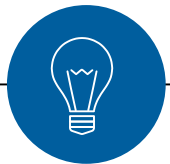


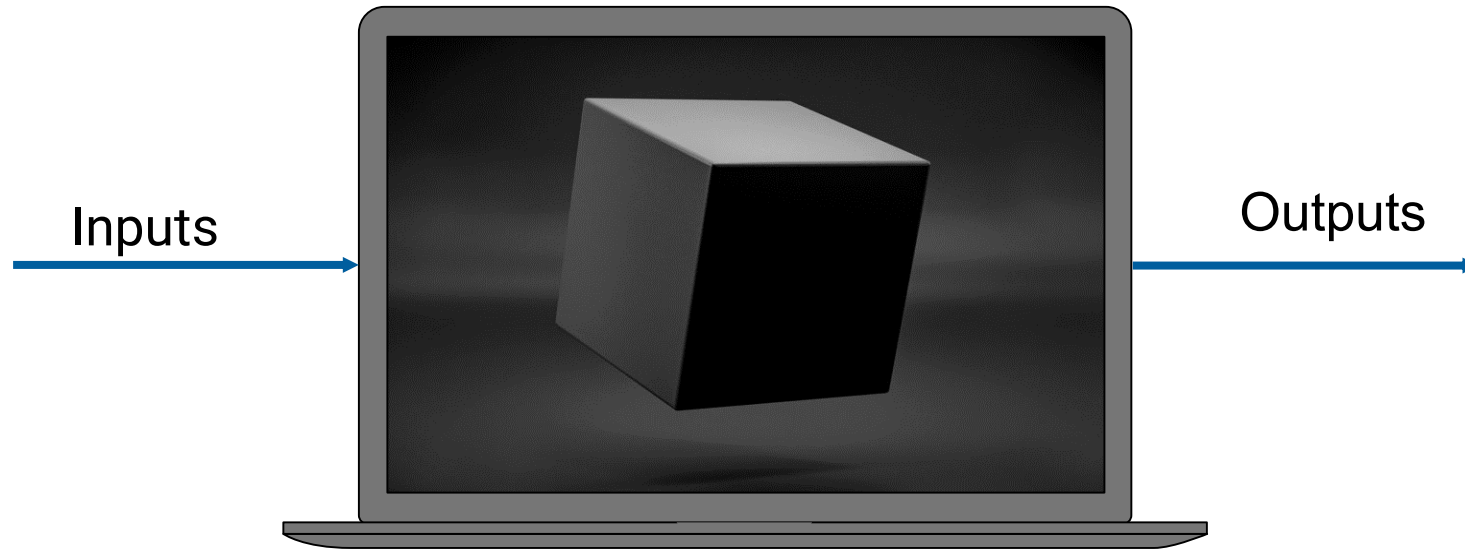
TOWARDS INCREASED INTERPRETABILITY OF AI-TOOLS IN THE ENERGY SECTOR WITH FOCUS ON THE WIND ENERGY TRADING APPLICATION



