

# Explainable AI: Machine Learning for Humans



Yannick Martel  
Bleckwen Director



Leonardo Noletto  
Bleckwen Senior Data Scientist

12 June, 2018



[@bleckwen\\_ai](https://twitter.com/bleckwen_ai)



[www.linkedin.com/company/bleckwen](https://www.linkedin.com/company/bleckwen)

# Artificial Intelligence is Everywhere



# Booking an appointment for you

## Google Assistant



*"Hi, I'm calling to book a  
women's haircut for a client."*

## Hair Salon



*"Sure, give me one second."*

# Beating the Go Master





# Carrying goods



# Defeating Breast Cancer

A 62-year-old woman presents for diagnostic mammography with a palpable mass in her right breast. Describe the finding.



You answered

- a. Infiltrating duct carcinoma of breast
- b. Lipoma of breast**
- c. Fibroadenoma of breast
- d. Fat necrosis of breast



Processing image

Analyze mammogram image in first view

Emphasize bright areas

Narrow down

Detect candidates

Refine candidate borders

Image Analysis

Watson Inference

Infiltrating duct carcin...

Lipoma of breast

Fibroadenoma of bre...

Fat necrosis of breast

Eyes of Watson

Analyzing Text

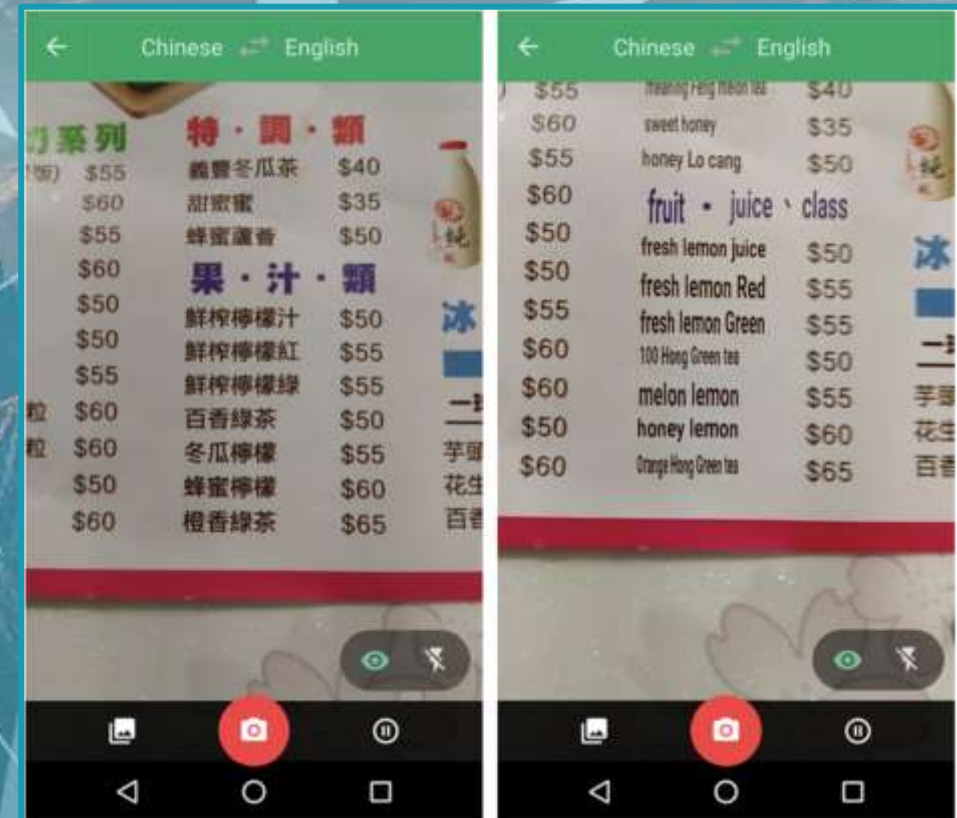
Analyzing Images

Reasoning

Selecting Answer



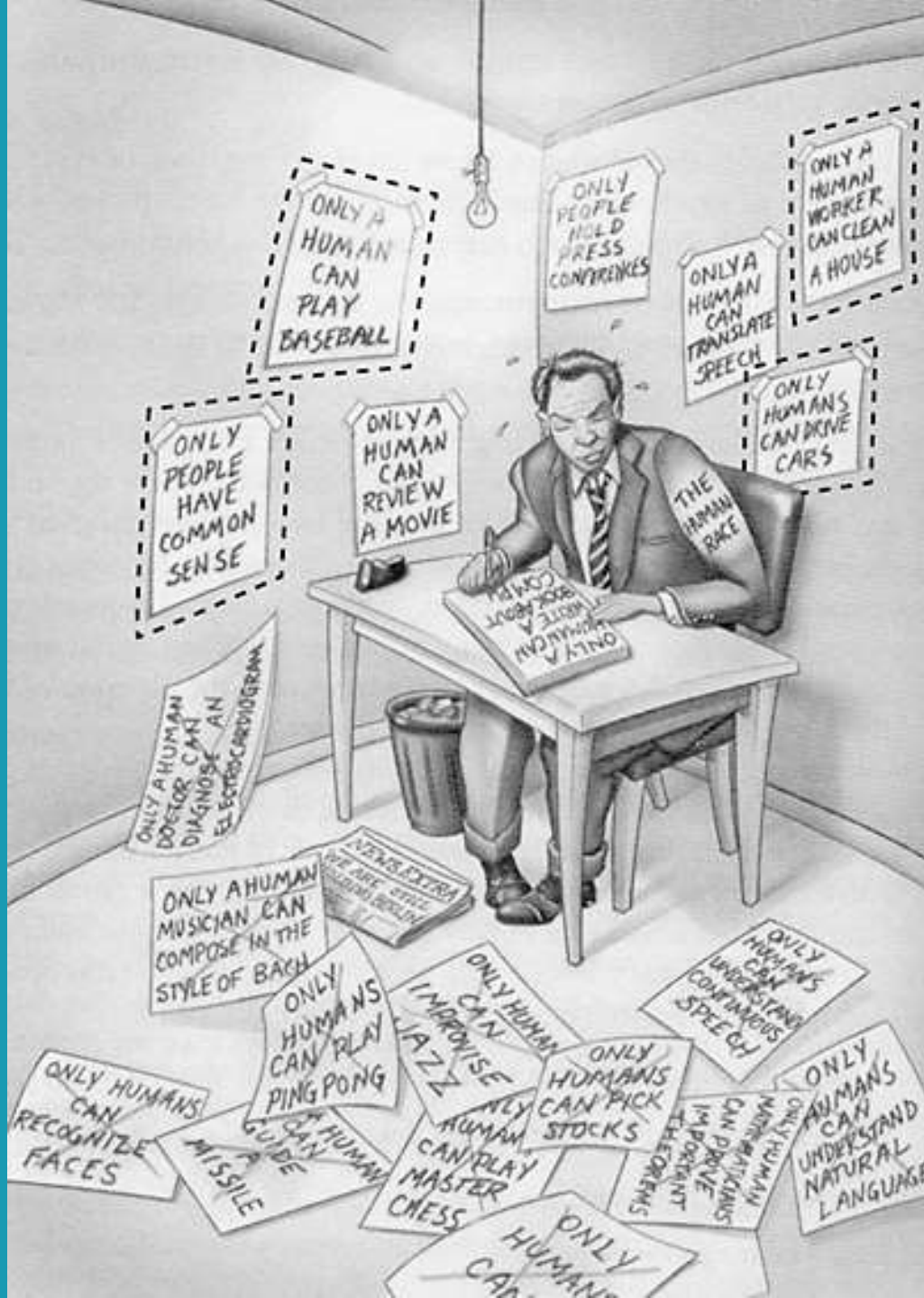
# Speaking Chinese



And even painting







When a task can be automated at a low cost, the task will be automated.





Augmented  
Human

Autonomous  
AI



Human-in-the-loop

Human-out-of-the-loop





We believe in a collaboration between  
Human and Artificial Intelligence.

Interpretability is the key enabler.





# Why you should care about Interpretability?

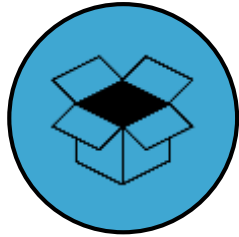


“Interpretability is the degree to which a human can understand the cause of a decision.”

