

The background of the slide is a vibrant, stylized illustration. It depicts Hiccup from the movie 'How to Train Your Dragon' riding his black dragon, Toothless. They are flying over a fantastical landscape with blue and green terrain, glowing plants, and a large white dragon in the background. The scene is set against a bright, hazy sky.

Ethical ML

Beyond Model Transparency and Explainability





and

How do you win over your enemy?

""

Most people are afraid of things
because they don't understand them.
For me, it was dragons.
Then one day, that changed, and
my greatest fear became my best friend
"" - Hiccup, How to Train Your Dragon



Beyond Model Transparency and Explainability

"""

Most people are afraid of things
because they don't understand them.
For me, it was AI & Emerging Technology.
Then one day, that changed, and
my greatest fear became my best friend

"""

- Prayson, How to Train Your Machine Learning Models



Beyond Model Transparency and Explainability

"""

Most people are afraid of things
because they don't understand them.

For me, it was GDPR.

Then one day, that changed, and
my greatest fear became my best friend

"""

- Prayson, How to Train Your Machine Learning Models



Beyond Model Transparency and Explainability

"""

Most people are afraid of things
because they don't understand them.

For me, it was ALTAI.

Then one day, that changed, and
my greatest fear became my best friend

"""

- Prayson, How to Train Your Machine Learning Models



Why should you care?

"""

Social and economic

Harm → (Allocation | Quality of Services | Environmental)

Technological advancement



"""



😞 We, (Data Scientists), have mess things up!

```
"""  
    Focused on the wrong metrics  
    Performance > Explainability > [Fairness]  
    Fairness > Performance > {Explainability}  
    """.replace("> {Explainability}", "")
```



🐱 We, (Data Scientists), are fixing it!

```
"""  
    Changing the Language  
    Definitions | calling a ♠ ♠  
    """
```

```
amend = {"p-question": "can we trust the model?"} | {"e-question": "is the model fair?"}
```




🐱 We, (Data Scientists), are fixing it!

```
"""  
    Changing the Language  
    Definitions | calling a ♠ ♠  
    """
```

```
amend = {"p-question": "can we trust the model?"} | {"e-question": "is the model fair?"}
```



 Show & Explain: Counterfactual Test!

```
def explain(counterfactual):  
    ...  
  
def show(counterfactual):  
    ...
```




Beyond Model Transparency and Explainability

"""

Let's do what we were supposed to have done yesterday.

Let's do it today before it is late.

Let's do it for the future of our social and economic wellbeing.

"""

-  Prayson



Beyond Model Transparency and Explainability

"""

Let's do what we were supposed to have done yesterday.

Let's do it today before it is late.

Let's do it for the future of our social and economic wellbeing.

"""

-  Prayson



Beyond Model Transparency and Explainability

"""

Let's do what we were supposed to have done yesterday.

Let's do it today before it is late.

Let's do it for the future of our social and economic wellbeing.

"""

-  Prayson



Beyond Model Transparency and Explainability

"""

Let's do what we were supposed to have done yesterday.

Let's do it today before it is late.

Let's do it for the future of our social and economic wellbeing.

"""

-  Prayson