LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

VISUALISATION AS A MEAN TO FOSTER ADOPTION OF AI BY POLICY MAKERS

Dr Ir Benoît Otjacques
Luxembourg Institute of Science and Technology
benoit.otjacques@list.lu

Applied Machine Learning Days 25-29 January, 2020, EPFL, Lausanne

PLAN

- Introduction
- Adoption of AI
- AI & Visualisation
- Conclusion

Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

Introduction

- Do we need AI?
  Yes, in some domains, for some specific tasks
- Do we want a wider use of AI?
  Yes, where it helps
- Does it require a larger acceptance of AI?
  Yes, AI shouldn't be imposed to citizens and customers against their will
- Do we need the support of policy makers to foster AI development?
  Yes, because AI has become a topic of public interest

What can be the **Visualisation** contribution in this domain?

Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

Acceptance of AI

Lasting large-scale **use** of a technology requires wide **acceptance** by citizens, customers, public bodies, business actors, regulators…

Same probably applies to AI

Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

Acceptance of AI

Acceptance ≠ Understanding

Most of the people taking a flight don't understand aerodynamics, turbines technology or the redundancy principles of embedded systems present on an aircraft…

and it is not a problem!

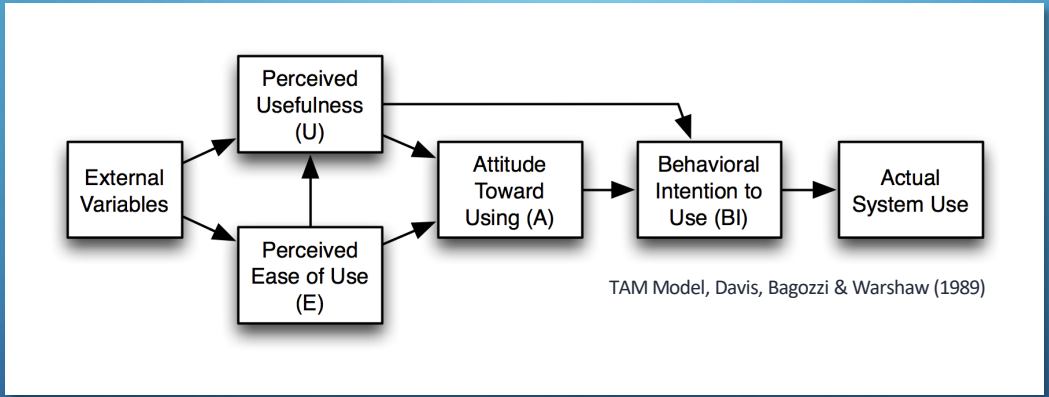
Same probably applies to AI

You don't need to understand how a technology works to use it

Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

Acceptance of AI



Understand how the technology works is not explicit in the TAM model

- Objective usefulness and ease-of-use is not enough
- Perception plays a critical role too
- No wide acceptance of a technology without **trust**
- No trust if people fear that they will face negative and/or unfair consequences

Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

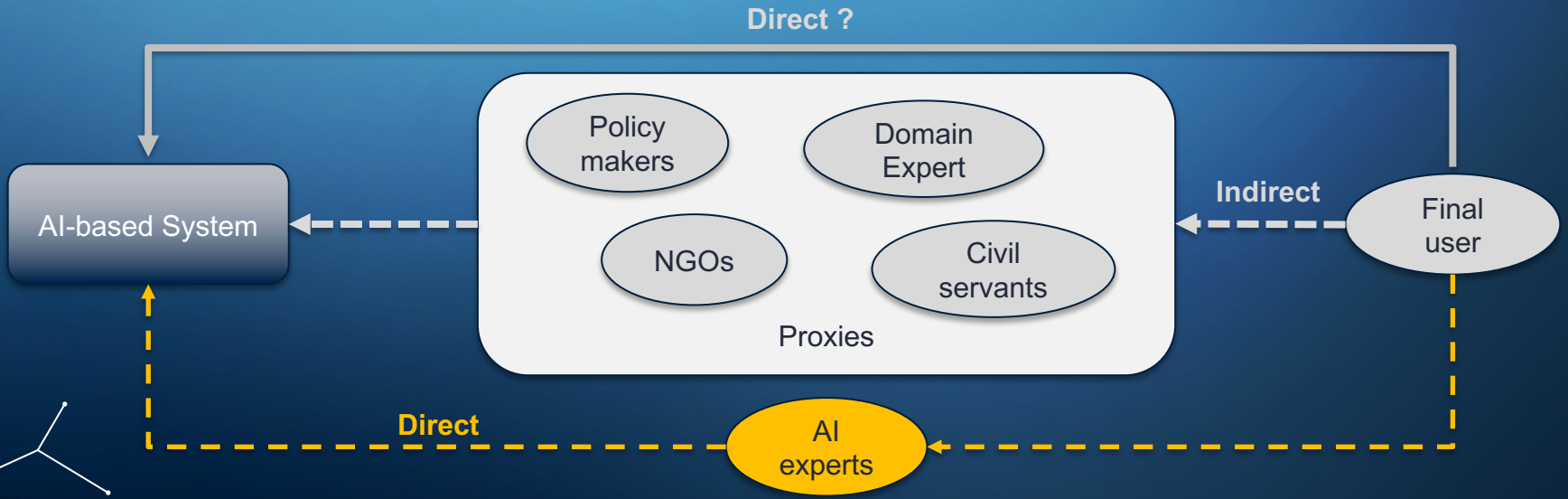
LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

