EUROCAE
SAE INTERNATIONAL

AAAI CONFERENCE KEYNOTE

EUROCAE WG114 – SAE G34: a joint standardization initiative to support Artificial Intelligence revolution in aeronautics

Speakers:

Christophe GABREAU, Airbus, co-chair of EUROCAE WG-114 Group

Beatrice PESQUET-POPESCU, Thales, co-chair of EUROCAE WG-114 Group

Fateh KAAKAI, Thales, Sub-Group Leader of EUROCAE WG-114 Group

Copyright © EUROCAE & SAE International. Further use or distribution is not permitted without permission from EUROCAE & SAE

Agenda

1. **General presentation of EUROCAE WG114 – SAE G34**
2. **Certification Challenges of Machine Learning**
3. **System considerations**
4. **Machine Learning Development Lifecycle (MLDL)**
5. **Conclusion**

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

General presentation of EUROCAE WG114 – SAE G34

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

Objective & Scope of EUROCAE WG-114

- **Creation: June 2019 (KOM end of August 2019)**

- **Objective: establish common standards, guidance material and any related documents required to support the development and the certification/approval of aeronautical safety-related products based on AI-technology**

- **Scope:**
  - Airborne: Aircrafts and UAS
  - Ground: UTM, ATM and Air Traffic Solution

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

A joint group with SAE G-34 (AI in Aviation)

500+ engineers

Researchers and AI scientists from across the globe, with representation from regulators and authorities (FAA, EASA, TCCA, ANAC, EDA, NASA, DOD, EUROCONTROL), major airframers, UAS/UAM/eVTOL manufacturers, engine manufacturers, component manufacturers, technology providers, and other stakeholders, including operators and airlines

Special thanks to all contributors

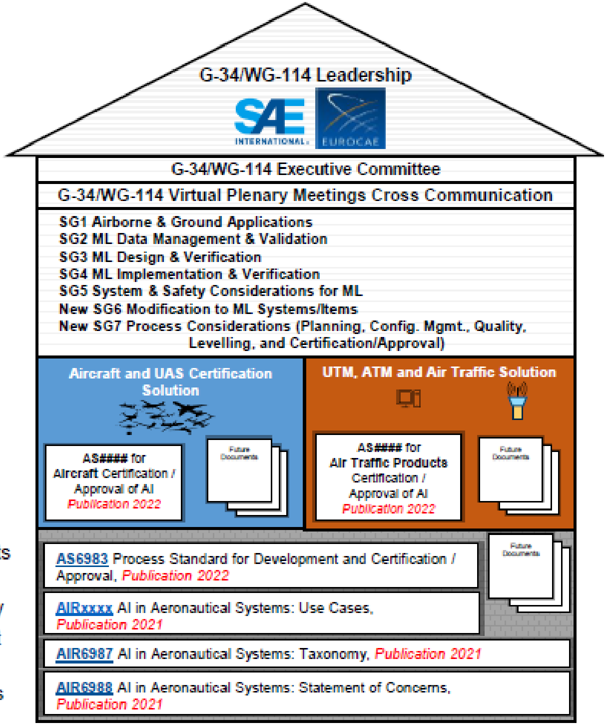
G-34/WG-114 focuses on implementation and certification related to AI technologies for the safer operation of aerospace systems and aerospace vehicles.

G-34/WG-114 (comprised of 500+ members) promotes and standardizes Artificial Intelligence in the entire aviation eco-system (both Airborne and Ground) addressing both manned and UAS.

G-34/WG-114's Global contributors: Boeing, Airbus, ATR, Embraer, Textron, Gulfstream, Dassault, Mitsubishi, Lockheed, Northrop Grumman, GA-ASI, HondaJet, Daher, IAI, ICAO, FAA, EASA, TCCA, ANAC, DGAC, CAA UK, CAA NZ, JCAB, ENAC, FOCA, DOD, EDA, Lilium, Aerion Supersonic, Amazon, DXC, SAP, IBM, Joby, EUROCONTROL, NASA, EDA, Honeywell, Collins, Thales, GE, P&W, RR, Safran, Raytheon, BAE, Elbit, L3Harris, Iridium, Japan Manned Space Systems, FedEx, UPS, AF-KLM, Nodein, Lufthansa, Audi, Toyota, IATA, Leonardo, Leidos, NVIDIA, Intel, Saab, Volocopter, ANSPs, Skyguide, Searidge, Woodward, Vertical Aerospace, Diehl, ADB Safegate, AVSI, ANSYS, BNAE, Copenhagen Airports, D-Risq, Daedalean AI, KIAST, Infosys, Afuzion, Patmos Engineering, QinetiQ, RelmaTech, Rockdale Systems, DLR, drR2, Federated Safety, MathWorks, SRI, Oak Ridge National Lab, etc.

