

Transparency

Data science that provides transparency –

How to clarify answers such that they become indisputable?

December 2, 2016

Jack van Wijk



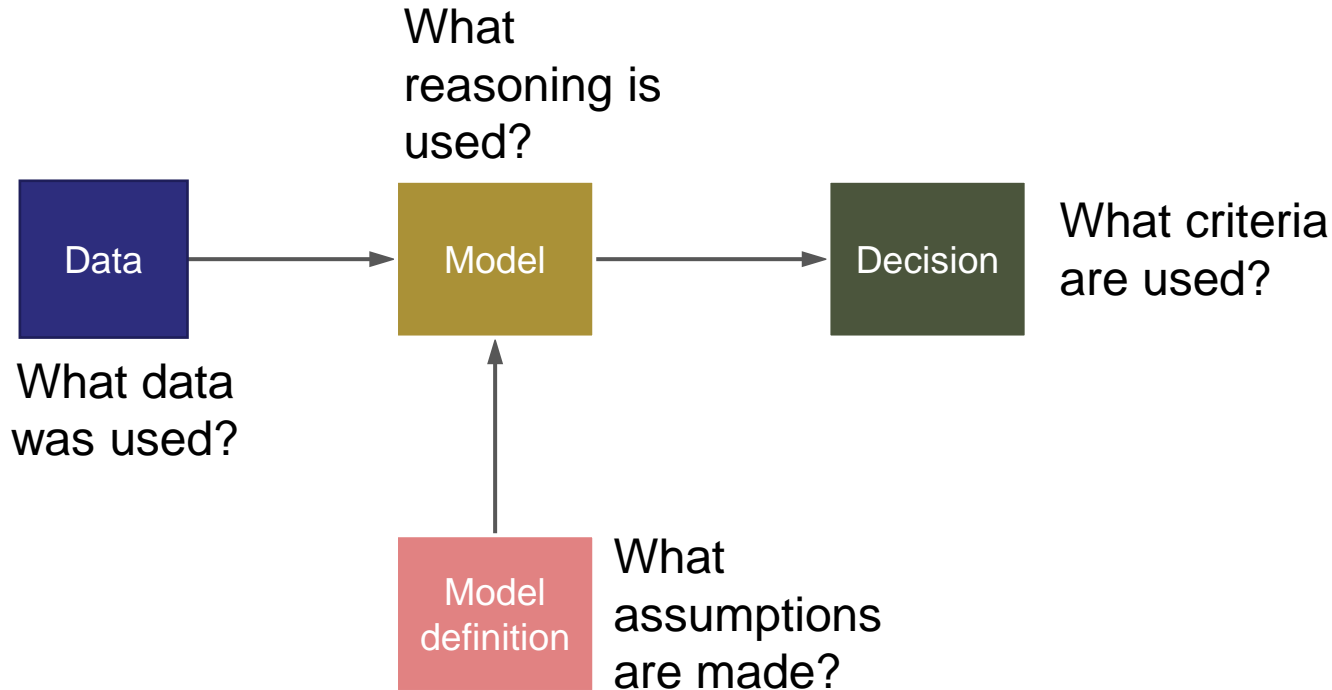
Jheronimus
Academy
of Data Science



We need
transparency!