Acceptance of AI

Direct vs. Indirect Trust

- Direct: I use this AI-based system because **I personally understand** how it works and I trust it (requires AI Literacy)

- Indirect: I use this AI-based system because **I trust some persons** who understand how it works (requires the set up of a chain of people and processes, i.e. trustworthy proxies)

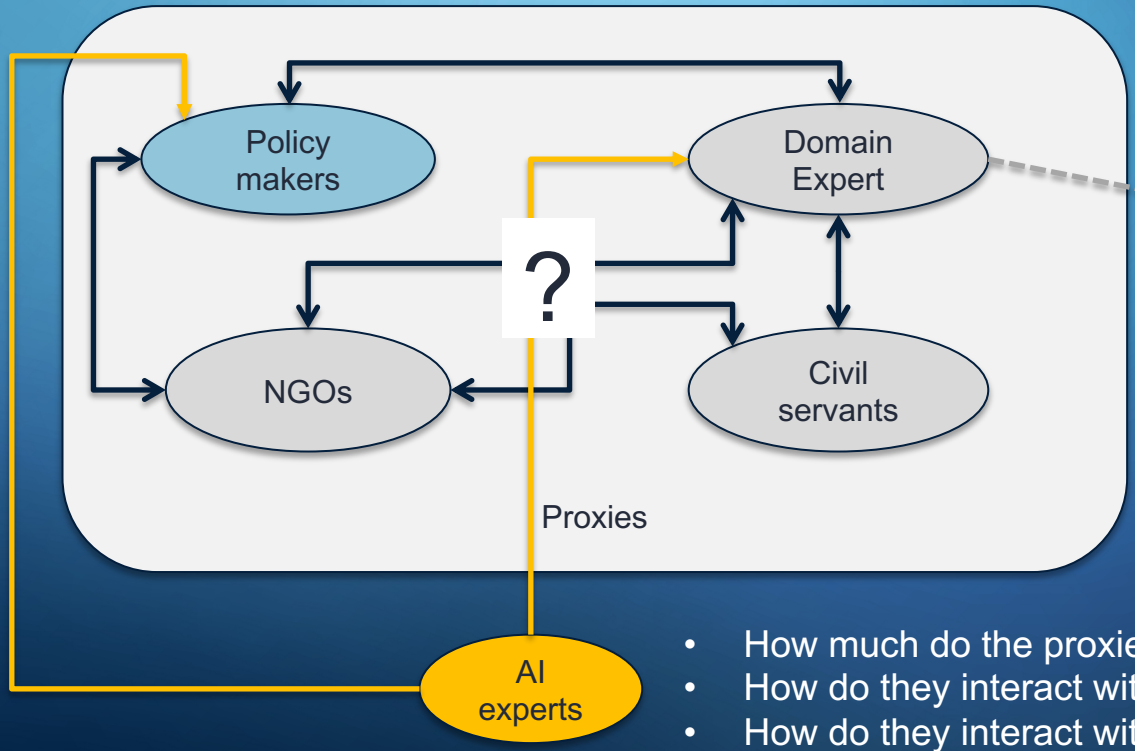


Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

Acceptance of AI



- Medical Doctors
- Lawyers
- Journalists
- Automotive Engineers
- Agronomists
- Bankers
- Energy Experts
- Marketing Professionals
- ….

- How much do the proxies understand AI?
- How do they interact with each other?
- How do they interact with the AI experts?