Works In Progress and deliverables:

- AS6983 Process Standard for Development and Certification/Approval of Aeronautical Safety-Related Products Implementing AI
- AIR6987 Artificial Intelligence in Aeronautical Systems: Taxonomy
- AIR6988 Artificial Intelligence in Aeronautical Systems: Statement of Concerns
- AIRxxxx Artificial Intelligence in Aeronautical Systems: Use Cases Considerations

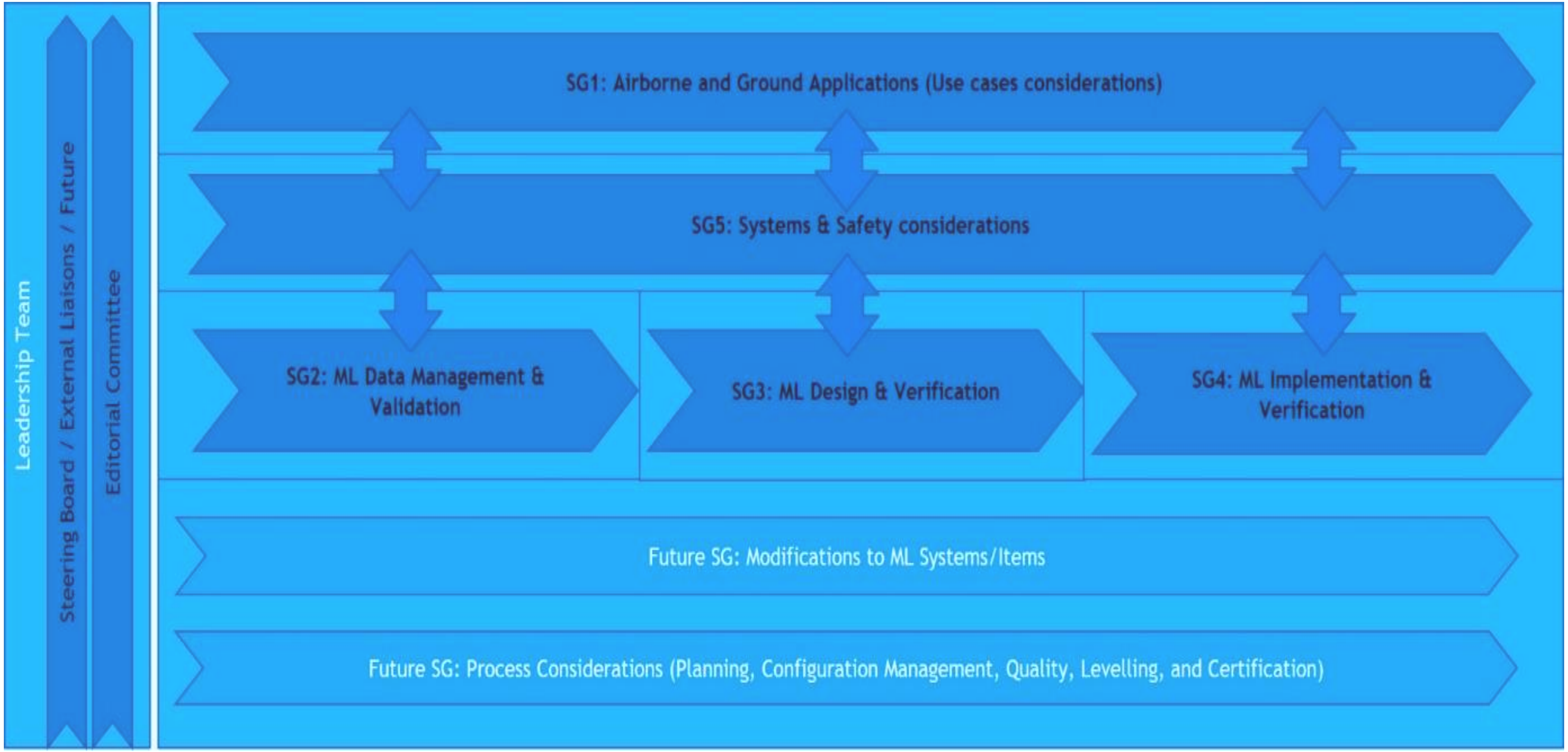


For more information and/or membership registration, contact: jordanna.bucciere@sae.org and/or anna.guegan@eurocae.net.

SAE INTERNATIONAL — Joint International Committee on Artificial Intelligence in Aviation Ecosystem — EUROCAE

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

WG-114/G-34 setup to write the standard



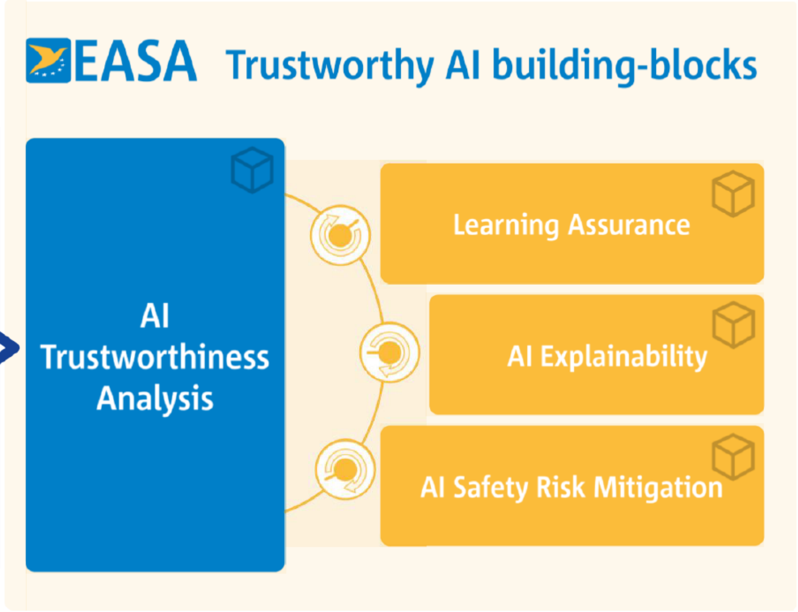
Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