Presenter: **Konstantinos Parginos**

In collaboration with: George Kariniotakis, Ricardo Bessa, Simon Camal

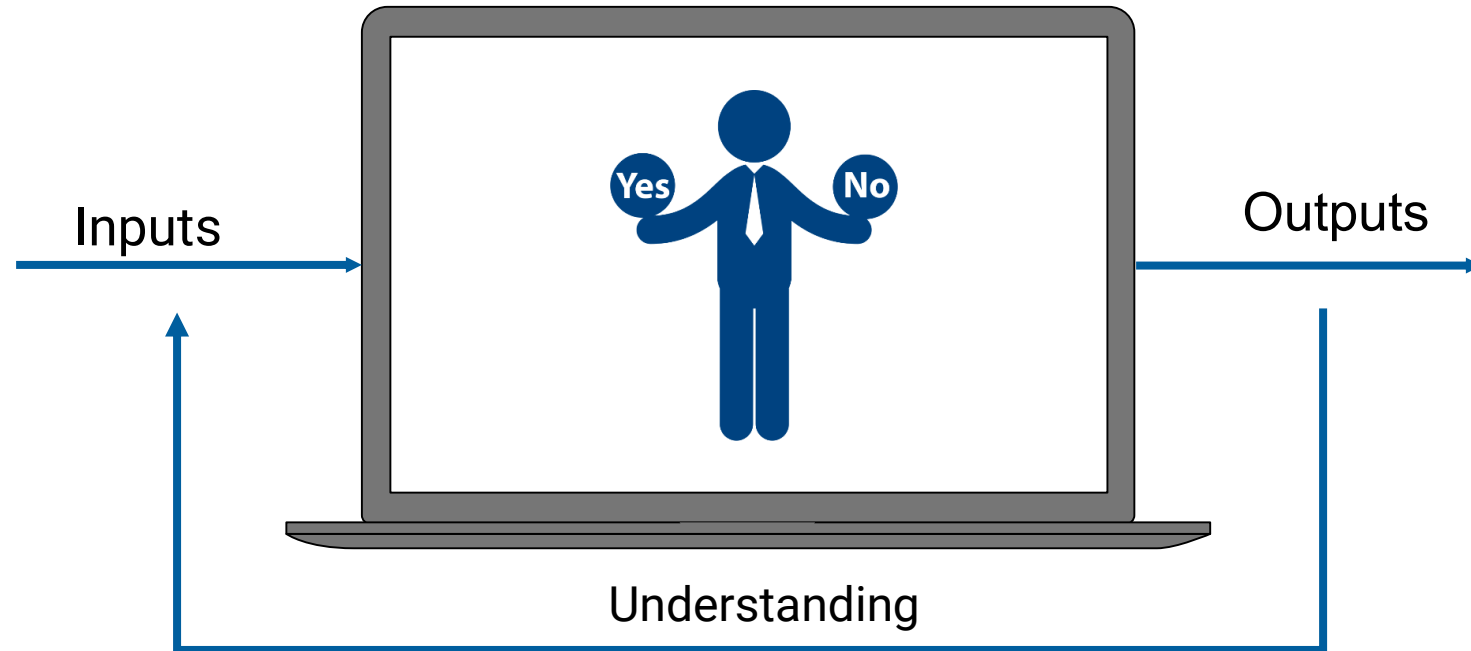
THE BLACK-BOX PROBLEM



Limitations

Difficulty to assess output in **High Dimensional** and **Non Linear** Problems

WHAT IS INTERPRETABILITY



Definition

“ The ability to explain or to present a model output in understandable terms to a human”,
Doshi-Velez & Been Kim, 2017



WHO NEEDS INTERPRETABILITY?



AI Model Builders:

- Why is the model not performing?
- How can I improve it?

End Users of AI systems:

- Should I trust the model's output?
- How should I respond to the predictions?

Public stakeholders:

- Is the model safe and fit for purpose?
- Does it comply with regulations?