How can **Visualisation** be useful in this context?

Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

AI & Visualisation

Visualisation for ML is currently studied for various purposes:

- to better identify biased data
- to better choose the most appropriate type of model
- to better fine tune models
- to trace back the rationale of a specific output of a model
- to better explain the overall behaviour of a model…

Proxies persons are a key target population of AI Visualisation

Benoît Otjacques, benoit.otjacques@list.lu

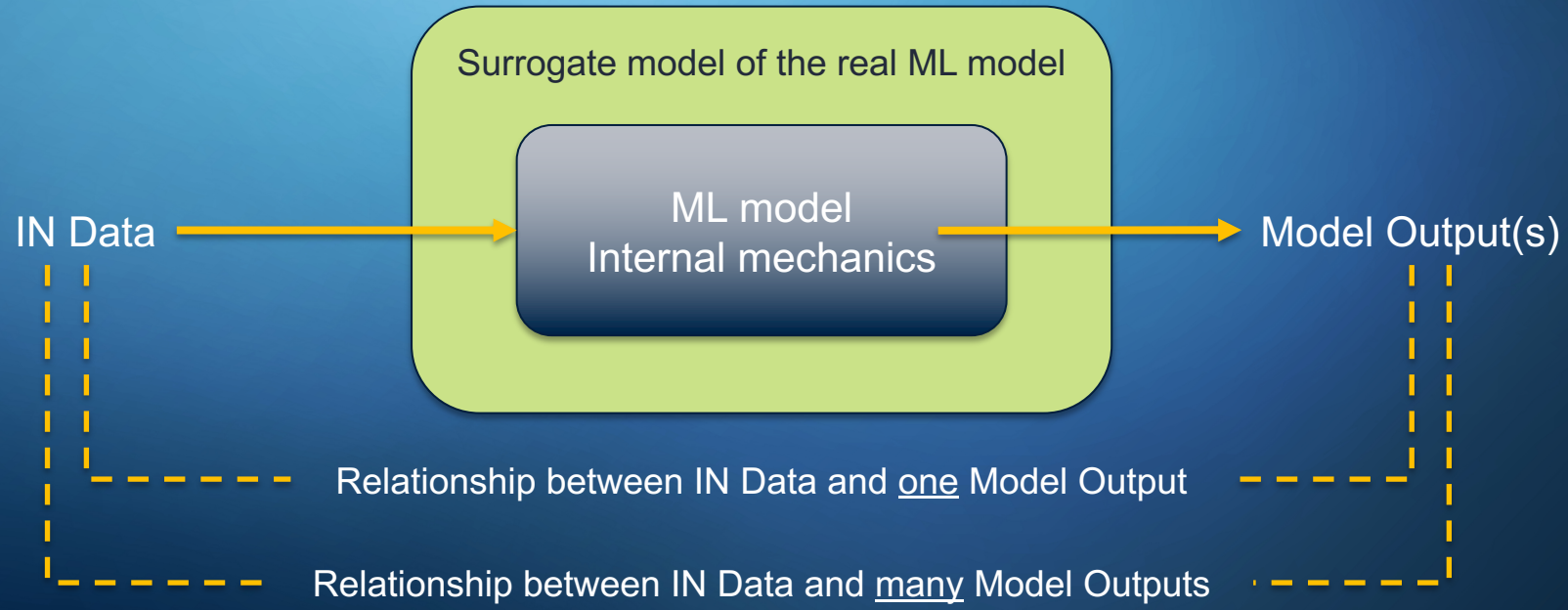
*Focus on ML (as a subtopic of AI)

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

AI & Visualisation

Visualisation: { what, for whom, why}



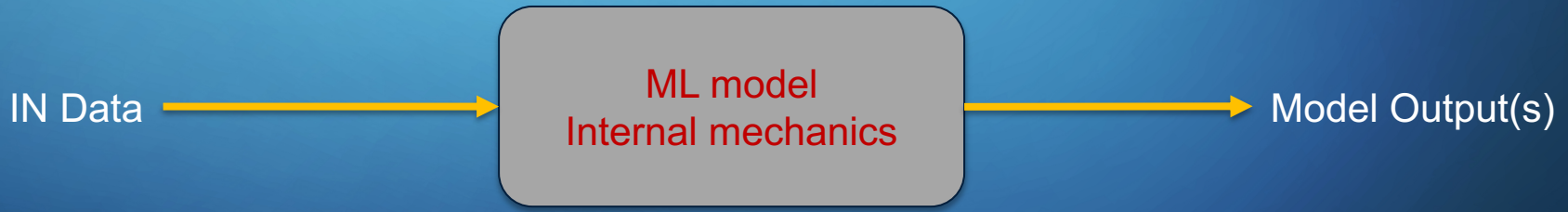
Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY LIST