Certification Challenges of Machine Learning

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

EC / EASA Challenges for AI Trustworthiness

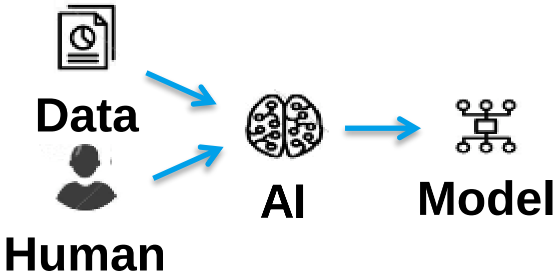
Executive Order issued on February 11, 2019 by the White House on "Maintaining American Leadership in Artificial Intelligence"



Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

1. Specification (and validation) Challenge

ML applied very often to complex problems, difficult to specify (e.g. pedestrian detection)



Data-driven algorithms, implicit model

« Black box »:

- ➢ Difficult to relate SW code to requirements
- ➢ How to specify/verify data requirements ?
- ➢ Quantifying model uncertainties

Trusting an ML model involves « opening the box » to a degree commensurate with its intended use

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

2. Data Challenge : Representativeness

Data Accuracy

Data Integrity

Data Completeness

Data Relevance

Data Traceability

Data Timeliness

Data Consistency

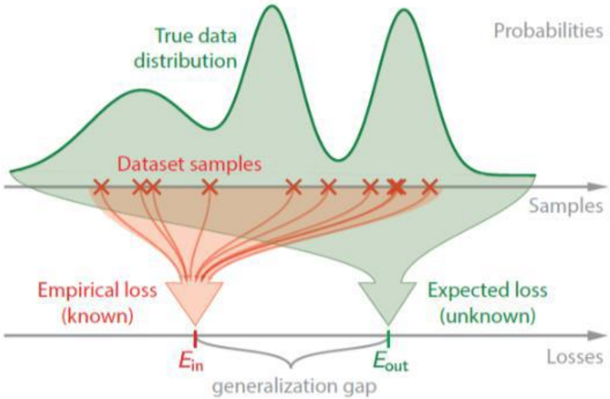
Data Accessibility

Main Data challenges:

- Detection and mitigation of bias and variance
- Dataset quality and completeness
- Change of paradigm: how datasets may contribute to the specification?

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

3. Robustness and Verification Challenges



Source: EASA CodaNN IPC report

- Evaluate robustness of an algorithm to changes in the training set
- Detect unintended and unexpected behavior of NN
- Detect abnormal or adversarial inputs to the NN
- Asses intrinsic robustness of trained ML models through formal or empirical methods
- Assess training methodologies that can enhance or guarantee robustness
- Manage performance / robustness tradeoff
- Define safety process analysis and relevant architectural mitigations (bounding, voting, diversity, etc)

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

4. Explainability Challenges

Specific challenges:

- ➢ **“Black-box” model**
- ➢ **Correlation does not imply causation:**
  - ML models rely on correlation
  - Explanations need causality
- ➢ **Prove the explanation is reliable and correct**
- ➢ **Meaningful explanation for:**
  - Data scientist, SW dev
  - End user (ATCO, pilot, maintenance operator)
  - Regulation authority
  - Accident investigator

(a) Husky classified as wolf (b) Explanation

Link with **Learning Assurance**: high level and low level features

Link with **operational monitoring**: OOD, performance

Link with **Human Factors** considerations

Link with **data recording and traceability** (inputs, internal states, outputs, derived features)

Explanation Accuracy: The explanation correctly reflects the system’s process for generating the output

Knowledge Limits: The system only operates under conditions for which it was designed or when the system reaches a sufficient confidence in its output