As currently proposed, AI tools when applied to high risk systems like energy systems, should be subject to the largest set of requirements



- human oversight
- transparency
- cybersecurity
- risk management
- data quality

reporting obligations

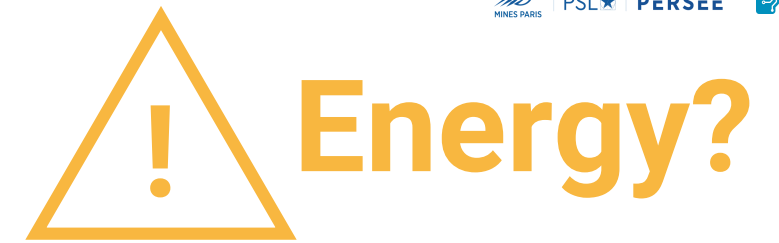
It is important to monitor the overall design as well as the system outputs.

Interpretability is a key characteristic that AI systems should provide in a user friendly way, to provide trust in its outcomes.

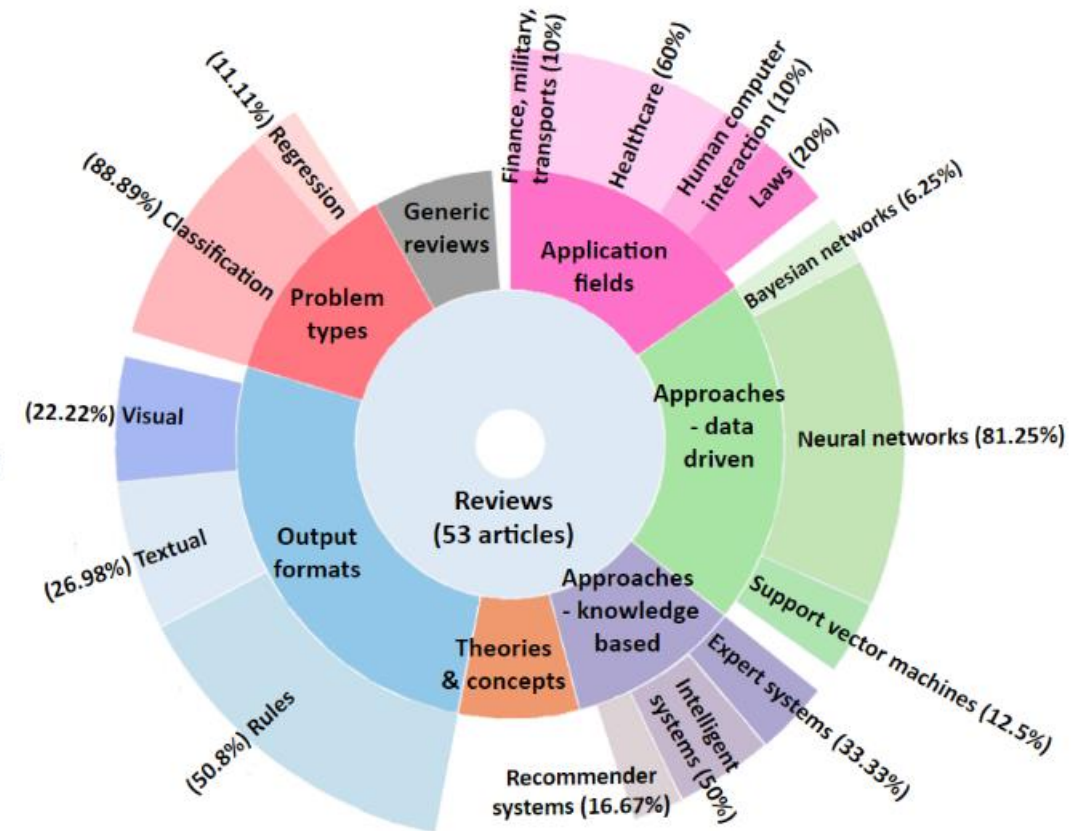
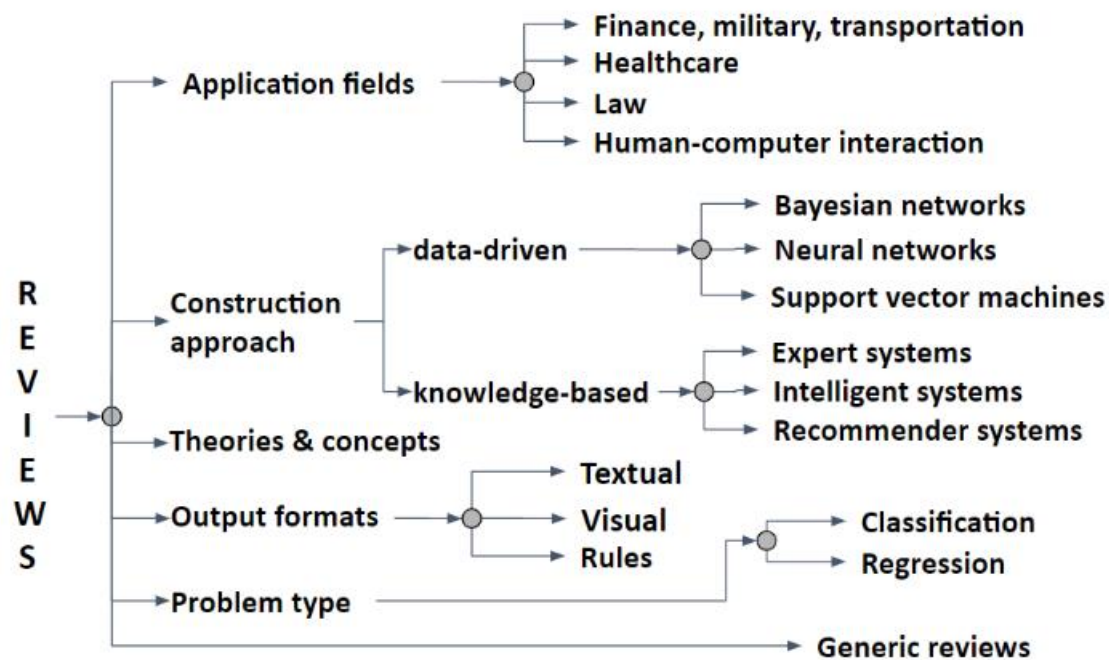
Deeper understanding of the models can help engineers fine tune their systems to perform consistently throughout their lifetime, **especially on outliers.**

