AIX

Why and how transparency increases the success of Artificial Intelligence solutions.

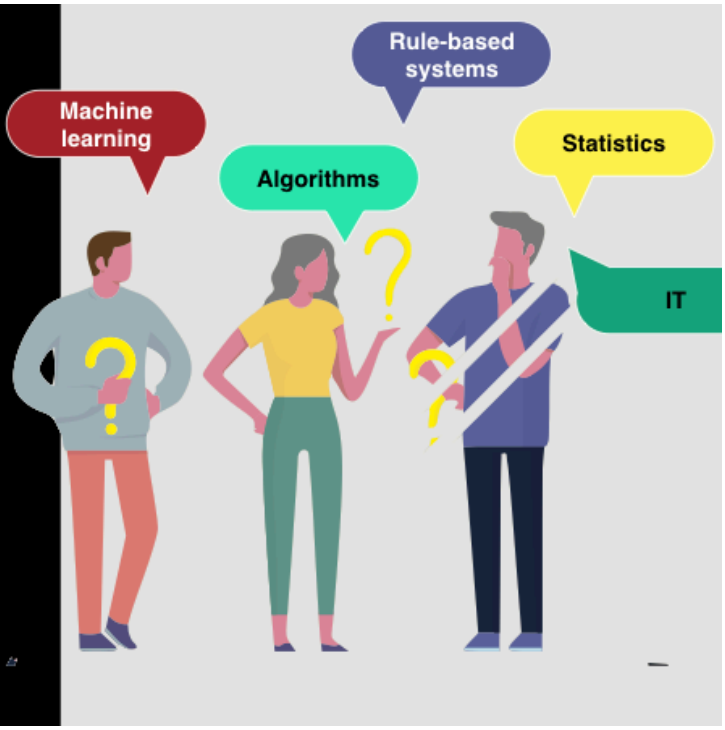
Silvie Spreeuwenberg

Copyright LibRT 2019

AI EXPLAINED

AI is a technology trend referencing any development to automate or optimize tasks that traditionally have required human intelligence.

Silvie Spreeuwenberg
AIX

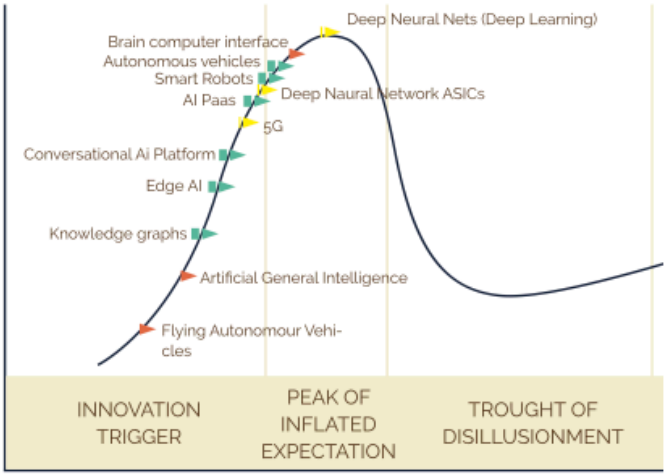


AI HYPE

Startups in AI raise more money.

Everyone will interact daily with AI-based within next decade.

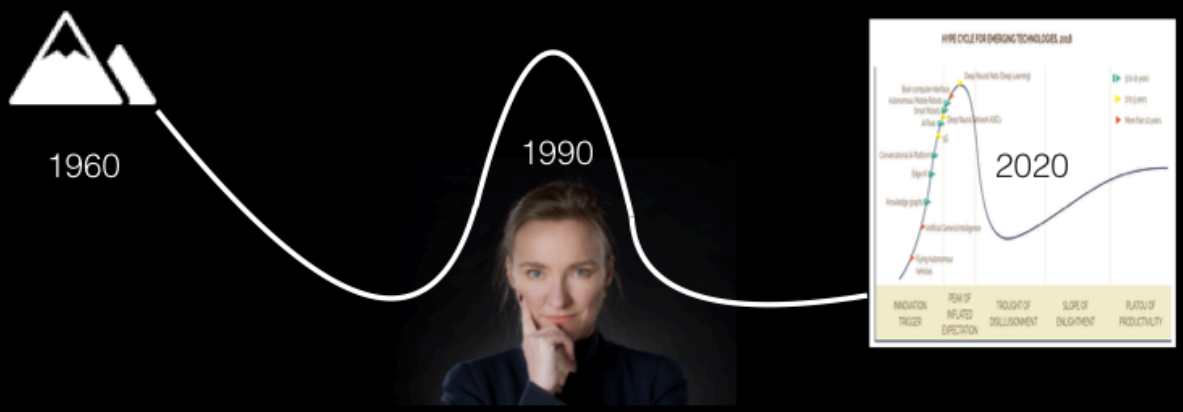
Gartner Hype Cycle for emerging technologies 2018