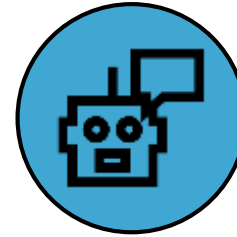
*Miller, Tim. 2017. “Explanation in Artificial Intelligence: Insights from the Social Sciences.”*



# Why do we need interpretability?



STRENGTHEN TRUST AND  
TRANSPARENCY



EXPLAIN DECISIONS



SATISFY REGULATORY  
REQUIREMENTS



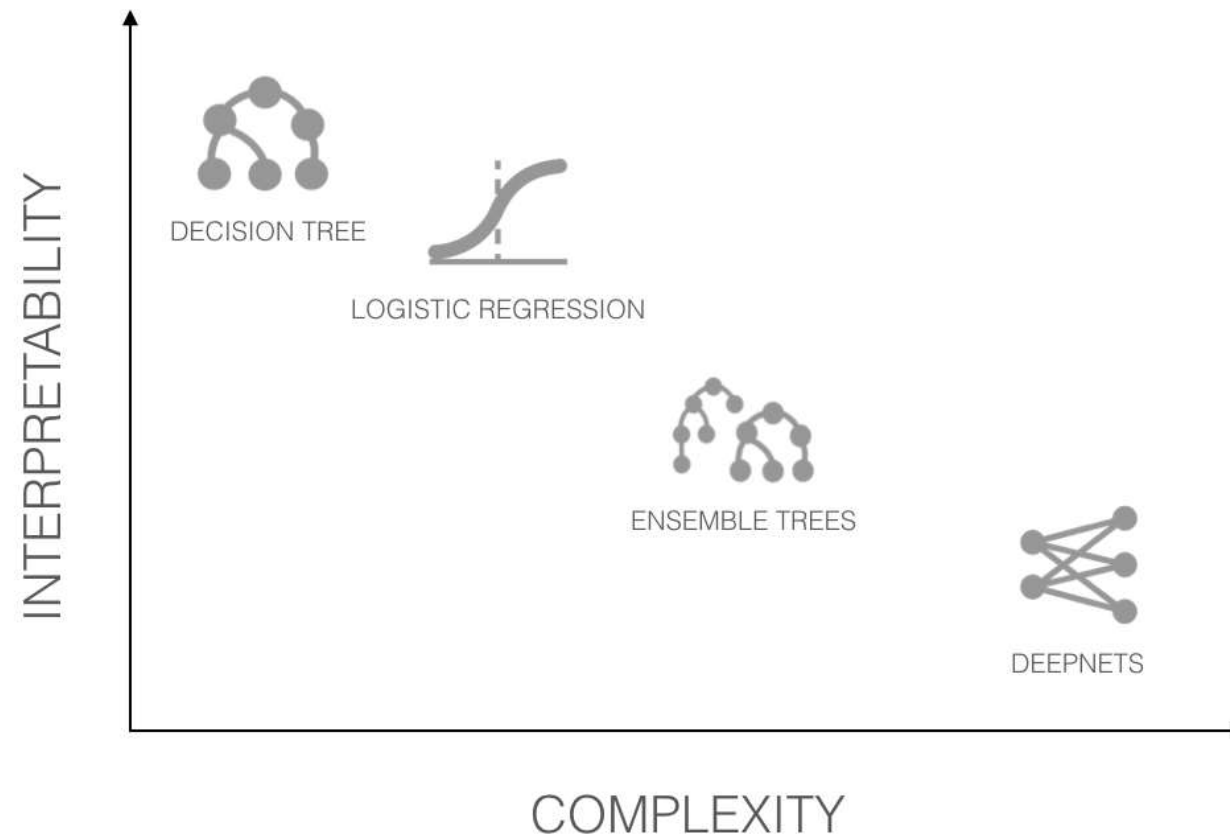
IMPROVE MODELS







But what about the trade-off Accuracy vs. Interpretability?