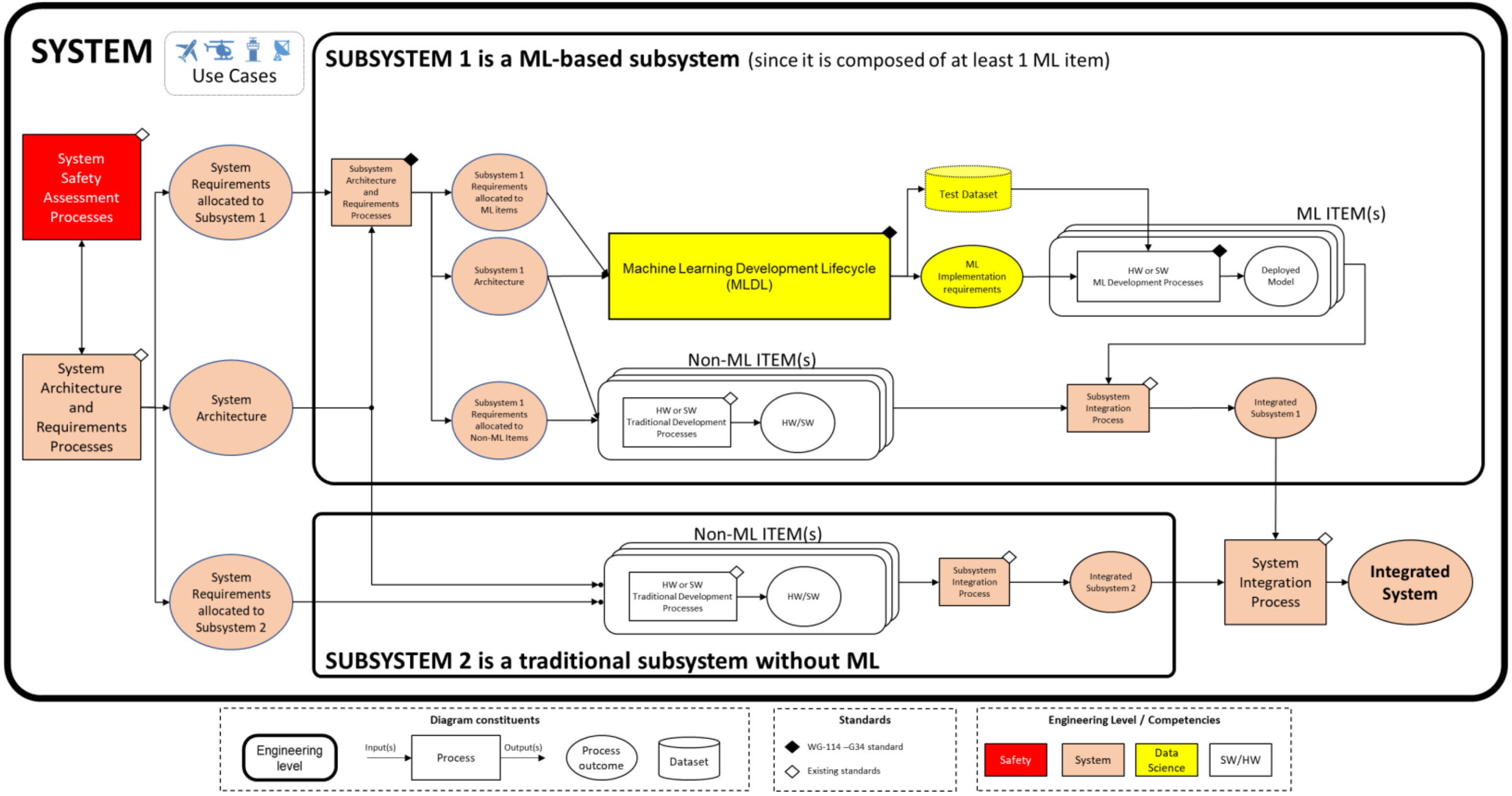
Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

System considerations

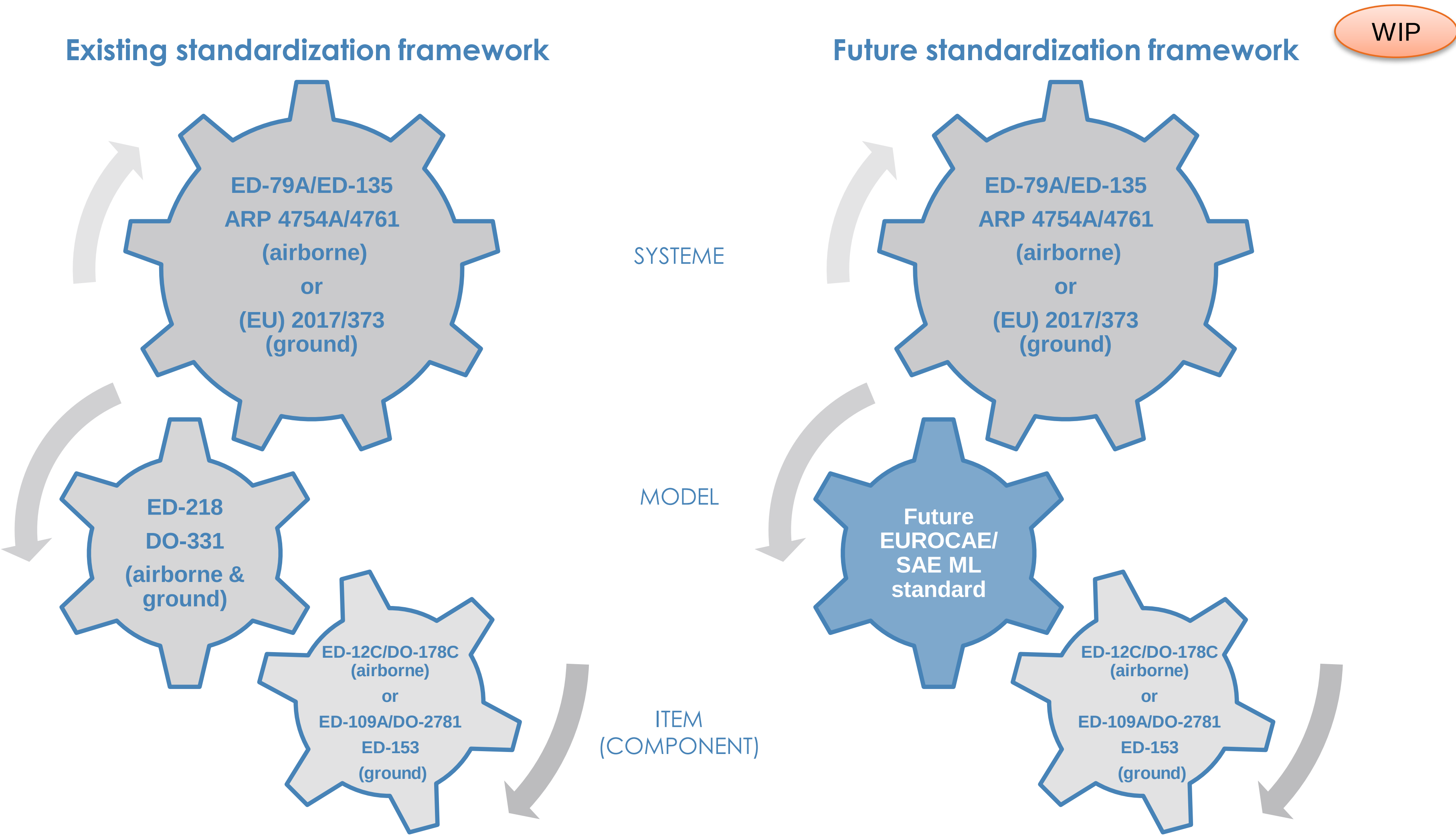
Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