AI & Visualisation

Visualisation: { what, for whom, why}



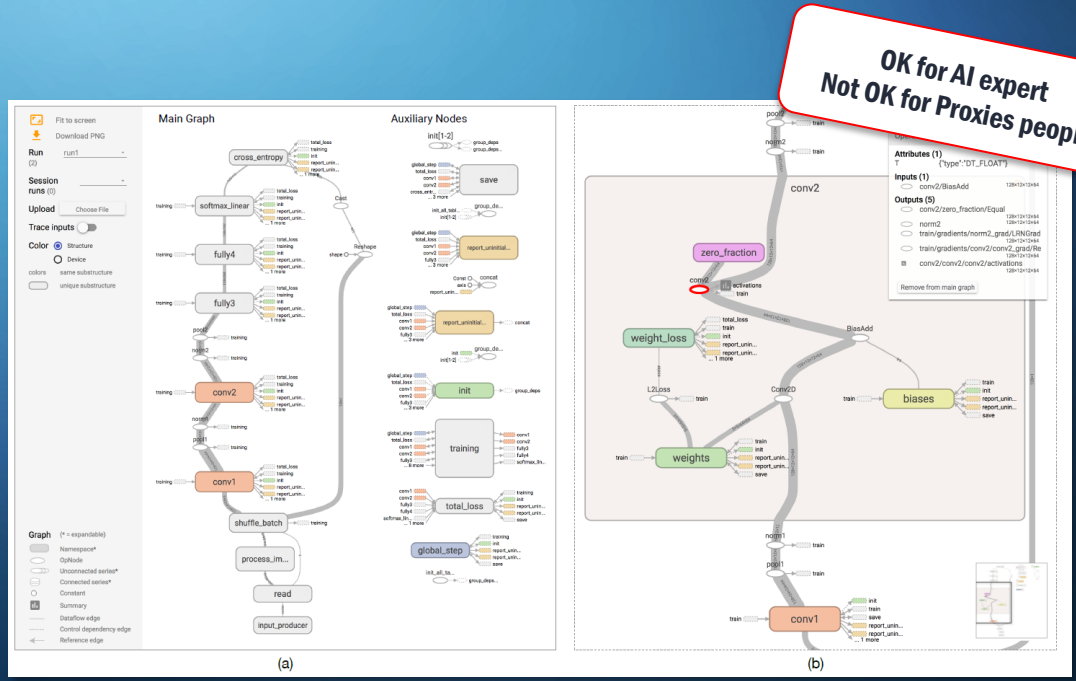
Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

AI & Visualisation

- Visualise the internal mechanics of a Machine Learning model

Helping to understand complex ML architectures by visualizing their underlying dataflow graphs



TensorFlowGraph Visualizer[1], 2018

Benoît Otjacques, benoit.otjacques@list.lu

[1] IEEE TCVG 24(1), 2018: Wohngsuphasawat K et al. Visualising Data Graphs of Deep Learning Models in Tensor Flow