# Explaining the taxonomy of Interpretability



local

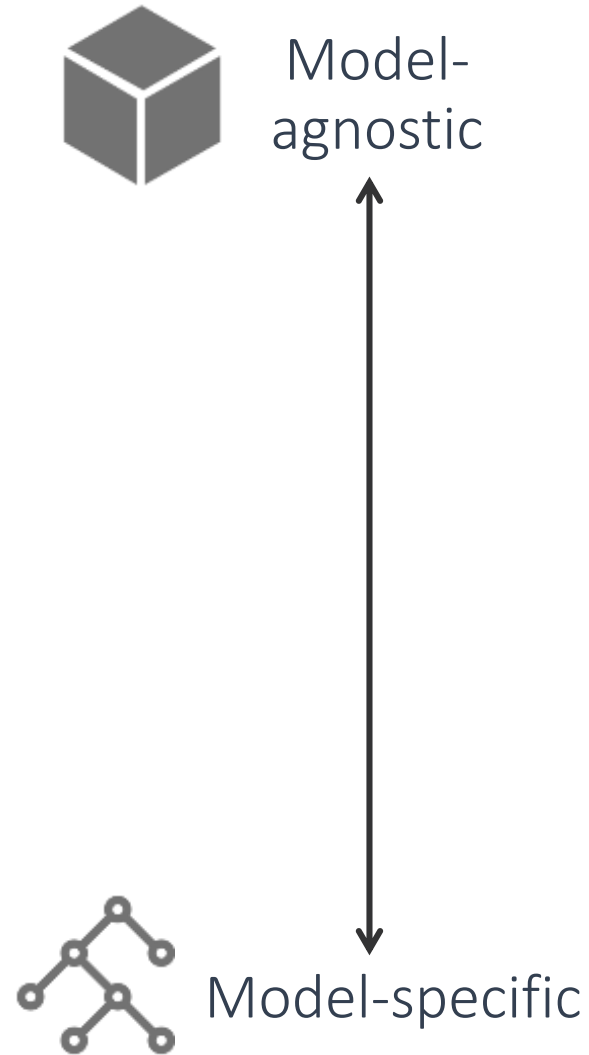


global





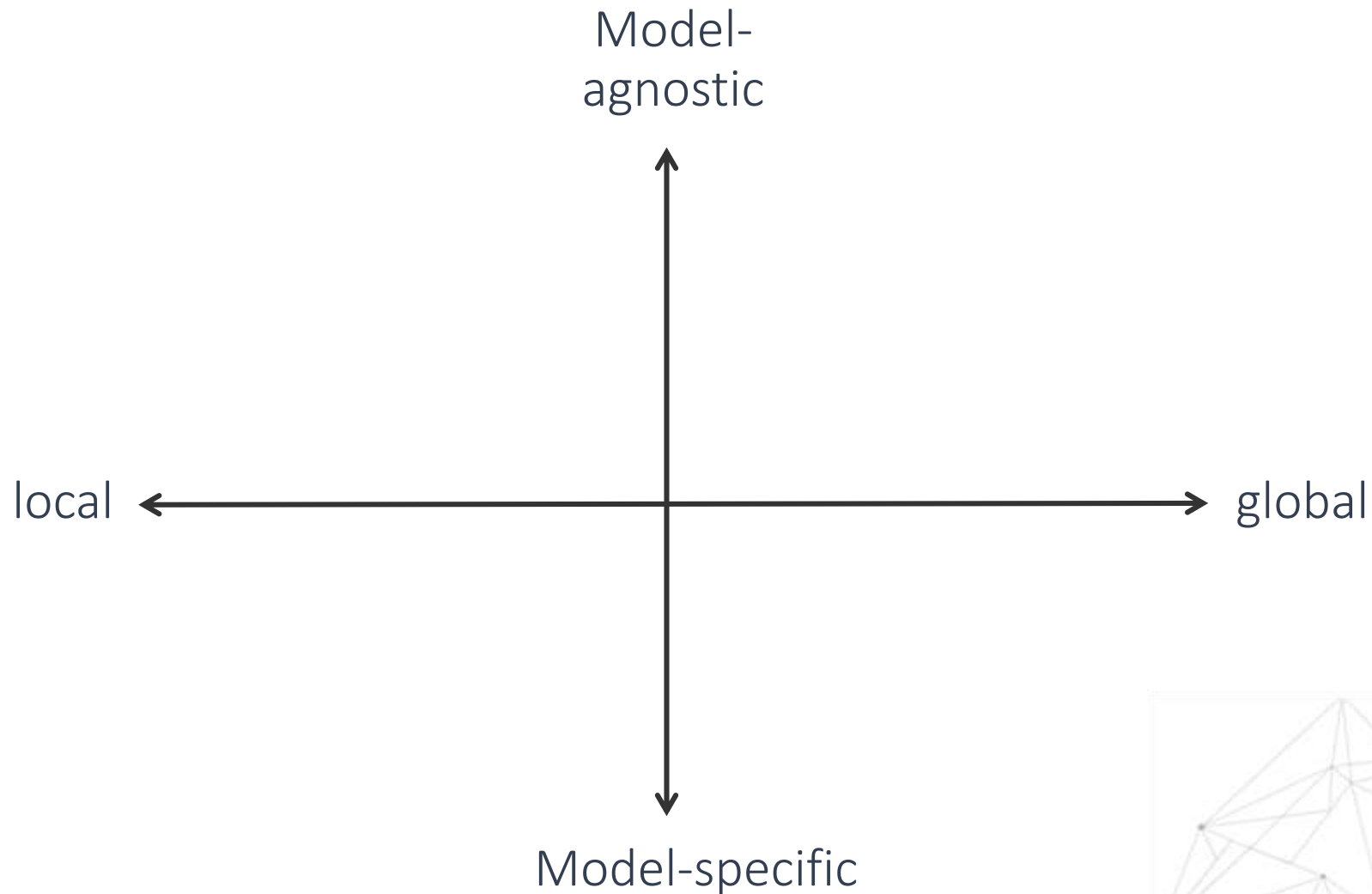
# Explaining the taxonomy of Interpretability





