DEEL

DEpendable & Explainable Learning

DEEL Project

DEEL
DEpendable & Explainable Learning

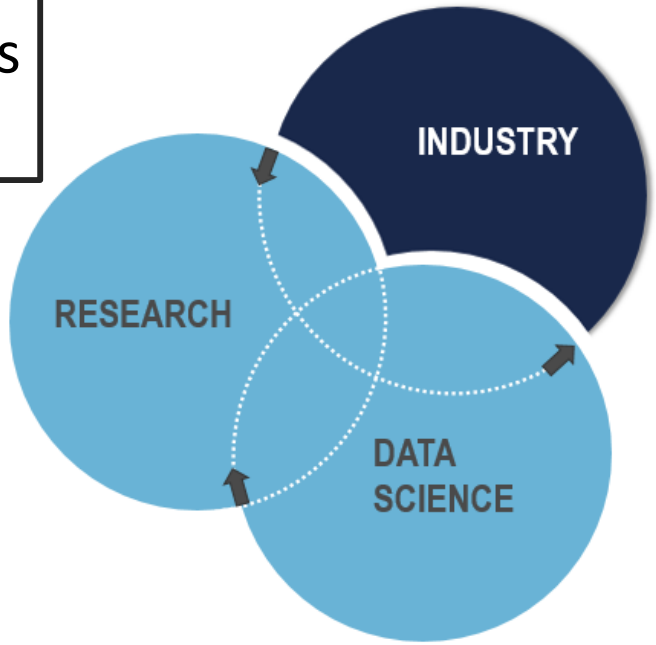
The project will provide industrial partners with the artificial intelligence (AI) tools and technological bricks that enable them to **secure and certify** in a short time the development of their critical systems integrating AI functions.

Main topics:

- Explainability,
- Robustness,
- Fairness,
- Certificability,
- Privacy

Toulouse

Montréal



26 M€

5 ans

24 partners

60

DEEL project

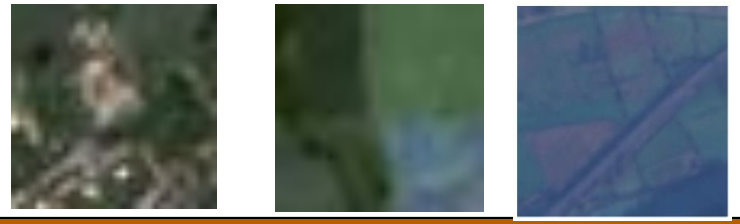
DEEL
DEpendable & Explainable Learning

Fairness challenge objectives:

- removing bias from training (training data, unfair decisions)
- effects of the learning sample in the ML process use it to improve training datasets
- collaborative training with separate and secret datasets

Deliverables:

- methodology, tools, metrics to measure bias
- algorithms to correct bias in models
- benchmark with industrial applications,



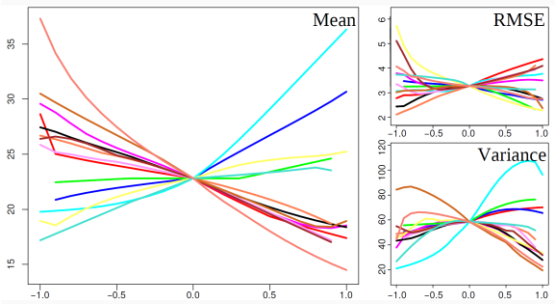
M. Serrurier, Jean-Michel Loubes, et E. Pauwels, « Fairness with Wasserstein Adversarial Networks », 2019.
Risser, Vincenot,Couellan, Loubes, 2019 https://arxiv.org/abs/1908.05783

Explicability challenge objectives:

- User concepts for explainability toolbox (industrial and certification level)
- Explainability of black box models
- Stability of models and interpretability
- Metrics

Deliverables:

- Methodology, tools, metrics for explanation
- Algorithms to explain black box models
- Turorials on industrial use cases



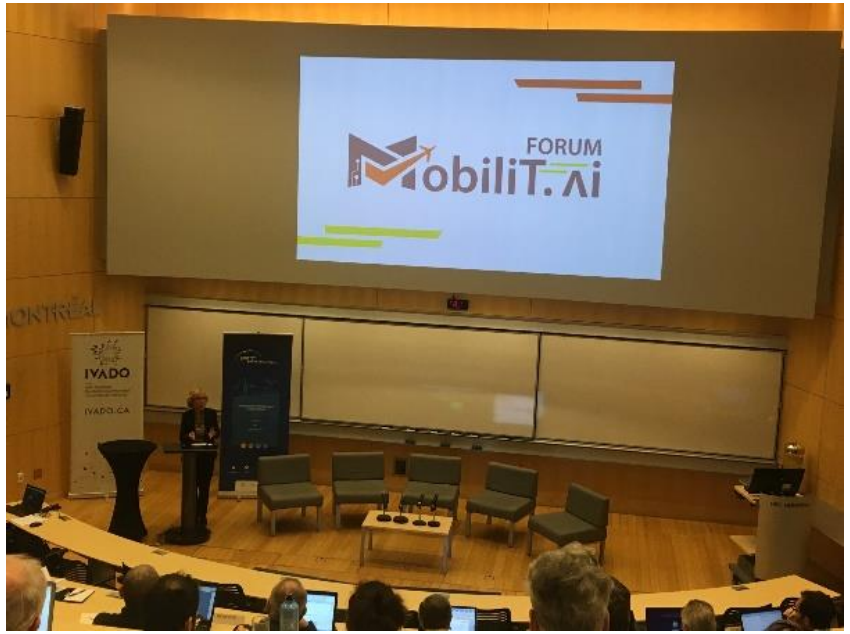
https://github.com/XAI-ANITI/ethik

« Certification of AI » mission presented yesterday (paper We.1.B.3)

FORUM MobiliT.AI

26 & 27 MAI 2020

TOULOUSE



www.mobilit.ai

DEEL
DEpendable & Explainable Learning

THANKS!

IVADO IRT SAINT EXUPERY CRIAQ ANITI ARTIFICIAL & NATURAL INTELLIGENCE TOULOUSE INSTITUTE