Source: "White Paper On Artificial Intelligence - A European approach to excellence and trust" by European Commission, 2020

Source: Operating AI systems in the electricity sector under European's AI Act – Insights on compliance costs, profitability frontiers and extraterritorial effects" by F. Heymann, Konstantinos Parginos, Ricardo Bessa, M. Galus, 2022



STATE OF THE ART



ARTIFICIAL INTELLIGENCE APPLICATIONS TO THE ENERGY SECTOR

Demand & RES Forecast



Anomaly detection



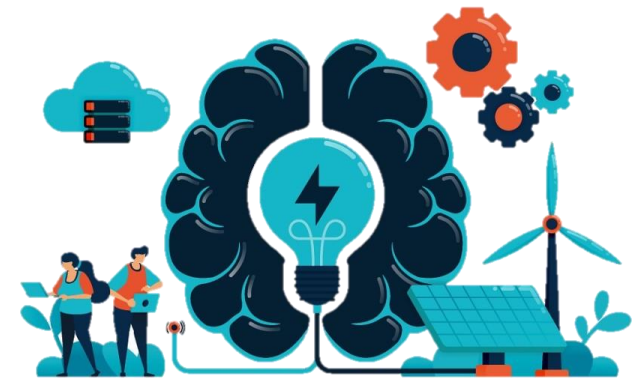
Predictive Maintenance



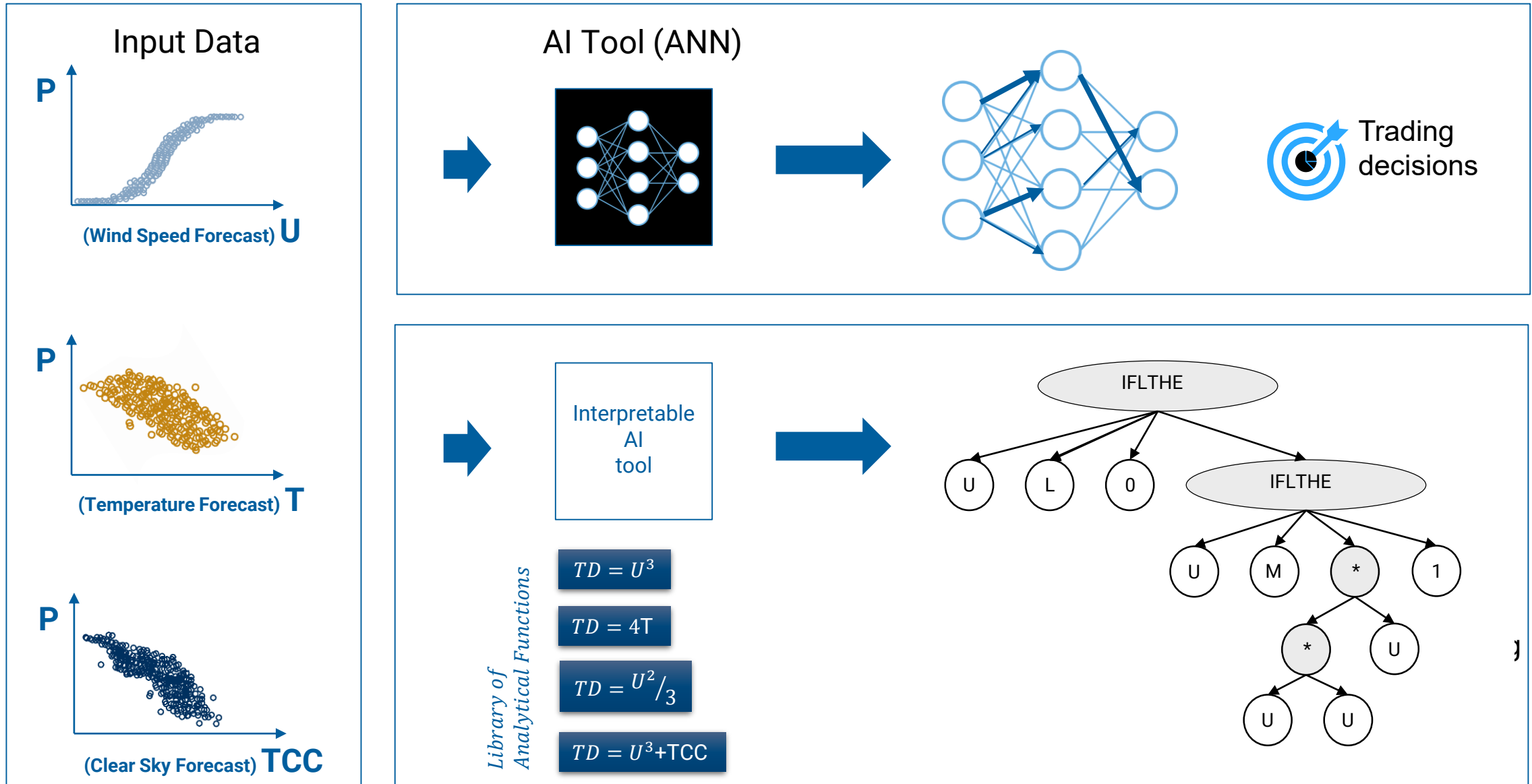
Power Trading



Demand Profiling



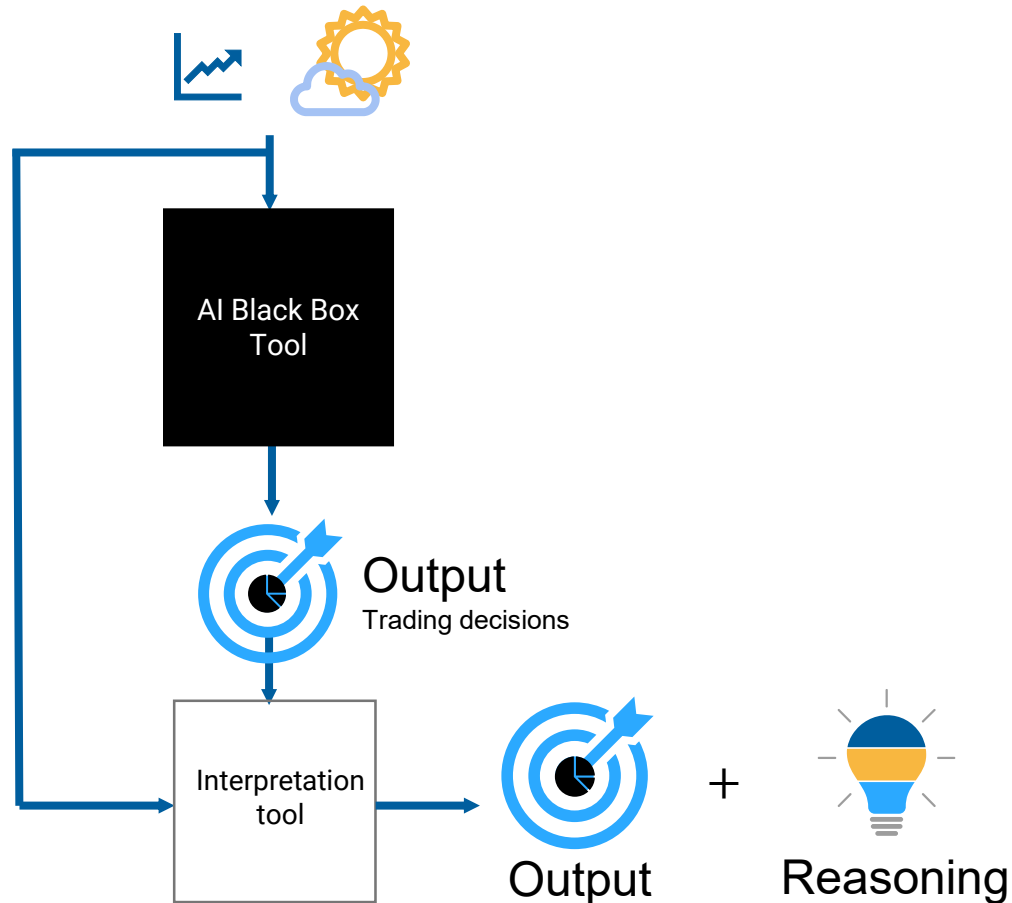
AI INTERPRETATION TO THE ENERGY SECTOR



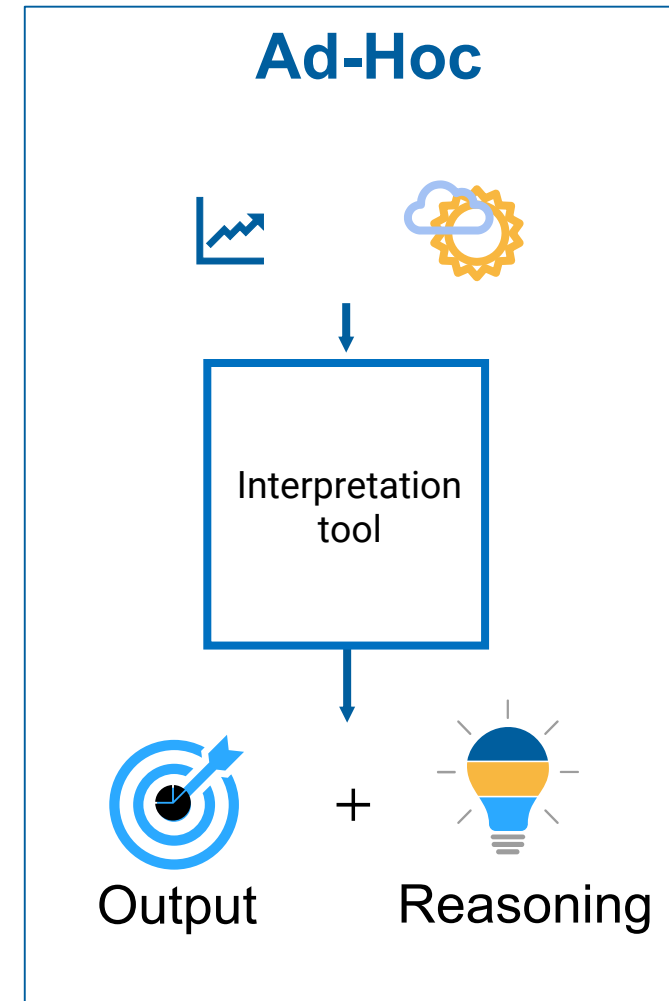
INTERPRETABLE METHODS

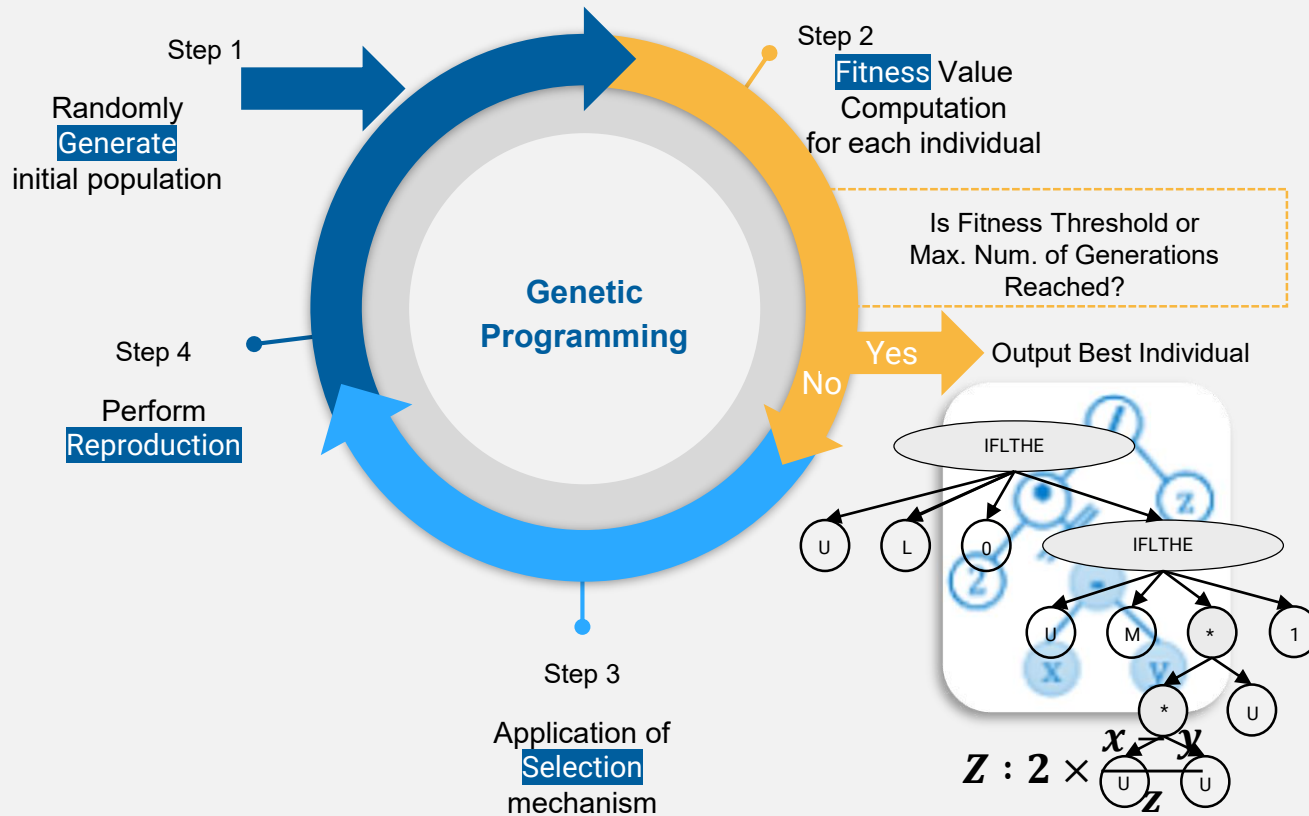
- Two main alternative interpretable approaches