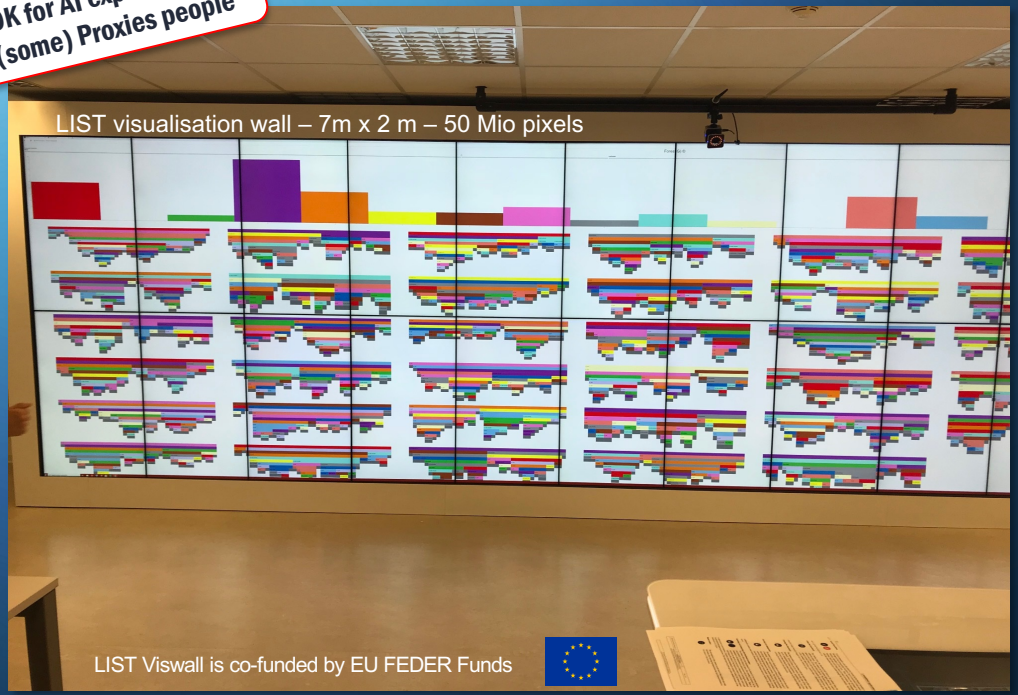
LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

VISUALISATION & ADOPTION OF AI

AI & Visualisation

- Collaborative large-scale visualization of ML models
  - To support data scientists to develop and improve models
  - To support domain experts to understand models
  - To explain ML models to domain experts…

OK for AI expert
OK for (some) Proxies people



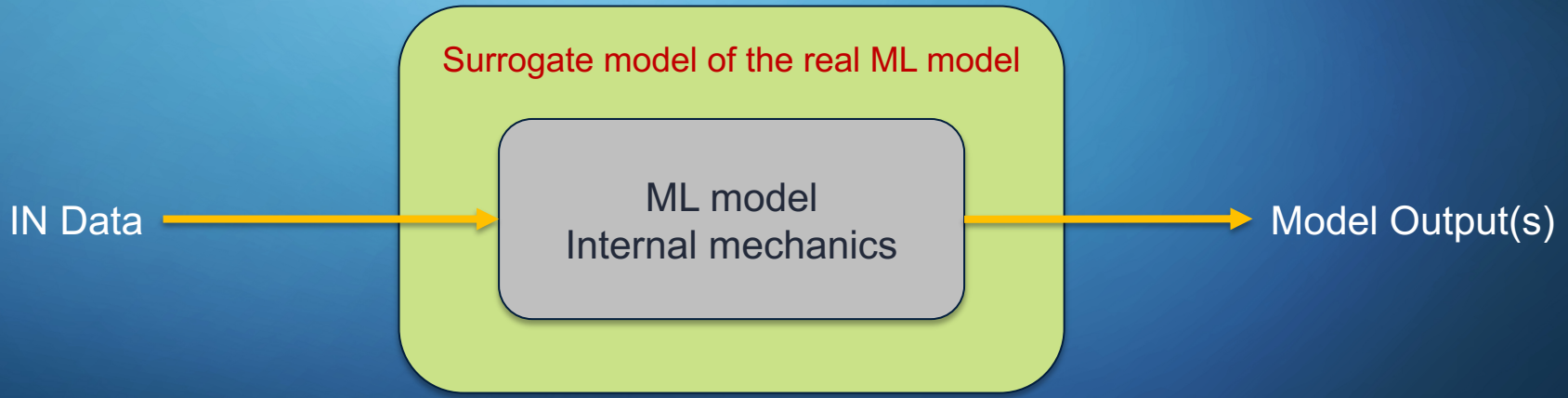
Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY LIST