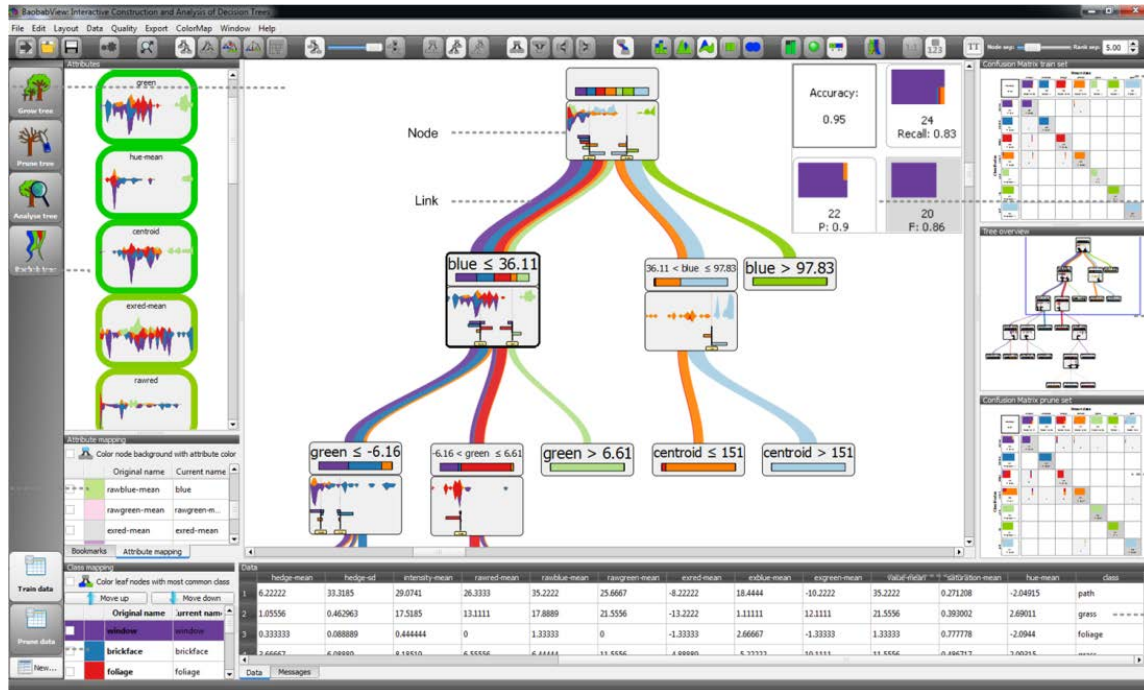
- you may also like that book
- you are not illegible for a loan
- you must get an expensive treatment
- you are fired
- you do not get parole

[2]



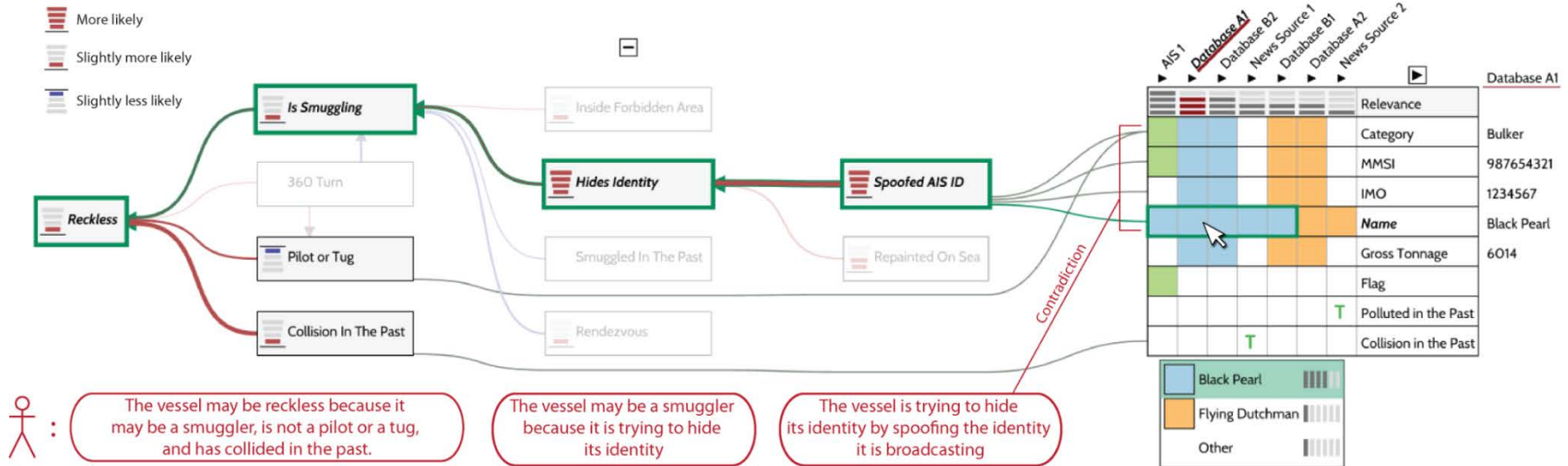
[3]