Post-Hoc



Ad-Hoc





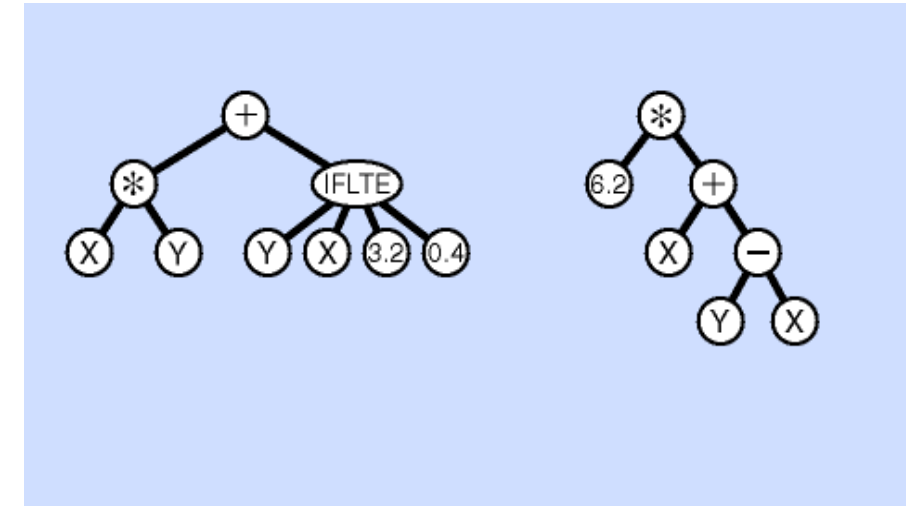
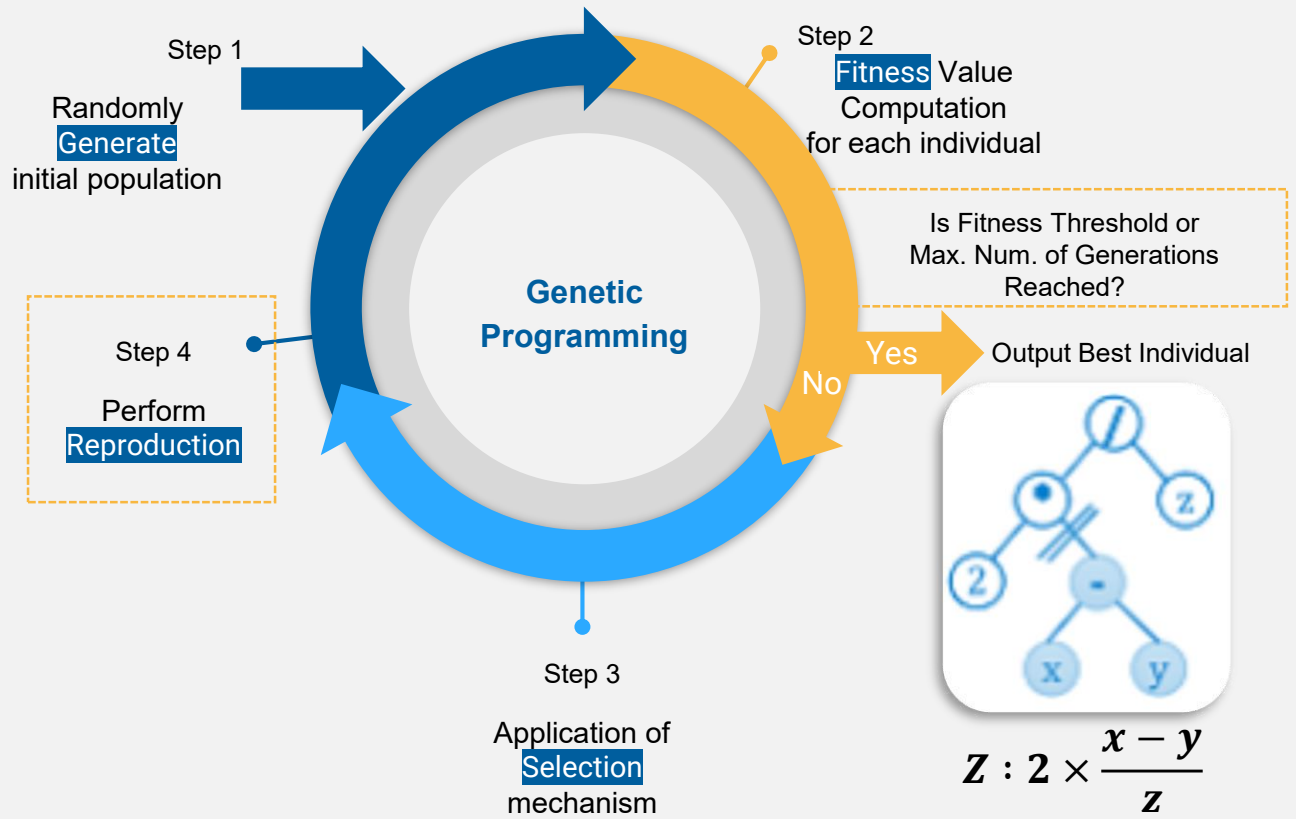
Using natural selection to build interpretable models

GP is an extension of evolutionary algorithms which the individuals in a population are programs instead of binary strings.

SR is a subset of GP where the solution obtained from the regression, represent the target model.