AI & Visualisation

Visualisation: { what, for whom, why}

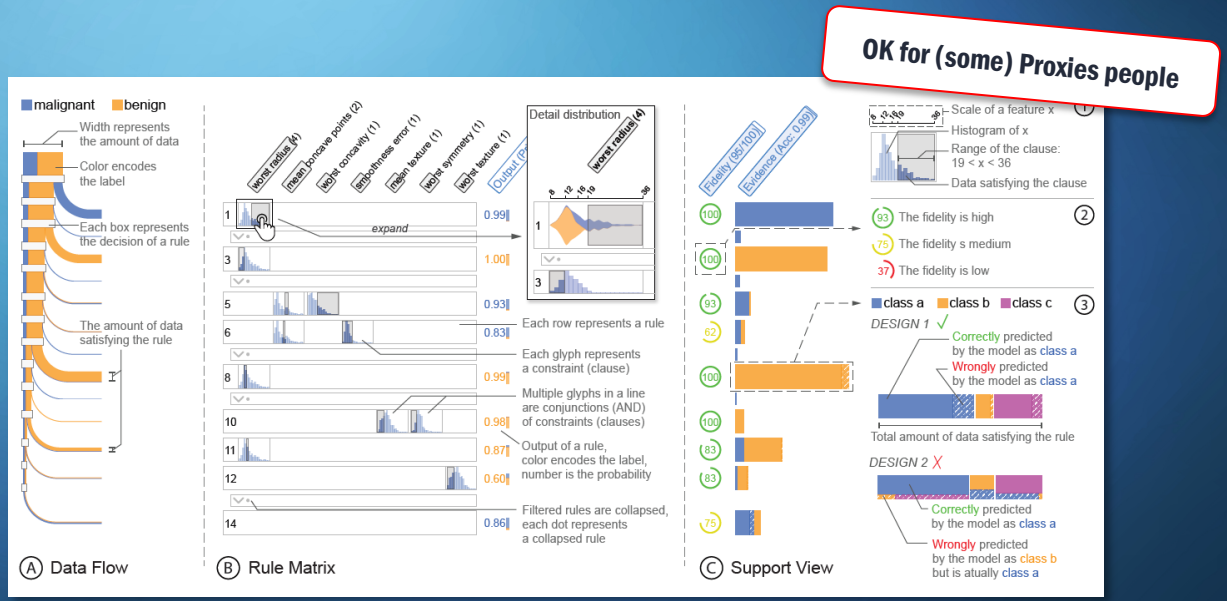


Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

AI & Visualisation

Visualise a surrogate model



Rule Matrix[2] , 2019.

- Supporting domain experts to understand ML models
  - Visualise the behavior of a Machine Learning model with another metaphor (Neural Network visualized as Rules-based approach)

[2] IEEE TCVG 25(1), 2019: Yao Ming, Huamin Qu, and Enrico Bertini RuleMatrix: Visualizing and Understanding Classifiers with Rules

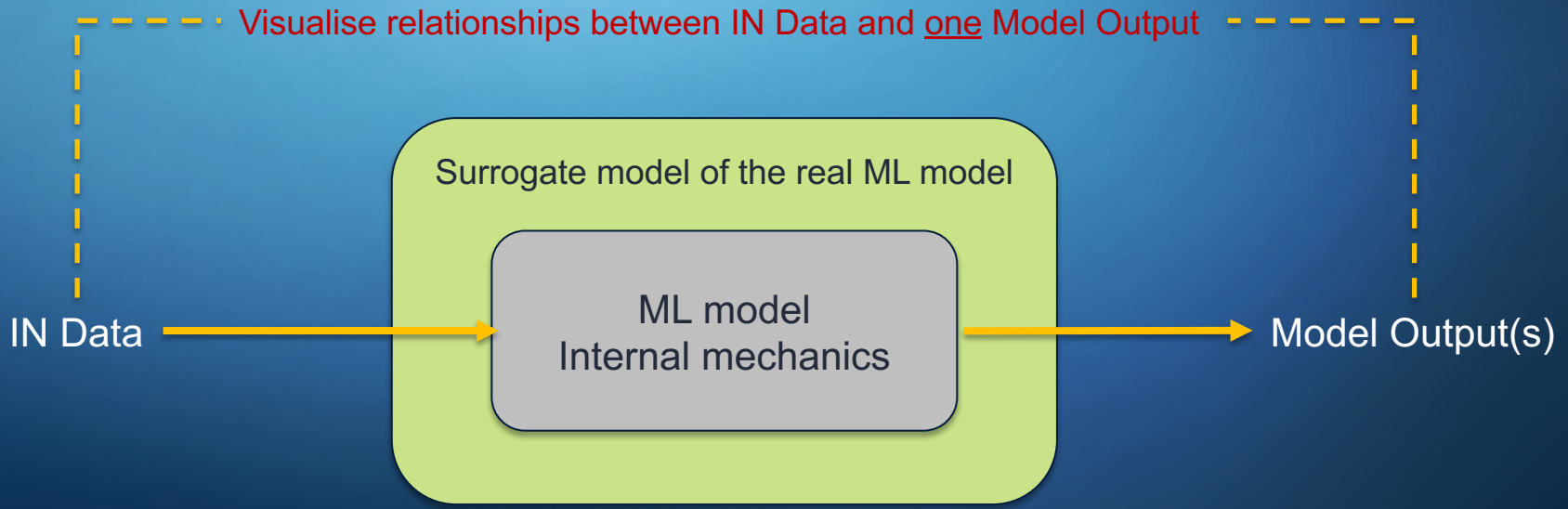
Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