Copyright LinkIT 2019

THREE AI WAVES

Connectionism **Expert systems** **Machine learning**



Copyright LinkIT 2019

WHAT AI CAN DO

The deceased artist Dali is talking to you, explaining about his life and making selfies with visitors as if he is still alive.

Dali Museum,
St. Petersburg, Florida

WHAT AI BRINGS TO FINANCE ?

Real time transaction monitoring to detect fraud by data analysis of past transactions, expert knowledge and machine learning.

It is not acceptable to close an account based on the risk score of a model that we don't understand.

AI FAILS

Google's algorithm that was trained to help them selecting new hires turned out to be highly gender biased towards man.

POOR DATA

BIASED DATA

MISSING FEATURES

MISSING SITUATIONS

HUMAN DECISION MAKING

"Nobody would say, 'I'm voting for this guy because he's got the stronger chin,' but that, in fact, is partly what happens."

Psychologist Kahneman, Nobel prize 2002

"Should I marry Ann?"

_David

"NO!"

_Oracle

libirt
Copyright LibIRT 2019

5 REASONS

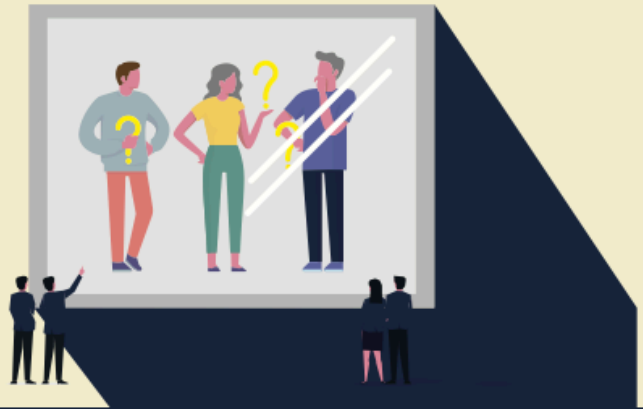
Why XAI is important?

- It will be required
- Better ROI
- More trusted
- Enhanced decision making
- It is feasable

libirt
Copyright LibIRT 2019

1

Society is demanding transparency



People have ethical concerns

Governments prepare legislation demanding explanations

AI technology can be used for the good and the bad.

… it can fake images, fake videos and make fake news.

… it can detect lies and guide people to make better decisions by using large amounts of data.

www.thispersondoesnotexist.com

www.fakenewschallenge.org

Copyright LibRT 2019

3 fundamental rights of democracy are threatened:

1.
Freedom of choice
2.
Access to information
3.
Equality of opportunity

European Commission
Report on trustworthy AI
July, 2019

Suggested readings and echoed or homogenized shared opinions provide us with a tunneled view based on our own patterns.

Copyright LibRT 2019

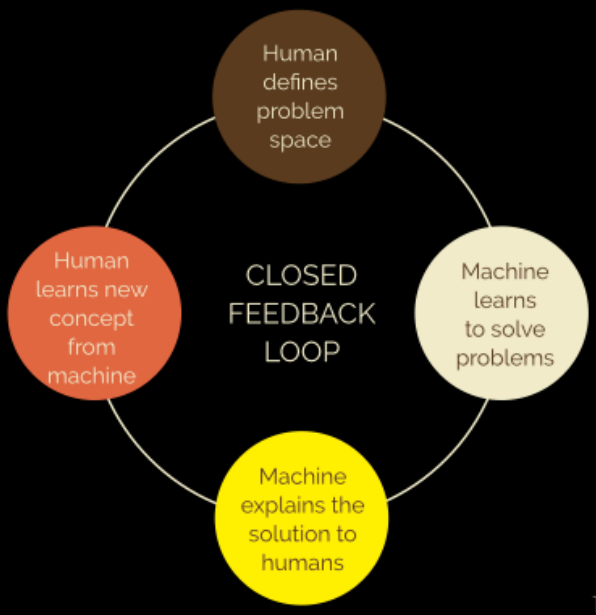
2

Systems that explain themselves have a better ROI

Explanations make a solution easier adaptable to changes

Explanations increase the success rate of AI solutions

Integrated in the value creation life cycle of your business means to adopt timely when your policy, customers or environment changes.