End-to-End System Lifecycle with Machine Learning

WIP



Copyright © SAE International. Further use or distribution is not permitted without permission from SAE


WIP
Existing standardization framework
Future standardization framework
ED-79A/ED-135
ARP 4754A/4761
(airborne)
or
(EU) 2017/373
(ground)
SYSTEME
ED-218
DO-331
(airborne &
ground)
MODEL
ED-12C/DO-178C
(airborne)
or
ED-109A/DO-2781
ED-153
(ground)
ITEM
(COMPONENT)
ED-79A/ED-135
ARP 4754A/4761
(airborne)
or
(EU) 2017/373
(ground)
Future
EUROCAE/
SAE ML
standard
ED-12C/DO-178C
(airborne)
or
ED-109A/DO-2781
ED-153
(ground)

Machine Learning Development Lifecycle (MLDL)

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

Scope and desired attributes of the MLDL

The MLDL should be:

Generic

- The MLDL is applicable to offline ML technologies considered in G34-WG114 scope
- Any technology-specific MLDL phase should be addressed as a second step (further updates of the MLDL)

Process/Environment Agnostic

- The MLDL does not impose a specific development process
- The MLDL does not impose a specific learning environment

Support certification/approval

- ML assurance objectives should be well organized consistently with MLDL
- ML assurance objectives should be simple and clear

Counter-examples

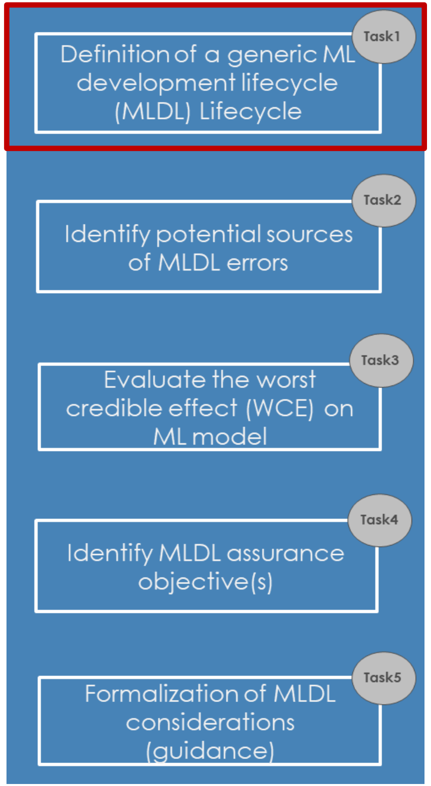
e.g. The MLDL is only applicable to supervised learning using Artificial Neural Network

- e.g. The MLDL is only applicable for V or W development process
- e.g. The MLDL is only applicable to ML models built using tensorflow framework

- e.g. ML assurance objectives are organized using phases and steps that are not consistent with the MLDL definition

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

Methodology to build Machine Learning Assurance Objectives



Task 1 Objective:

Define a generic ML development lifecycle (MLDL) to support:

- the analysis of fault injection all along the ML developement lifecycle
- the identification of ML development assurance objectives (MLDAO) to avoid fault injection or detect resulting errors
- the evaluation of proposed MLDAO with appropriate use cases.

This MLDL should be approved by the full SG3 group

Task 2 Objective:

Identify the possible source of errors called either ML development fault injection cases or ML development failure modes. They are described with at least the following attributes : Name, Rationale (if not obvious)

The completness of the failure modes should be assessed using appropriate method(s).

The list of MLDL failure modes can be classified per MLDL phase and should be approved by the full SG3 group

Task 3 Objective:

Study the worst credible effect (WCE) on the ML model of all ML development failure modes. The adverserial effects that are considered to establish WCEs come from SG5 safety objectives (e.g. impact on ML model integrity, performance, explainability, etc.). When not obvious, a rationale should be provided to explain WCEs. When there is no adverserial effect on safety, the WCE should be « No identified effect ».