AI & Visualisation

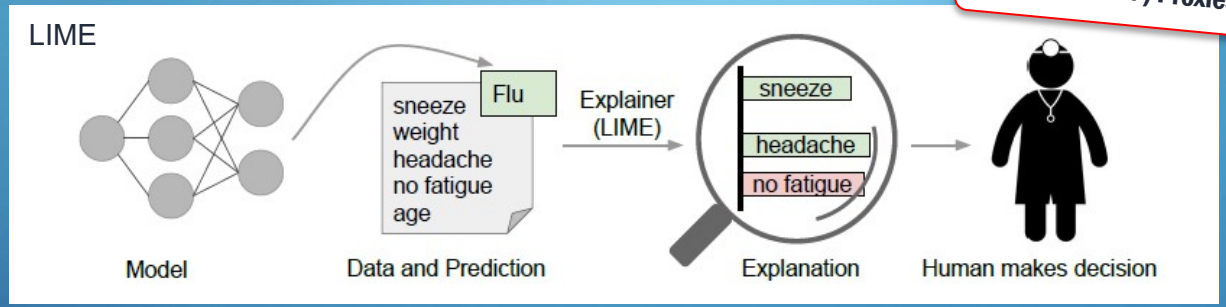
Visualisation: { what, for whom, why}



Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

AI & Visualisation



OK for (more?) Proxies people

Ribeiro et al., KDD 2016

- Supporting domain expert to understand the specific result of a ML model
  - Visualise on which basis a ML model has given this answer for this case
- Does the user sufficiently trust a specific prediction to accept to take action on this basis?

Domain Expert combines his/her (tacit) knowledge and/or experience with the result of ML system (visually displayed as an histogram) and takes decision on this basis

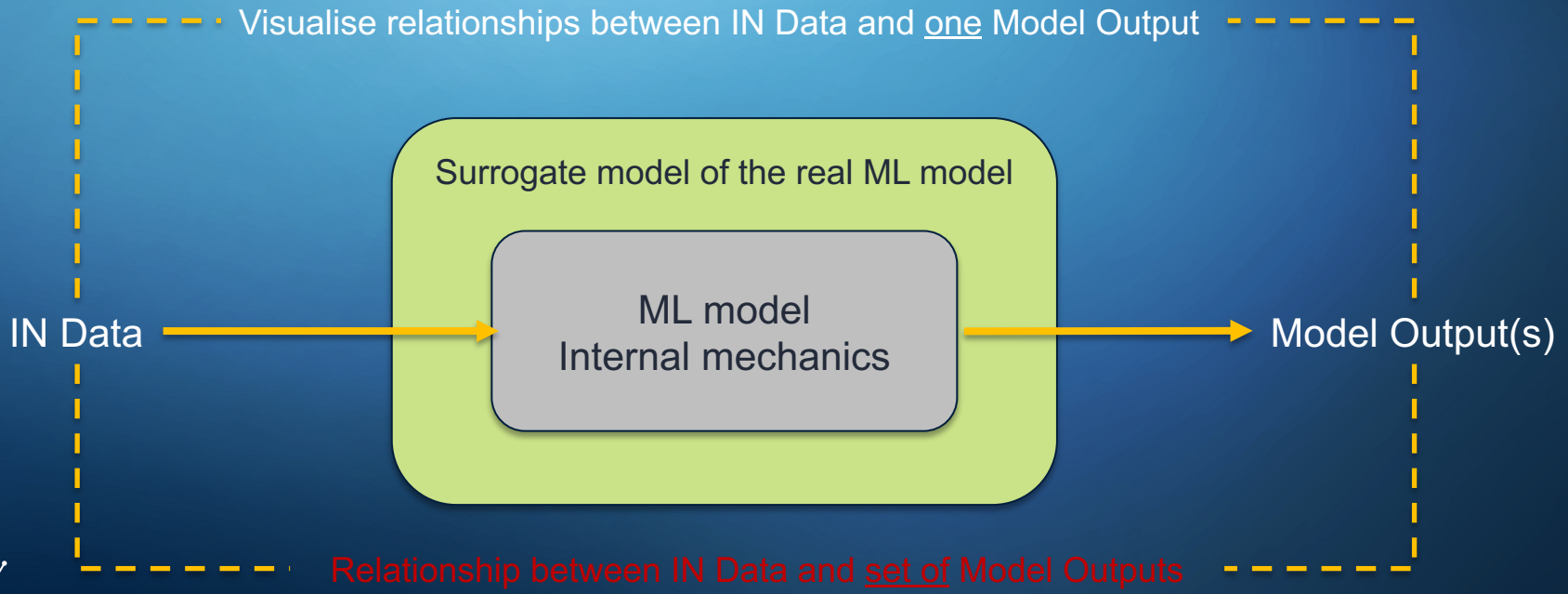
Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