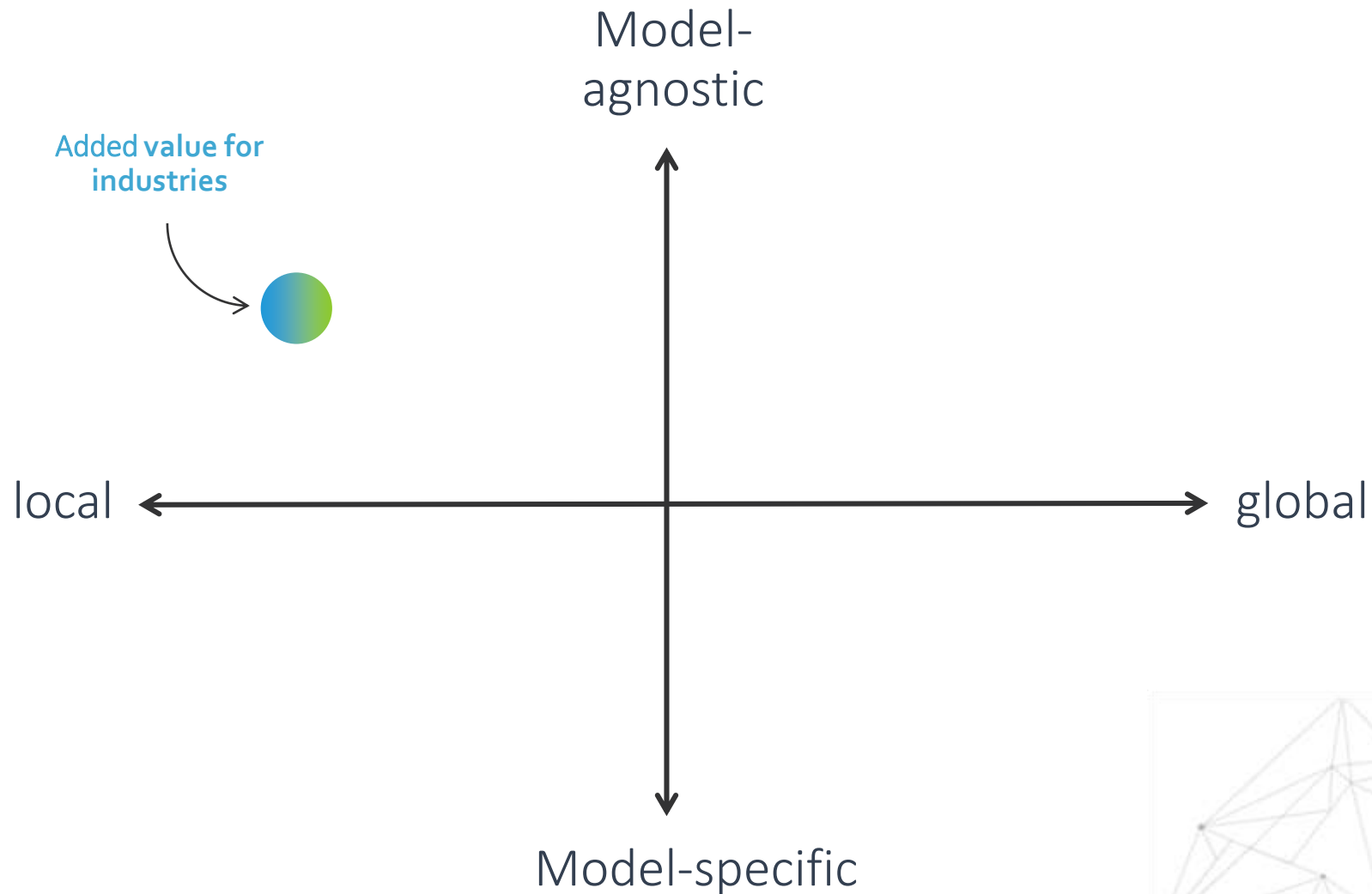


# Explaining the taxonomy of Interpretability





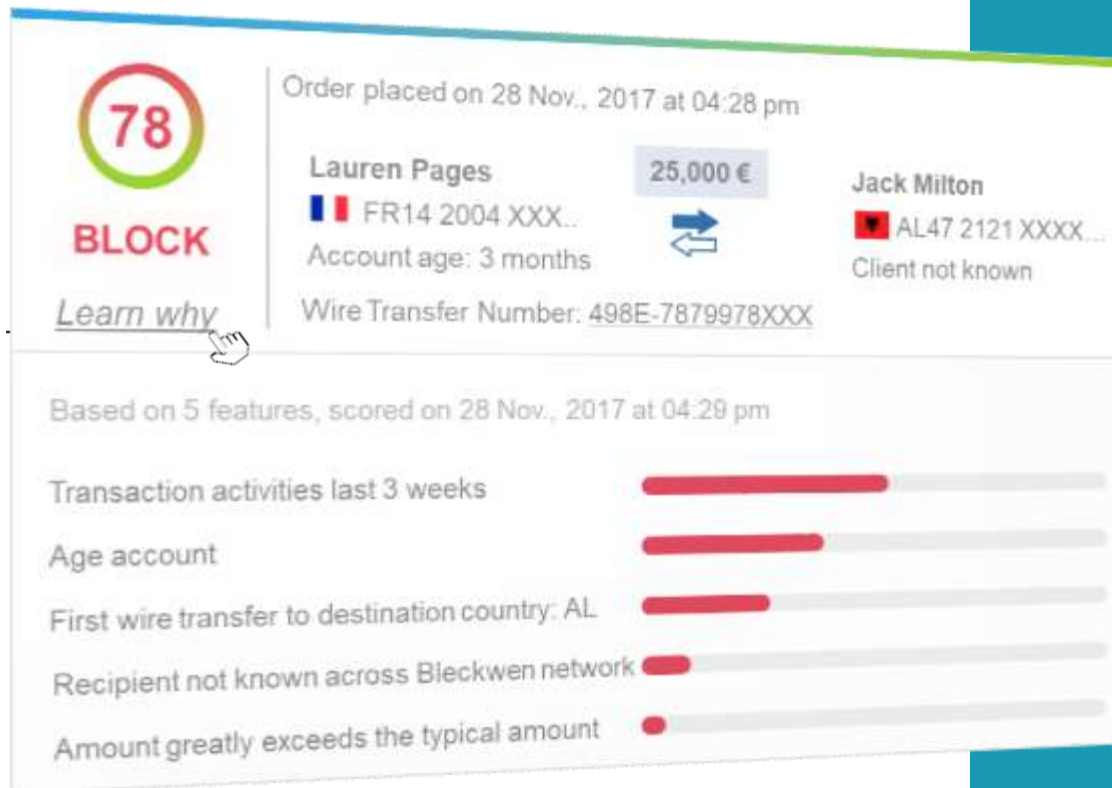
# Explaining the taxonomy of Interpretability



How we are helping the banking industry accelerate  
the adoption of AI?