Task 4 Objective:

Identify MLDL assurance objective(s) to mitigate any adverserial WCE on SG5 safety objectives allocated to the ML model (e.g. adverserial impact on ML model integrity, performance, explainability, etc.). MLDL assurance objectives should be classified by DAL/AL/SWAL levels. A gradation of these assurance objectives is expected according to the DAL/AL/SWAL levels. Airborne and Ground specificities should be taken into account. A rationale should be provided to explain each MLDL assurance objective. When there is no adverserial effect on safety (i.e., WCE = No identified effect »), no assurance objective is needed.

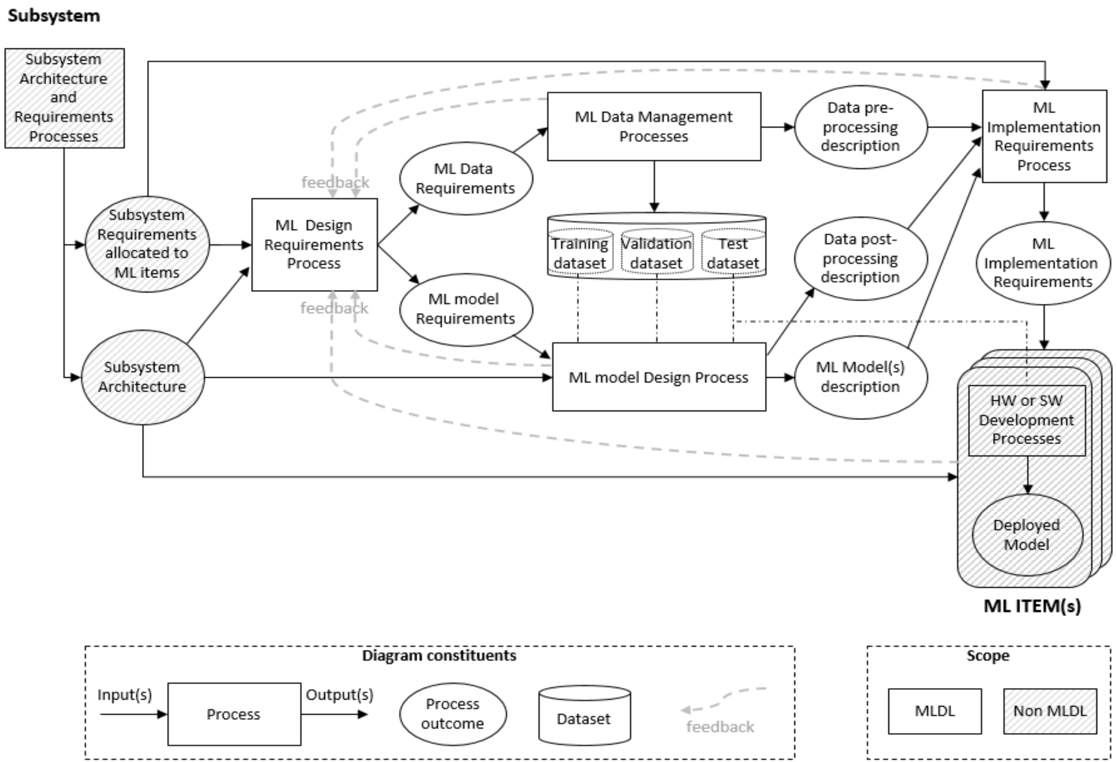
Task 5 Objective:

Formalize the outputs of all tasks into a guidance material that follows AS6983/ED-XXX Outline. This guidance is expected to be part of the final AS6983/ED-XXX standard. The need to issue a FAQ should be assessed by SG3 leaders.

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

Machine Learning Development Lifecycle (MLDL)

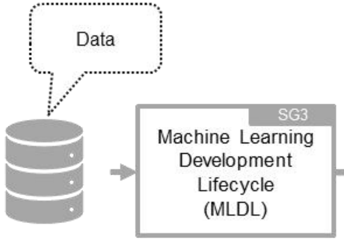
WIP



Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

Outcome of the MLDL to implementation phase

WIP



Pre-processing requirements

1rst layer

- Recover the entries in the form of a vector of dimension 3
- Multiply this vector by the matrix:
  [[1.8514004 , 0.06981966, 1.2454928 , -0.23673481, 0.33582166, 1.144316 , -0.07928362, 0.85065633], [0.2689074 , -1.6337383 , -0.09429459, -1.5041184 , -0.8323707 , -0.19164966, 1.7850695 , -1.332562], [0.4509967 , -0.6398565 , 0.30318516, -0.5797693 , -0.28540373, -0.11622415, -0.93806374, -0.36459136]]
- Add the vector:
  [0.0436854 , 0.66848093, -0.34943226, 0.82454354, -0.941798 , -0.1605606 , 0.308173 , -0.01794778]
- For each element, apply the **ReLU** function : (x) -> (x si x >0, 0 else)
  (NB: the result of the 1st layer)