AI & Visualisation

Visualisation: { what, for whom, why}



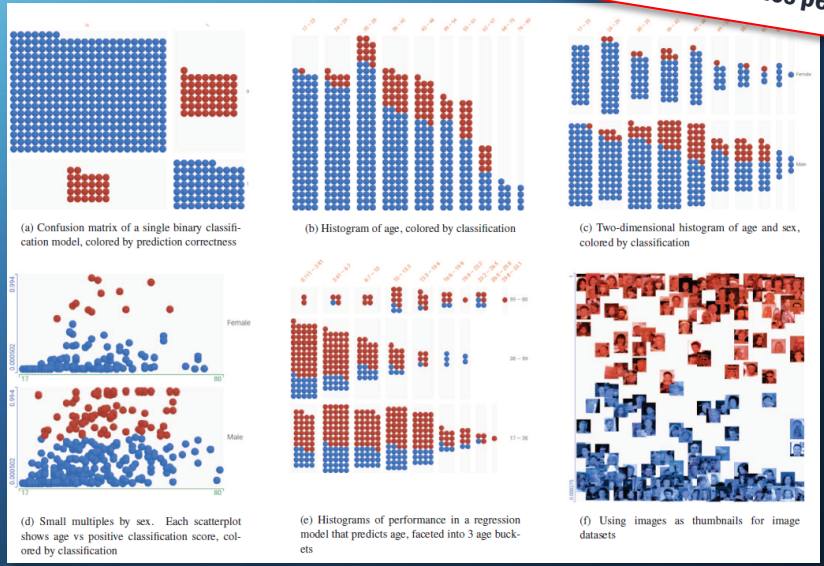
Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

AI & Visualisation

- Visualise how a model performs with various input and outputs
- Show how perturbations in input affect the output (what-if scenario)
- Useful if you don't have access to the internal mechanics of the model

OK for (more?) Proxies people



(a) Confusion matrix of a single binary classification model, colored by prediction correctness

(b) Histogram of age, colored by classification

(c) Two-dimensional histogram of age and sex, colored by classification

(d) Small multiples by sex. Each scatterplot shows age vs positive classification score, colored by classification

(e) Histograms of performance in a regression model that predicts age, faceted into 3 age buckets

(f) Using images as thumbnails for image datasets

[3]What-If tool, 2020

Benoît Otjacques, benoit.otjacques@list.lu

[3] IEEE TCVG 26(1), 2020: Wexler J. et al. The What-If Tool: Interactive Probing of Machine Learning Models

Conclusion

Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

Conclusion

Focus on People acting as **proxy** between
the citizens and the AI experts

- AI diagnosis systems → **Doctors** → Patients
- Predictive Justice systems → **Judges, Lawyers** → Defendants
- Autonomous cars → **Engineers** → Drivers
- AI-based portfolio management → **Bankers** → Investors
- AI-based HR systems → **Recruitment Officers** → Applicants
- AI-based Governance → **Policy makers** → Citizens
- Governance of AI → **Policy makers** → Citizens
- …

Benoît Otjacques, benoit.otjacques@list.lu

VISUALISATION & ADOPTION OF AI

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

Conclusion

(1) Visualisation is emerging as a mean to support AI experts

(2) Visualisation should also be used more to support proxy people

- to better understand AI,
- which will later help them to take informed decisions about AI
- and to communicate with citizens, patients, drivers, applicants…
- ultimately leading to a **wider acceptance and use of AI**.

People at the interface between the users / citizens and the AI-based system are key in the adoption process (proxies)

Benoît Otjacques, benoit.otjacques@list.lu

LUXEMBOURG INSTITUTE OF SCIENCE AND TECHNOLOGY | LIST

Thank you

Dr Ir Benoît Otjacques

benoit.otjacques@list.lu