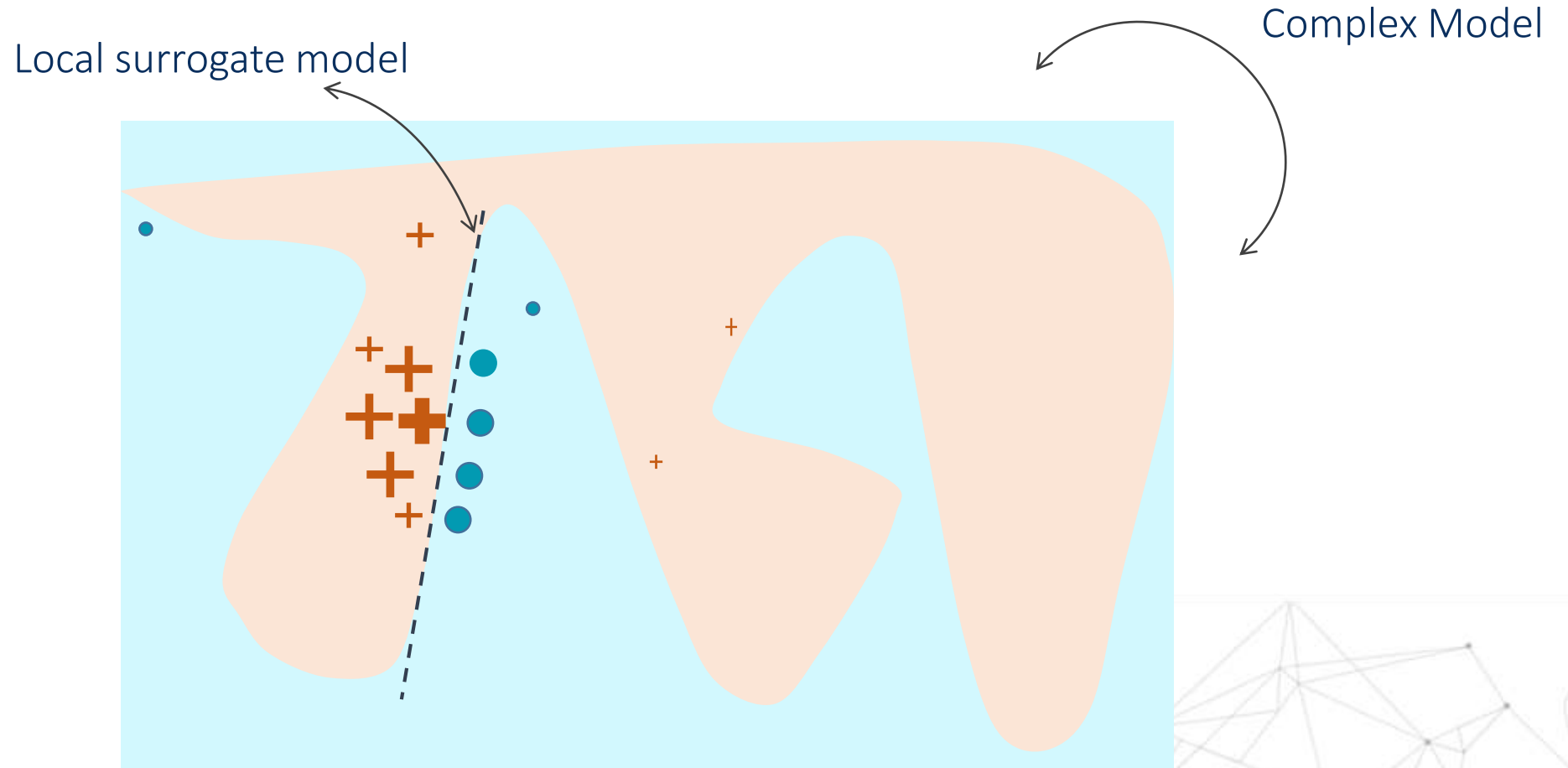
Our clients are fighting against financial fraud with the power of Machine Learning and the Behavioral Profiling while keeping decisions interpretable.