3

Your business will be more trusted

Make more credible decisions

Being a responsible business

The financial sector is commonly held to a higher standard as trust is essential for the system.

When it comes to AI, strong operational controls could be more important then capital and liquidity buffers.

De Nederlandse Bank report on AI in financial sector 2019

3

Explanations enhance decision making

Unwanted decision biases come to the surface

Missing commonsense knowledge is detected

AI model presented with a foil makes the wrong decision.

When it sees ‘baby on blanket’, it concluded ‘animal in wild nature’.

Explainable AI is feasible using a six-step approach

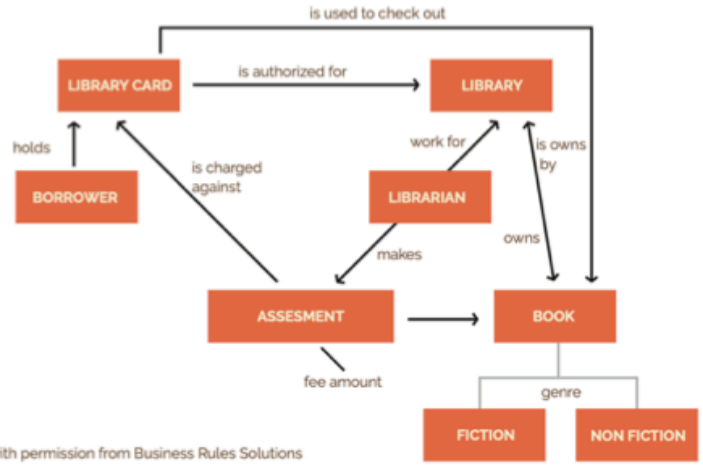
5

1. Get a shared understanding of the domain
2. Understand the task and select the right scope
3. Collect the right data and improve its quality
4. Select AI technology that deliver results
5. Generate good explanations
6. Evolve the system over time

Perhaps your organization is working on an enterprise ontology or data dictionary?

These are good initiatives to align terminology between departments. Hopefully your business stakeholders are involved!

EXAMPLE CONCEPT MODEL OF A LIBRARY USING FACT MODELLING



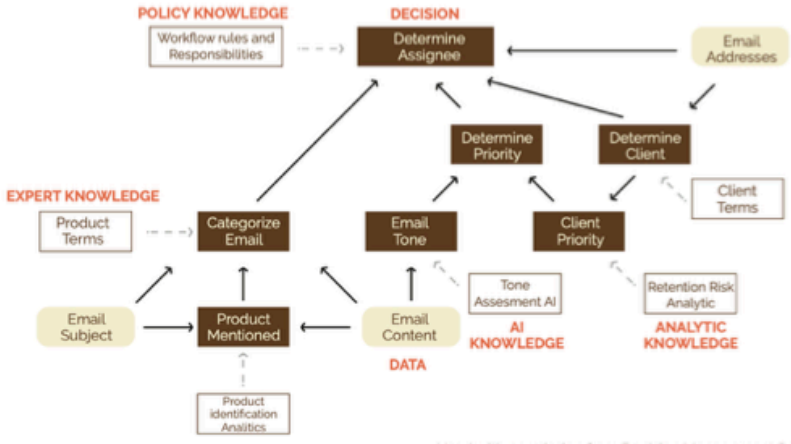
Used with permission from Business Rules Solutions

Copyright LibRT 2019

Break-up the question into potential influencers and phrase them as a testable hypothesis.

Test them using basic statistical analysis on the available data and create a root-cause analysis.

