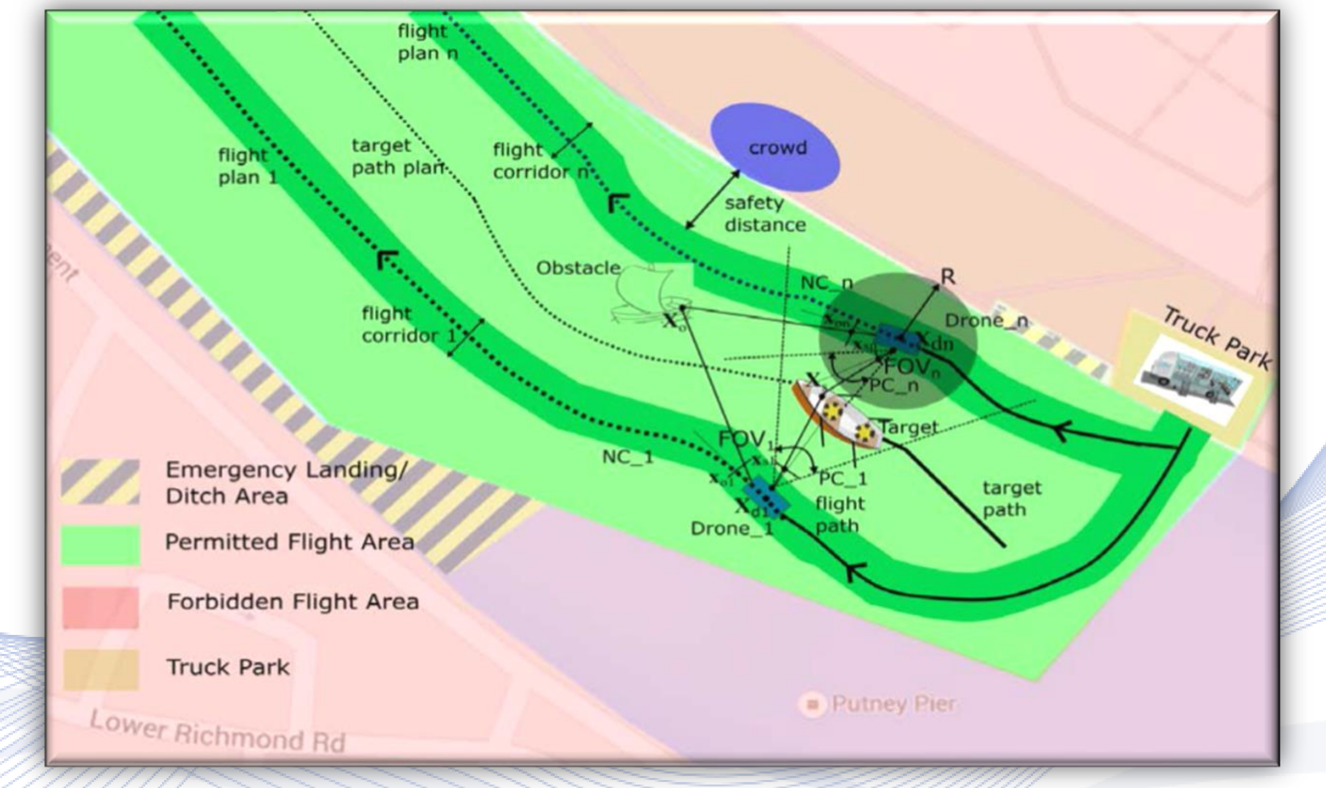
SOCIETY OF AUTOMOTIVE ENGINEERS (SAE) AUTOMATION LEVELS

0	1	2	3	4	5
No Automation	**Driver Assistance**	**Partial Automation**	**Conditional Automation**	**High Automation**	**Full Automation**
Zero autonomy; the driver performs all driving tasks.	Vehicle is controlled by the driver, but some driving assist features may be included in the vehicle design.	Vehicle has combined automated functions, like acceleration and steering, but the driver must remain engaged with the driving task and monitor the environment at all times.	Driver is a necessity, but is not required to monitor the environment. The driver must be ready to take control of the vehicle at all times with notice.	The vehicle is capable of performing all driving functions under certain conditions. The driver may have the option to control the vehicle.	The vehicle is capable of performing all driving functions under all conditions. The driver may have the option to control the vehicle.