2nd Layer

- Multiply by the matrix:
  [[0.08302072 -0.34755808 0.3101814 -0.29662135 -0.16163935 0.09059773 0.16717902 0.74825156] [-0.10736815 1.0281937 -0.11446434 0.237721 0.7667631 -0.07044937 0.03289383 0.04712383] [0.62960064 0.60489106 0.40853637 0.3820236 -0.02992251 0.40818375 0.71154916 0.6853567] [0.27913433 0.57993895 -0.1088046 -1.4035206 0.05937768 0.19022994 0.33831477 0.5348654] [-0.49616838 -0.59313494 -0.13610965 -1.602424 -0.8092544 -1.4723336 -0.74360555 0.14003202] [0.9558404 -0.38831693 0.54558367 1.1563224 -0.1213242 0.96744263 -0.04928763 -0.10160095] [0.29998606 -0.5457609 0.34783924 0.2050478 -0.02698133 0.3551015 0.19448055 0.9439171] [-1.1834544 -0.73129165 -0.83786845 0.25230122 -2.1236877 -0.6258286 -0.9858019 0.1222824]]
- Add the vector:
  [-0.33079636 0.09964079 -0.06536419 -0.02690533 0.05694228 -0.24229142 -0.18812867 0.8922489]
- Apply the **ReLU** function to each element : (x) -> (x si x >0, 0 else)
  (NB: result of the 2nd layer)

3rd Layer

- Multiply by the matrix :
  [[-0.3034037 -0.11246069 -0.6560938 1.2929432 0.9277852 0.79638636 -0.6643924 0.10672873] [-1.6083356 0.25883913 -1.0975031 1.3155658 1.5284146 1.7666548 -0.9037737 1.9760858] [0.11462954 -1.0091568 -0.4328157 1.0937432 1.1971915 1.2098296 -0.91365033 1.0630316] [-1.0262014 1.3004675 -0.61168474 1.7266848 0.58779174 1.0225899 -0.14789295 0.9648] [-2.7417295 -0.02741075 -2.4579542 3.2859435 3.4163947 3.6450598 -2.7008405 3.6044834] [-0.60329115 0.45157087 0.17973106 0.9708737 0.45990798 0.10065918 0.10000039 0.97939295] [-0.09893946 -0.7206996 -0.18443836 0.5532775 0.8732048 0.31475598 -0.07816552 -0.08506049] [0.43368313 -0.47053978 0.8148482 -0.35335994 -0.3584372 -0.27395764 0.4629211 -0.24606809]]
- Add the vector :
  [1.4406947 -0.17694788 0.99267095 -0.24365164 -0.20841813 -0.2159943 1.0041207 -0.40997046]
- Apply the **ReLU** function to each element : (x) -> (x si x >0, 0 else)
  (NB: result of the 3rd layer)

- Multiply by :
  [[3.0001943] [-1.1565807] [1.3064175] [-1.6839781] [-2.2671492] [-2.198915] [1.8382062] [-2.066988]]
- Add :
  [0.9438011]
- Apply the **tanh** function to the result (*=Output of the network*)

Example source: NN and DO331 - C. Travers/A-C Kerbrat/T. Ther (Dassault Aviation) 31/01/2020

Post-processing requirements

Implementation constraints (performance, safety, etc.) coming from system considerations

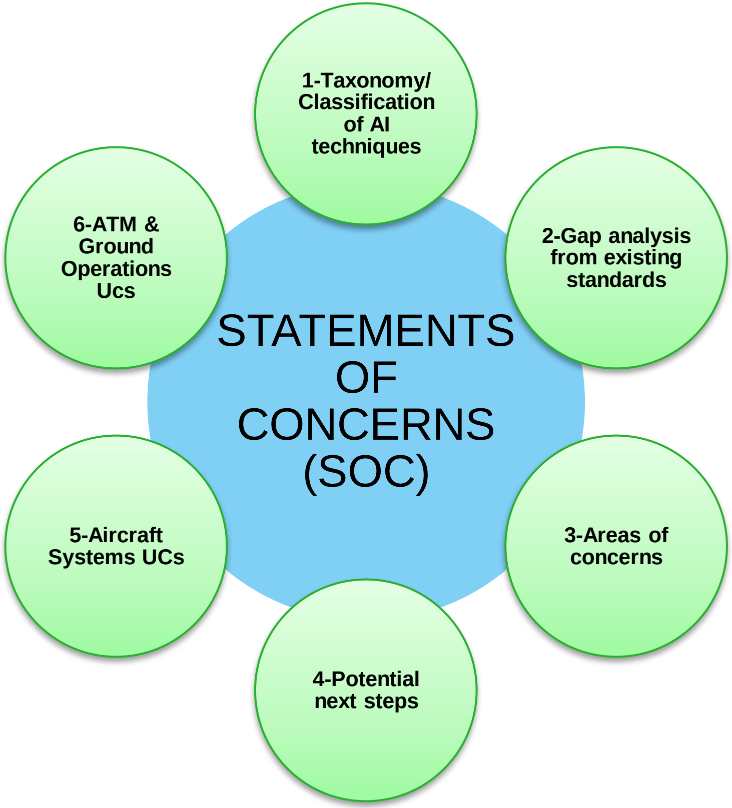
Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