THE FINAL OUTPUT IS AN INDIVIDUAL TREE LIKE ANALYTICAL EXPRESSION: Z

SYMBOLIC REGRESSION





BENEFITS OF SYMBOLIC REGRESSION (SR)

- Data-Driven analytical solutions
- Optimize both weights and feature relations
- Could be used both as Post-Hoc and Ad-Hoc

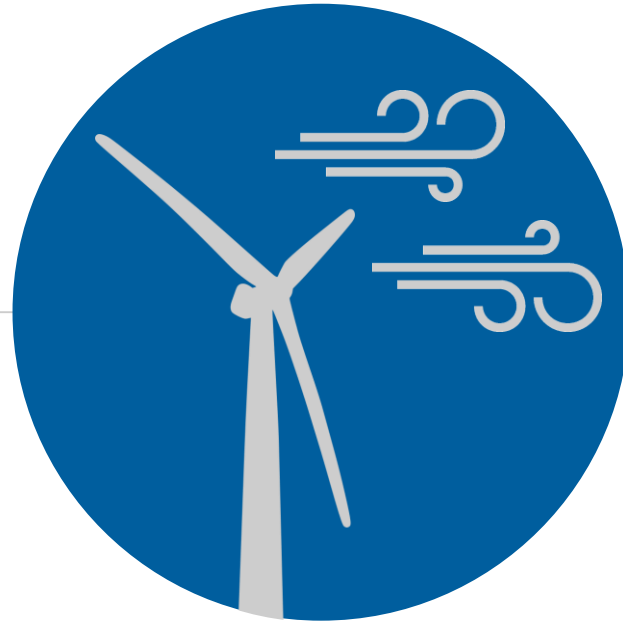
$$f_{w_1, w_2}(x_1, x_2)$$

y = output

$x_{1,2,\dots}$ = input

$w_{1,2,\dots}$ = weight

$$\begin{array}{c} \downarrow \\ y = w_1 x_1 + w_2 x_2 \\ \uparrow \quad \uparrow \end{array}$$

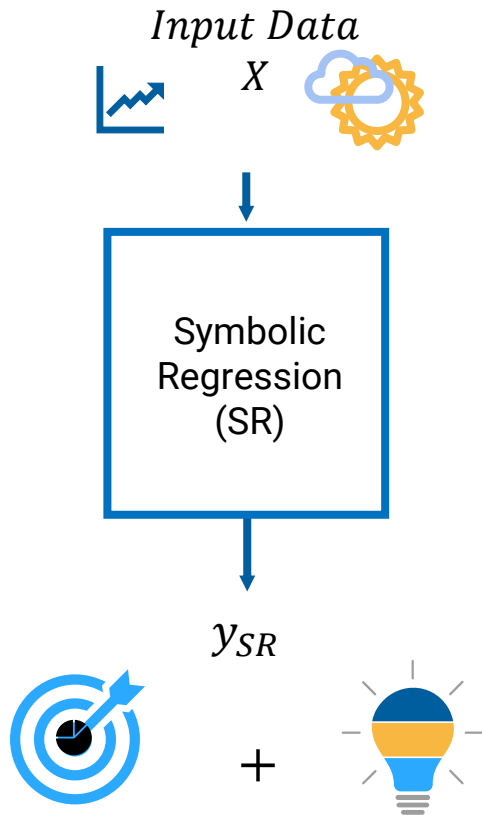


CASE STUDY

Optimizing Bidding Decision of Wind Turbines
in Day-Ahead Market

PROBLEM FORMULATION

Ad-Hoc



Definitions

$$y_{SR} = G(Z_{E,\theta}, X)$$

$$Z_{opt} = \underset{\{Z_{E,\theta} \in \Phi\}}{\operatorname{argmin}} L(G(Z_{E,\theta}, X), y)$$

Formulation of Fitness Function L to optimize Trading Value:

$y = p^E$: Actual Energy Produced

$\lambda^\uparrow, \lambda^\downarrow$: Imbalance Prices

$$L = \underbrace{\left[-\lambda^\uparrow (p^E - y_{SR})^- + \lambda^\downarrow (p^E - y_{SR})^+ \right]}_{\text{imbalance cost}}$$

Two Approaches

Non Clustered Data

SR_{Global}

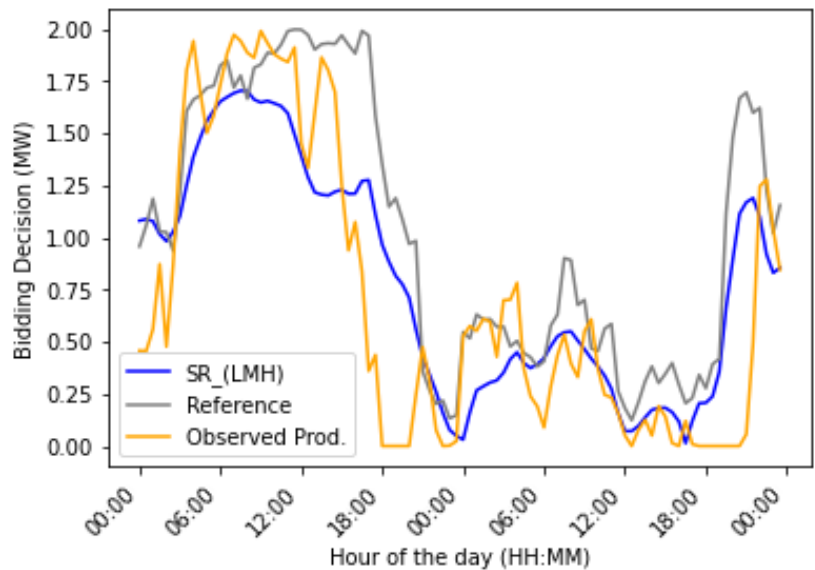
Clustered Data

$SR_{(Low, Medium, High)}$

RESULTS

4 Wind Turbines (2MW each), located in France for the period of March 2019-May 2020

Trading Strategy	WT ₁
Global	$0.11w_{speed}$
Low w_{speed}	$0.04w_{speed}^2$
Med w_{speed}	$-0.027 + \frac{0.87w_{speed}^3}{Temp}$
High w_{speed}	$\frac{1.77w_{speed}}{2TCC+w_{speed}}$