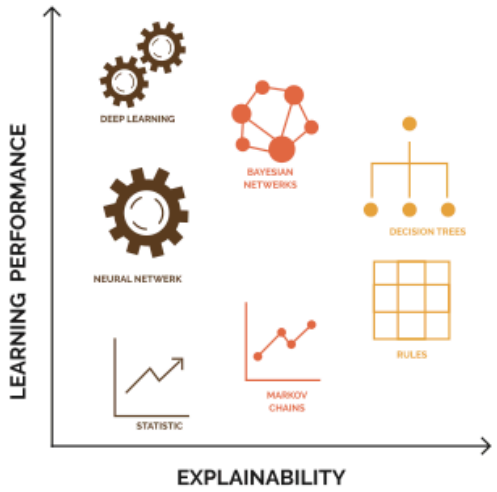
EXAMPLE DECISION MODEL USING DECISION MODELLING NOTATION



Used with permission from Decision Management Solutions.

Copyright LibRT 2019

XAI combines the technology of the three AI waves

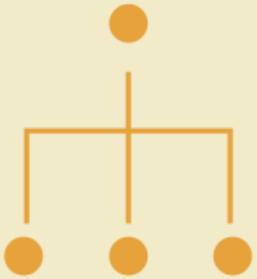


Copyright LeiRT 2019

Clinicians find it more important that the system help them to justify their clinical decision.

S. Tonekaboni et al, Contextualizing Explainable Machine Learning for Clinical End Use, 2019

Copyright LeiRT 2019



MACHINE LEARNED MODEL

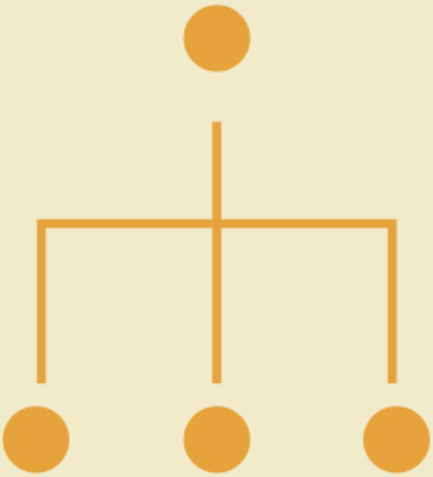
EXPLANATION MODEL

INTERACTION MODEL

Accuracy, robustness and precision are not the way to convince a domain expert.

Instead we need to develop seperate models to support communication of an explanation in a context

Copyright LiuKT 2019



Be alligned with the user's beliefs.

Indicate what relevant data is used.

Indicate when the model i s not applicable.

Be transparent and follow common sense logic

EXPLANATION MODEL

INTERACTION MODEL

Copyright LiuKT 2019

An explanation does not need to be complete

$$\text{GOOD EXPLANATION} = \frac{\text{Believable} + \text{Relevant} + \text{Distin-guishable}}{\text{Complexity}}$$