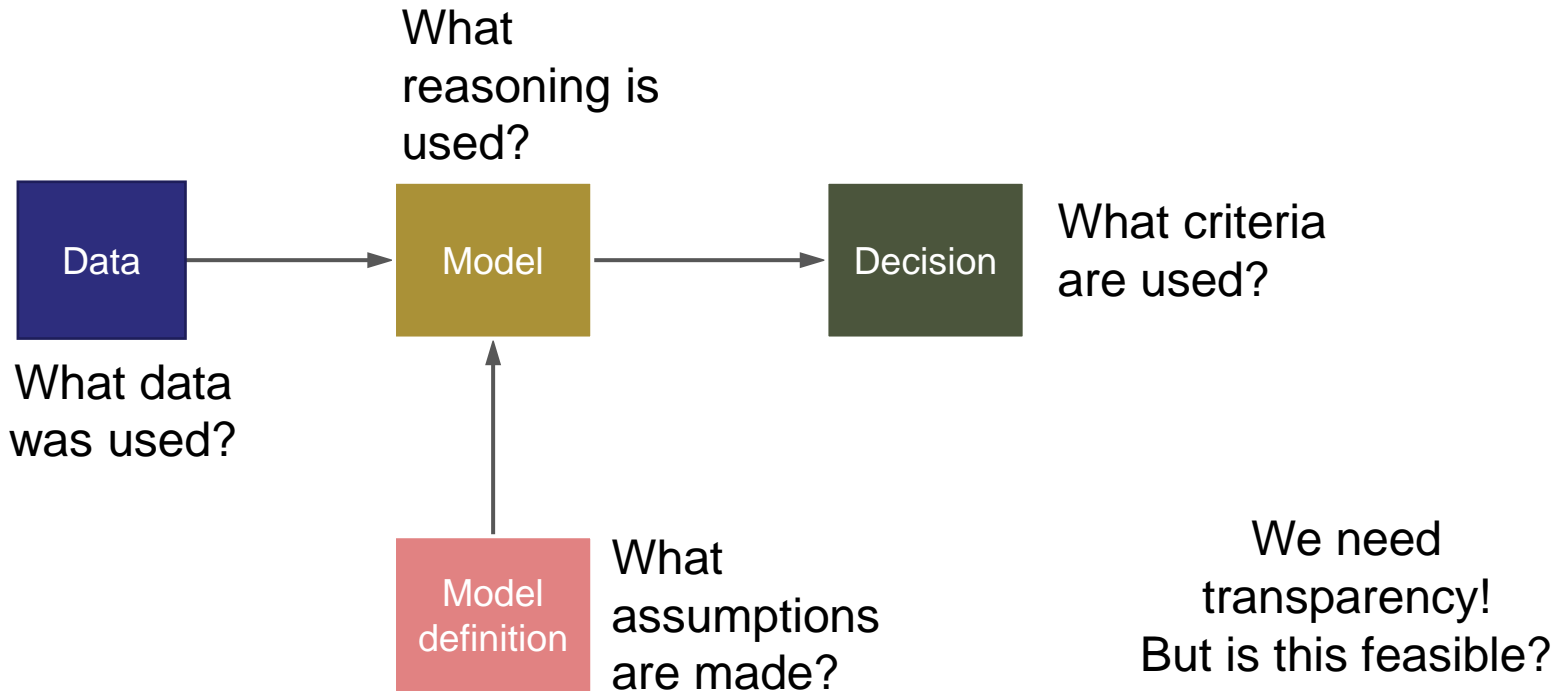
BaobabView: Visualization of decision trees (vd Elzen et al., 2011)