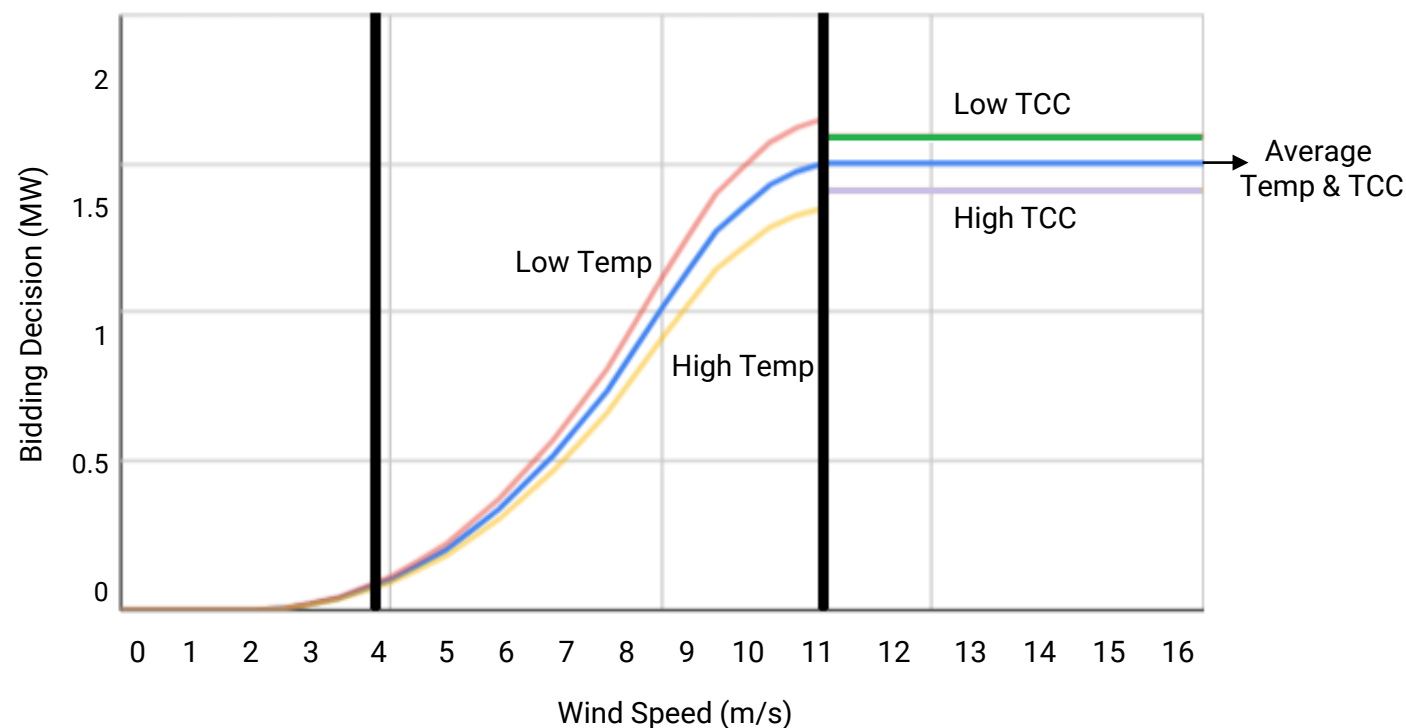


Selected Strategy	Total Revenue (€)				Comparison
	WT ₁	WT ₂	WT ₃	WT ₄	w./ Ref. Bidding
<i>Bidding Perfect hindsight</i>	272.209	264.449	254.960	253.068	7.15%
<i>Reference Bidding (Opt. Quantile)</i>	257.158	241.241	238.468	238.145	0.00%
SR_{Global}	238.632	228.124	219.630	219.744	-7.06%
$SR_{(Low,Medium,High)}$	250.051	241.338	232.002	231.072	-2.11%

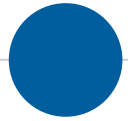
Data driven analytical expressions derived from SR perform better after subsampling the dataset according to different wind speed values

INTERPRETATION OF RESULTS

	Low wind speed	Medium wind speed	High wind speed	
SR expression	$0.04w_{speed}^2$	$-0.027 + \frac{0.87w_{speed}^3}{Temp}$	$\frac{1.77w_{speed}}{2TCC+w_{speed}}$	$Temp \in [271K, 311K]$ $TCC \in [0, 1]$



We can identify the key features that drive our model as well as how they interact with each other in different wind speed conditions



CONCLUSIONS

- Interpretability of AI tools is an essential enhancement for applications to the energy sector
- EU regulation on AI will impose a dedicated framework for high-risk sectors such as energy systems
- A knowledge transfer of interpretability tools from other sectors to energy applications is needed
- We adapted the Ad-Hoc Interpretable method of Symbolic Regression to the DA Trading Case study
- Further evaluation of the interpretability of the results is needed



THANKS!



Artificial Intelligence
for the Sciences



Marie
Skłodowska-Curie
Actions



Any **questions** ?

You can find me at



[linkedin.com/in/konstantinos-parginos](https://www.linkedin.com/in/konstantinos-parginos)



konstantinos.parginos@minesparis.psl.eu