Copyright LibRT 2019

RESEARCH ON XAI

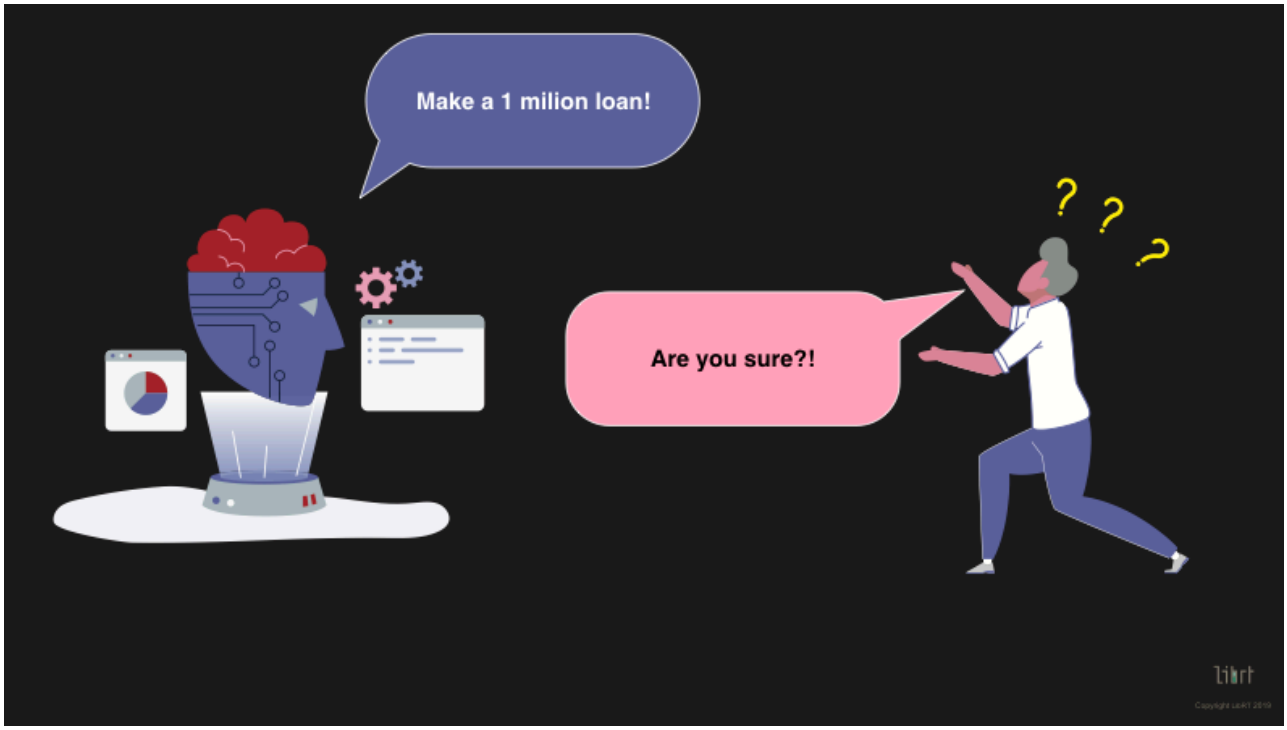
DARPA
Allen institute for AI

Explainability, next to accuracy and precision, is one of the optimizations criteria in the model.

Improve techniques that learn explainable models such as Bayesian networks.

Understand how to use techniques to generate explanations, such as decision tree generation, from black box models.

Copyright LibRT 2019


Make a 1 milion loan!
Are you sure?!

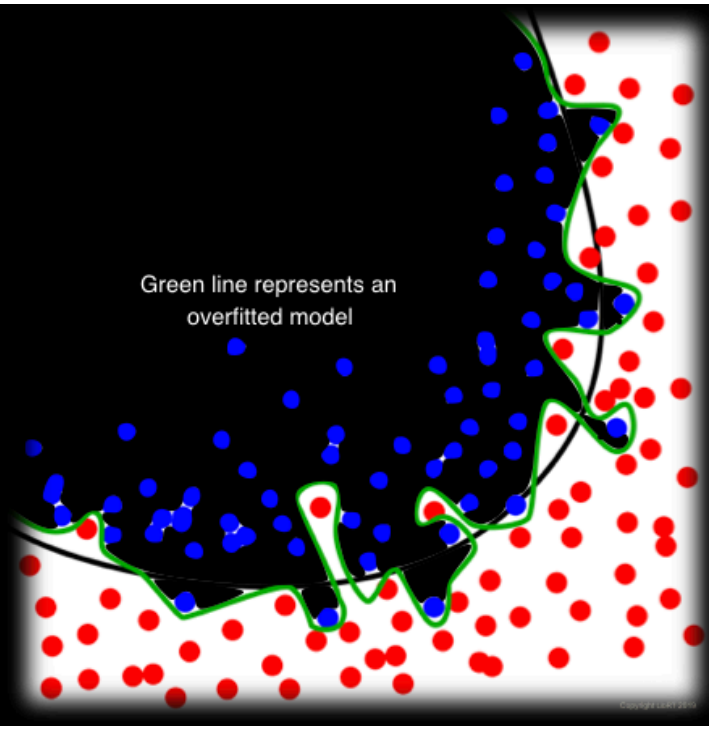
Don't make that loan.
Why?
KEEP THE HUMAN IN THE LOOP

ARE WE FIT?

How te measure the explainability of a model?

How to incorporate common knowledge in machine learned models?

How to present an explanation as a meaningful conversation?



www.xaix.com

THANK YOU

AI^X

Artificial Intelligence needs explanation: pre-order now!

available at amazon

Discount code 20%

AIX-RULEML-19

Follow me

Silvie Spreeuwenberg

Copyright LibRT 2019