How is that possible?




# Introducing LIME

*Local Interpretable Model-Agnostic Explanations*





# Introducing LIME

*Local Interpretable Model-Agnostic Explanations*






  
**BLOCK**  
[Learn why](#)

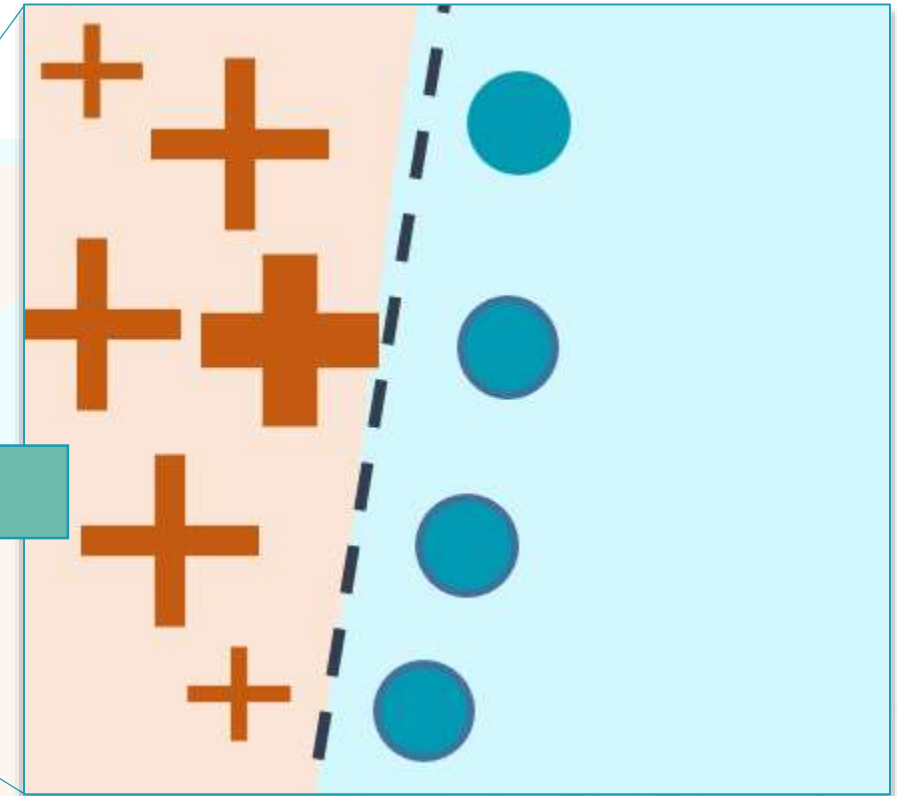
Order placed on 28 Nov., 2017 at 04:28 pm

<b>Lauren Pages</b> FR14 2004 XXX.. Account age: 3 months	<b>25,000 €</b>  	<b>Jack Milton</b> AL47 2121 XXXX.. Client not known
---	---	--

Wire Transfer Number: 498E-7879978XXX

Based on 5 features, scored on 28 Nov., 2017 at 04:29 pm

Transaction activities last 3 weeks	
Age account	
First wire transfer to destination country: AL	
Recipient not known across Bleckwen network	
Amount greatly exceeds the typical amount	



Extracting explanations from a linear model



We combine best-in-class Machine Learning models with  
**interpretation technologies** to get the best from  
a collective artificial and human intelligence



Fraud on payments  
b2b • b2c • p2p



Fraud on credit  
applications



AML / ATF



Collaboration between  
Human Being and Artificial Intelligence will  
secure the world.



Thank you for your attention.



BLECKWEN

DEEP TRUTH REVEALED

---

Anti-fraud solution for banks

---

[www.bleckwen.ai](http://www.bleckwen.ai)