Conclusion

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

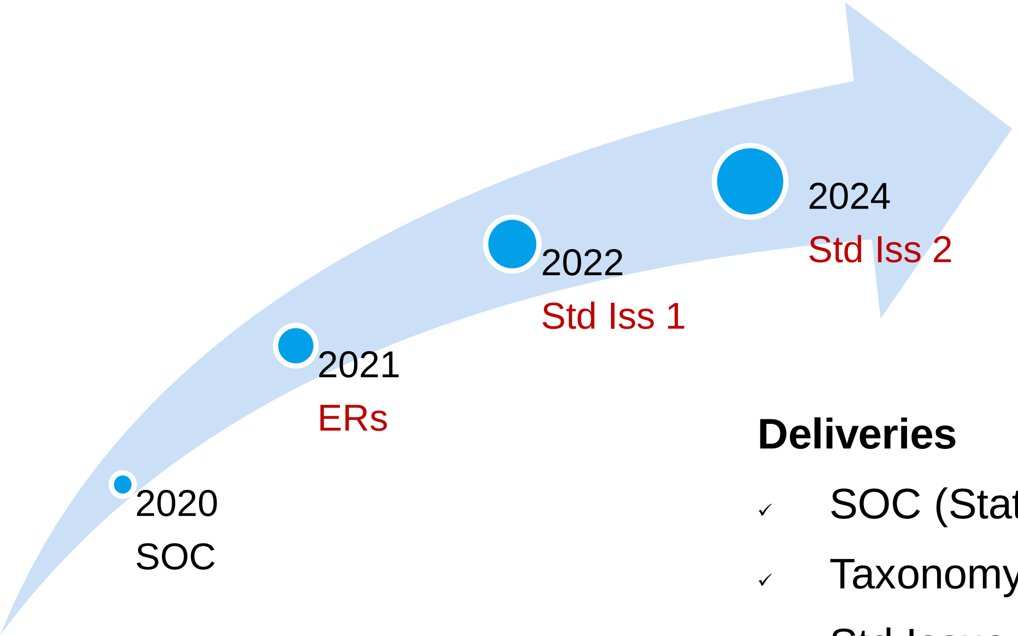
2021 Outcomes: Statement of Concerns

Worldwide industries aligned on the same concerns



Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

WG-114/G-34 Roadmap



Deliveries

- ✓ SOC (Statement of Concerns) – ER/AIR
- ✓ Taxonomy, Use Cases – ER/AIR
- ✓ Std Issue 1: ML (Offline Learning) – ED/AS
- ✓ Std Issue 2: Other AI Technologies – ED/AS

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

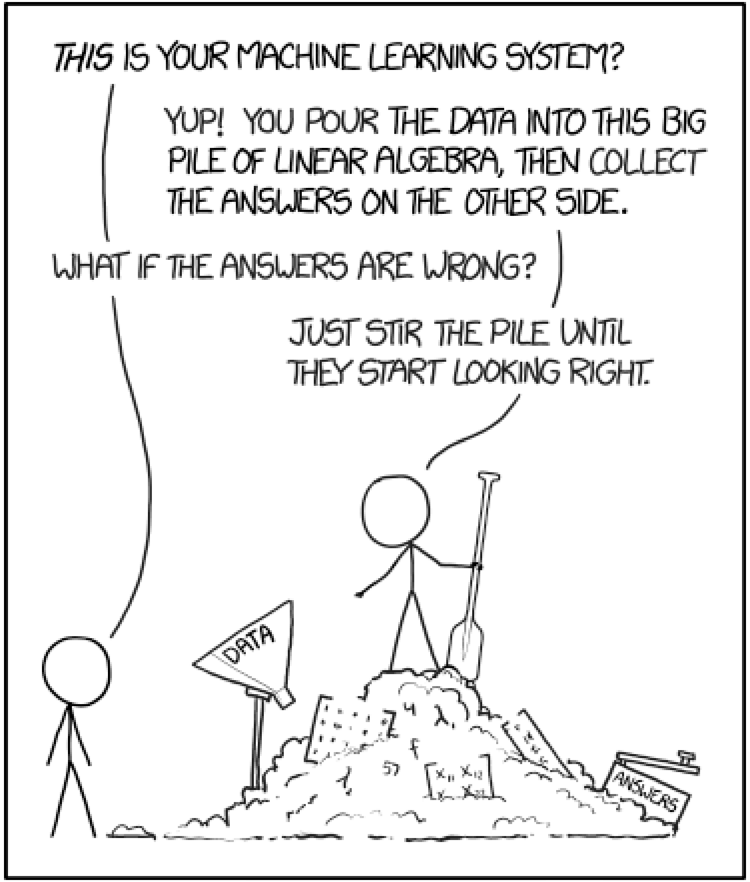
Liaisons with other Groups
(*) active ones are bolded

- EUROCAE
  - **WG-63 (Complex A/C systems)**
  - **WG-72 (Aeronautical Systems Security)**
  - WG-105 (Unmanned A/C Systems - UAS)
  - WG-112 (Vertical Take-Off and Landing – VTOL)
  - WG-117 (Topics on SW advancement)
- SAE
  - **S-18 (Complex A/C systems & UAS Autonomy)**
  - **G-32 (Cyber Security)**
- Others
  - **AVSI - AFE87**
  - **EUROCONTROL AI High Level Experts Group**
  - **ISO/IEC JTC 1/SC 42**
  - French Grand Defi CONFIANCE.AI
  - JARUS (Joint Authorities for Rulemaking on Unmanned Systems)
  - ASTM

Copyright © SAE International. Further use or distribution is not permitted without permission from SAE

THANK YOU FOR YOUR ATTENTION !

Questions ?



Source: https://xkcd.com/