

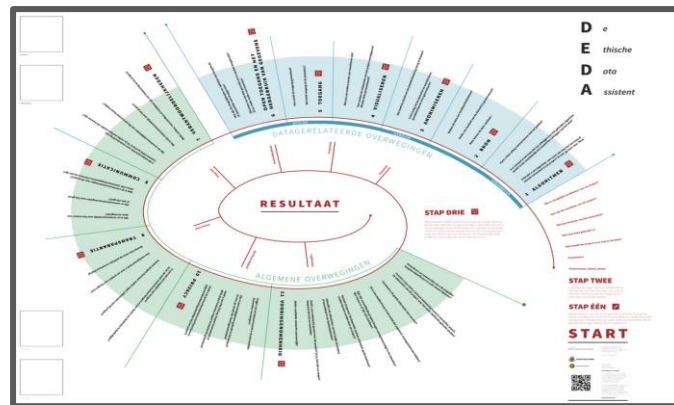
# Vraagstukken rondom Responsible AI

Voor de inzet en werking van specifieke algoritmes, alsook de uitkomsten en de betekenis daarvan bij het vormen van een gegeven besluit, geldt dat de volgende vraagstukken van belang zijn:

1. Zorgvuldige praktijk rondom ontwikkeling en inzet van algoritmes
2. Inzicht in, en beheersing van, potentiële impact en risico's
3. Uitlegbaarheid, transparantie en veiligheid
4. Toetsing / auditeerbaarheid / reproduceerbaarheid
5. Eerlijke behandeling van verschillende bevolkingsgroepen



# Zorgvuldige praktijk ontwikkeling & inzet van algoritmes

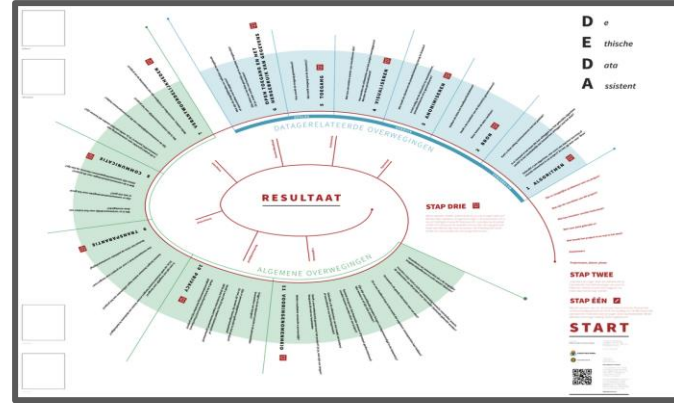


## An iterative approach to successfully deliver data products in your organisation

An approach based on hands-on experience and real world use cases with a strong focus on delivering valuable results for end users.

Whitepaper  
April 2021 - v.1.0.0

# Inzicht in, en beheersing van, potentiële impact en risico's



# Uitlegbaarheid, transparantie en veiligheid

What-If Tool

**What If...**

you could inspect a machine learning model,  
with no coding required?



IBM Research Trusted AI

AI Explainability 360

**im** InterpretML

**Types of Models Supported**



**Glass-Box**

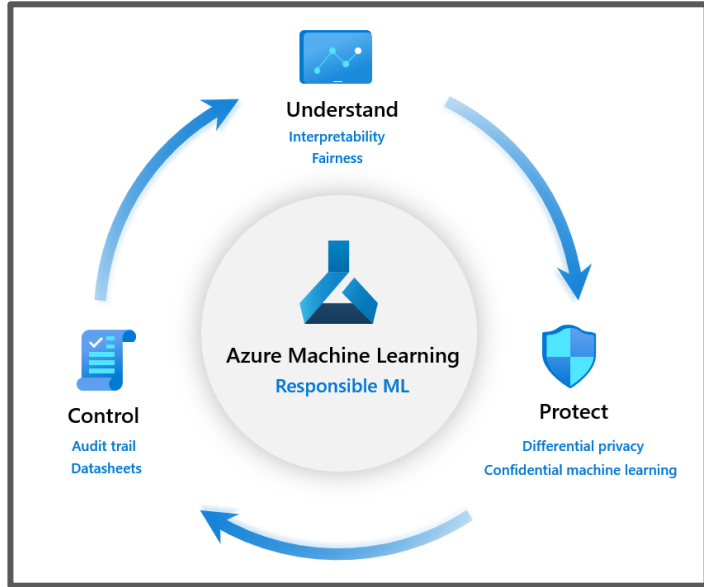


**Black-Box**

Glass-box models are interpretable due to their structure. Examples include:  
Explainable Boosting Machines (EBM), Linear models, and decision trees.

Glass-box models produce lossless explanations and are editable by domain experts.

# Toetsing / auditeerbaarheid / reproduceerbaarheid



**Amazon SageMaker Pipelines**

# Eerlijke behandeling van verschillende bevolkingsgroepen

 Fairlearn

What-If Tool

IBM Research Trusted AI

AI Fairness 360

**Amazon SageMaker Clarify**

Detect bias in ML models and understand model predictions