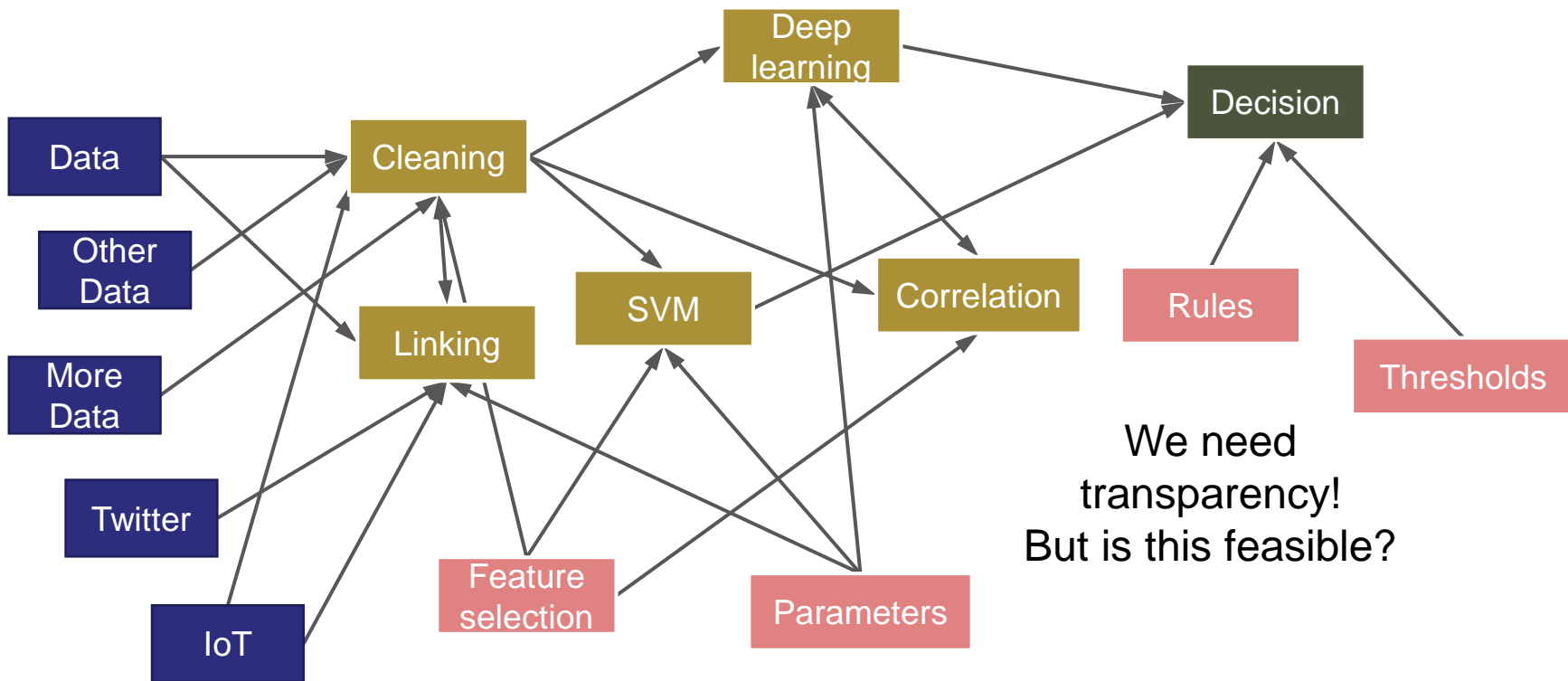
[4]

Rationale visualization (Scheepens et al., 2015)





[6]



We need transparency!
But is this feasible?

[7]

Transparency cannot be achieved, because

- **Processing pipelines are too complex**
 - *no way we can understand what a deep learning network is doing*
- **Understanding requires too much domain expertise**
 - *you can only understand this if you have studied medicine*
- **It blocks innovation**
 - *explaining everything is expensive, methods like deep learning cannot be used*
- **It opens the door to fraud**
 - *tell the outside world what heuristics and proxies we use? no way!*

[8]

To achieve more transparency, we need

A) Better technology?

- *to explain what complex methods are doing*

B) Legislation?

- *enforce that decisions are always clear and transparent*

C) Make companies accountable?

- *companies must be able to explain what they do*

D) Free markets?

- *such that the most transparent companies will win*

E) Consumer organizations, press attention?

- *blame and shame ill-behaving companies and governments*

[9]