Level of automation:
https://www.urbanismnext.org/technologies/autonomous-vehicles

Flight safety

- ***Safety distance from crowds***
  - Crowd detection and avoidance.
- ***Landing sites***
  - Foreseen landing sites;
  - Emergency landing site detection;
  - Person/obstacle avoidance.

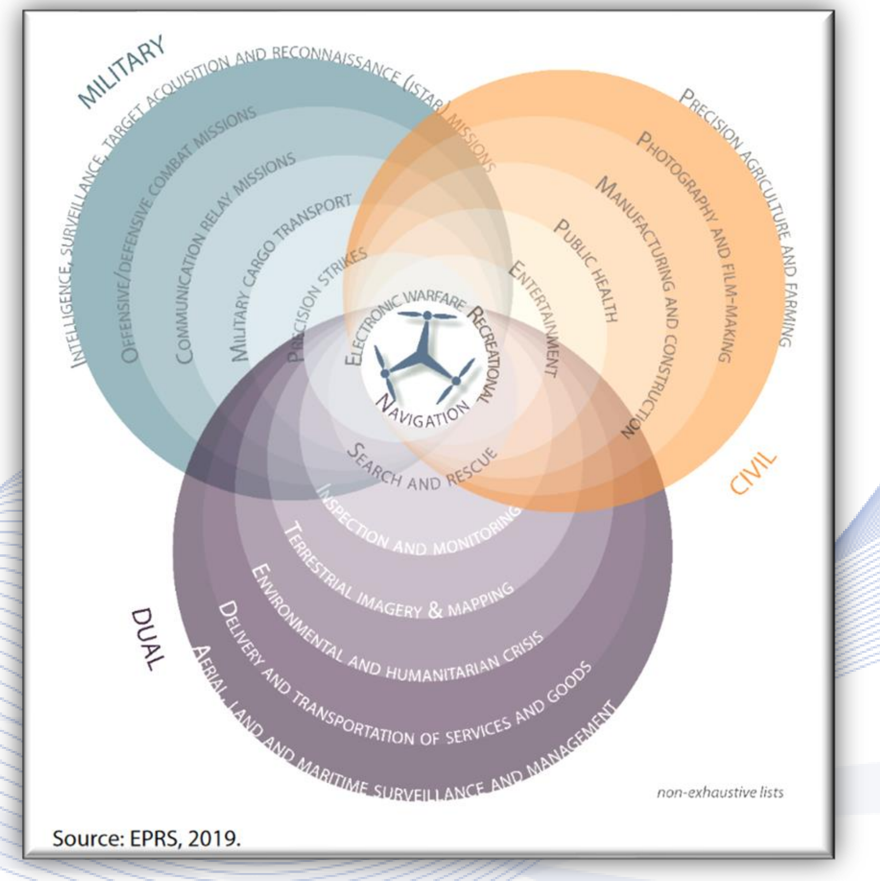


Privacy Protection, Ethics and Regulations for Autonomous Systems

- Data Security
- Privacy Protection
- Moral Machine
- Safety and Regulations
- **Dual use**

Dual Use

Definition: As dual-use products can be defined items, services, and technologies ***that take into account the needs of both defense and civilians***. The EU controls exports, transits and brokering of dual-use items, which makes an effective contribution to global peace and security, avoiding the Mass Destruction Weapons' further proliferation.


MILITARY
Intelligence, surveillance, target acquisition and reconnaissance (ISTAR) missions
Offensive/defensive combat missions
Communication relay missions
Military cargo transport
Precision strikes
Electronic warfare
CIVIL
Precision agriculture and farming
Photography and film-making
Manufacturing and construction
Public health
Entertainment
Recreational
Navigation
DUAL
Search and rescue
Inspection and monitoring
Terrestrial imagery & mapping
Environmental and humanitarian crisis
Delivery and transportation of services and goods
Aerial, land and maritime surveillance and management
non-exhaustive lists

Source: EPRS, 2019.

Most common drone uses [LAT2019]

Designing drone prototypes: Risk mitigation

Key export control documents

- ***WASSENAAR Arrangement*** (41 members): international/regional stability and security of transferred conventional weapons, and dual-use products;
- ***EU export control regime***: governed by EC 428/2009 Regulation;
- ***US Export Administration Regulation (EAR)***: managed by BIS (Bureau of Industry and Security).

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr