

Thinking outside of the box – building reliable and scalable data analytics products

Dr. Stephanie Thiemichen

Senior Data Scientist @ TÜV SÜD Digital Service GmbH

German Data Science Days 2020

An Example



**Mehr Wert.
Mehr Vertrauen.**

**Add value.
Inspire trust.**

TÜV SÜD at a glance



150+
YEARS OF
SAFETY, SECURITY
& SUSTAINABILITY



1,000+
LOCATIONS
WORLDWIDE



€2.5
BILLION
IN ANNUAL
REVENUE



24,500+
EMPLOYEES*



41%
OF REVENUE
OUTSIDE GERMANY^



574,000
CERTIFICATES



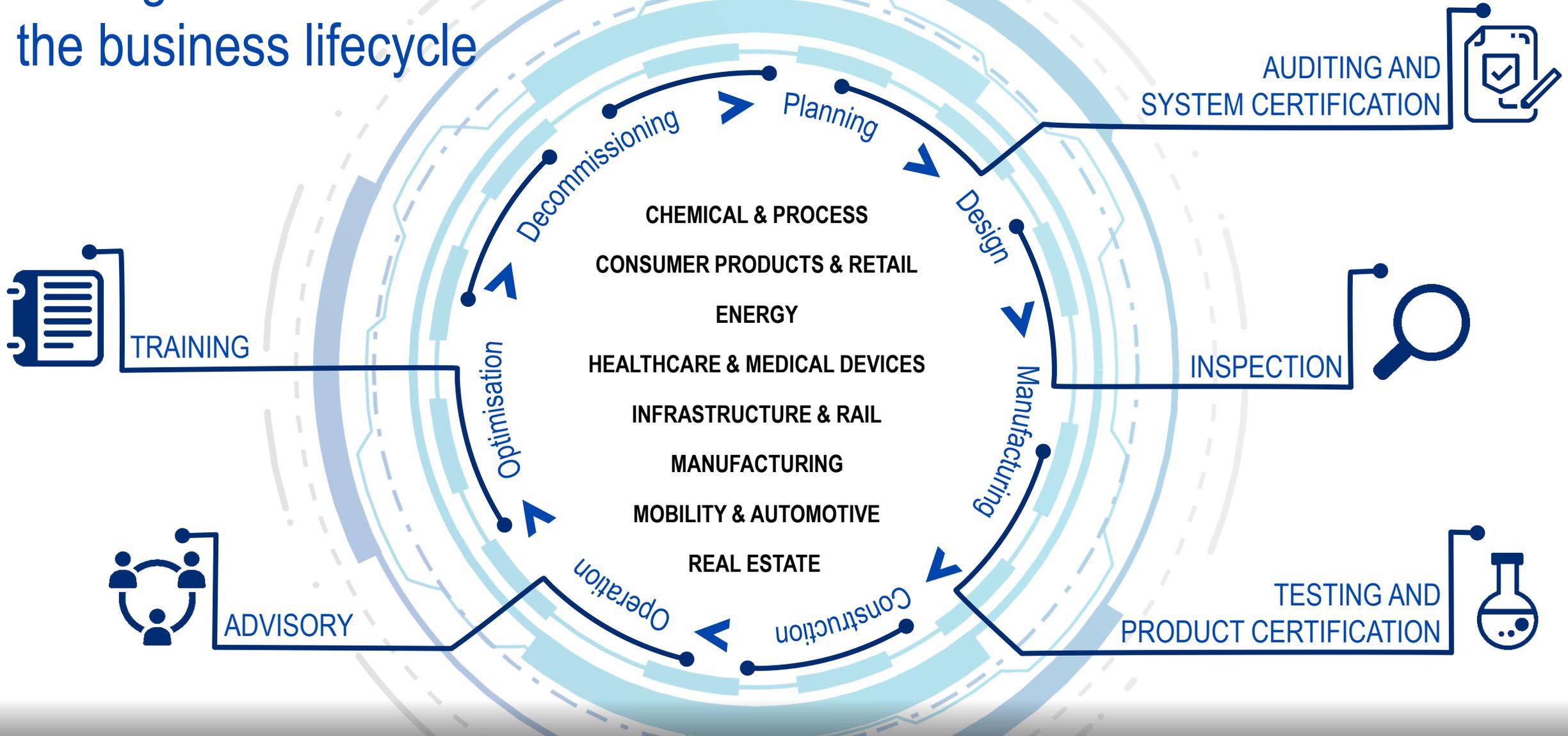
100%
INDEPENDENT
& IMPARTIAL



1-STOP
SOLUTIONS
PROVIDER

*As of 2018-12-31
^Based on clients' locations
Note: Figures have been rounded off.

Adding value across the business lifecycle



Agenda

1 Setting

2 Challenges

3 Solution

Agenda

1

Setting

2

Challenges

3

Solution

Task: Automated Damage Assessment



Setting

Use Case Examples

- Used Car Value Assessment (OEM)
- Risk Transfer (Fleet Rental)
- Fast Feedback of Car Condition (Logistic)

DVS – Digital Vehicle Scan



Setting

Box Setup (Indoor)

- Drive-through solution for the assessment of the condition of vehicles during their collection and return
- Vehicle's condition is recorded in only 30 seconds
- Up to 22 cameras scan the entire vehicle (including the underside)



Setting

Agenda

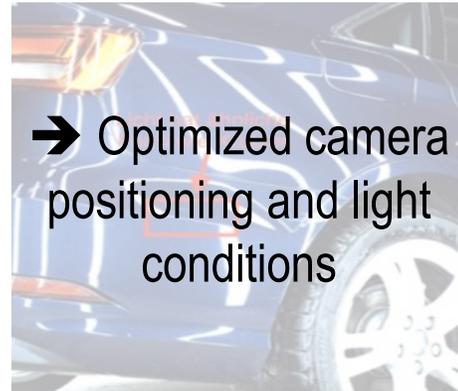
1 Setting

2 Challenges

3 Solution

Data Quality – was an Issue in the Beginning

- A human subject matter expert needs to be able to see the damage on the pictures.
- The more controlled and standardised the process / the environment the better.



Reflections



Water Marks



Overexposure



Resolution too low for certain vehicle parts

Computer Vision is a Challenge

Despite the recent advances in the field of computer vision, it is still a challenging task with respect to

- The amount of labelled data needed
- The computing infrastructure needed for training and scoring
- The available time and budget
- The people / expertise needed to build an end-to-end scalable computer vision based system (Data Scientists, Solution Architects, Software Engineers)
+ the subject matter experts





Agenda

1 Setting

2 Challenges

3 Solution

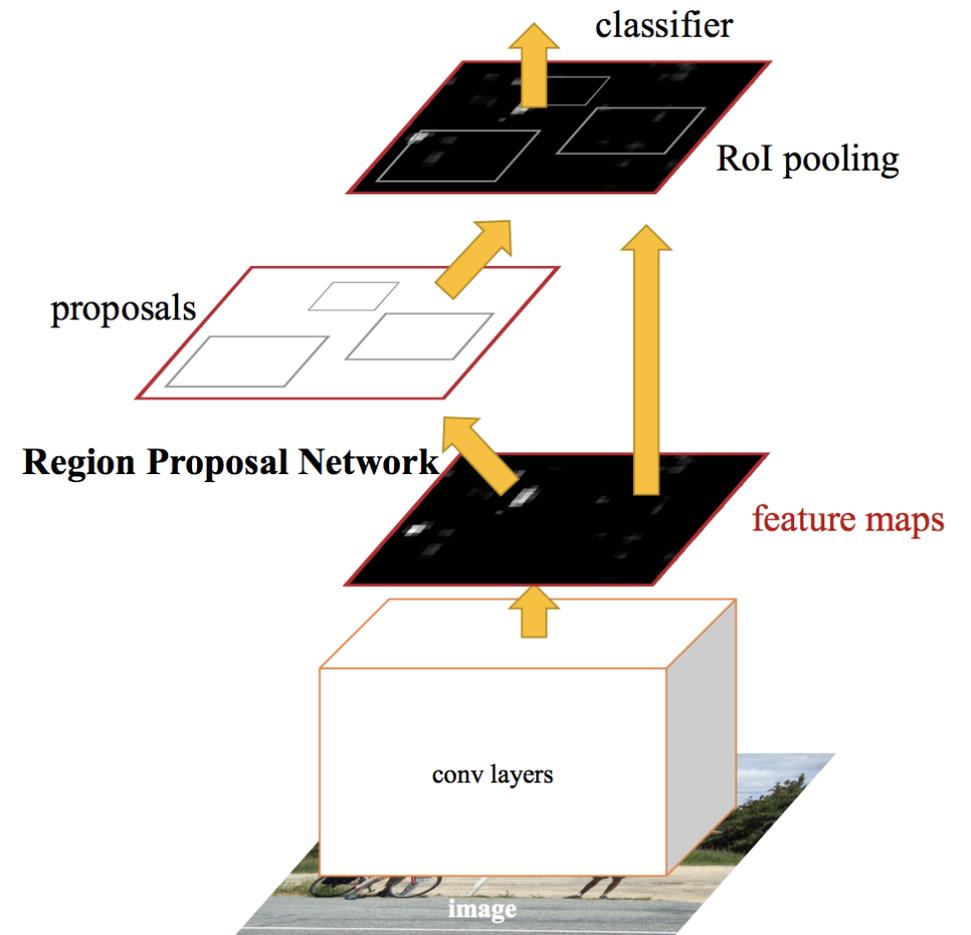
Data Science Approach for Object Detection

What?

- Convolutional Neural Network for simultaneous localization (where?) and classification (what?) of damages
 - Faster R-CNN / RetinaNet (<https://arxiv.org/abs/1708.02002>)
- Transfer-Learning to reduce amount data needed
 - Pretrained on Imagenet / COCO

How?

- Train, validate and test with labelled images of defects
- Active Learning Pipeline for retraining and improving model over time based on expert feedback (Backoffice)



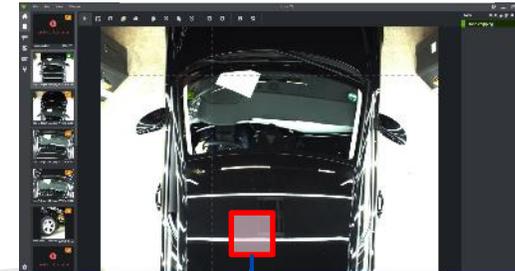
Source: <https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e>

Architectural System Overview

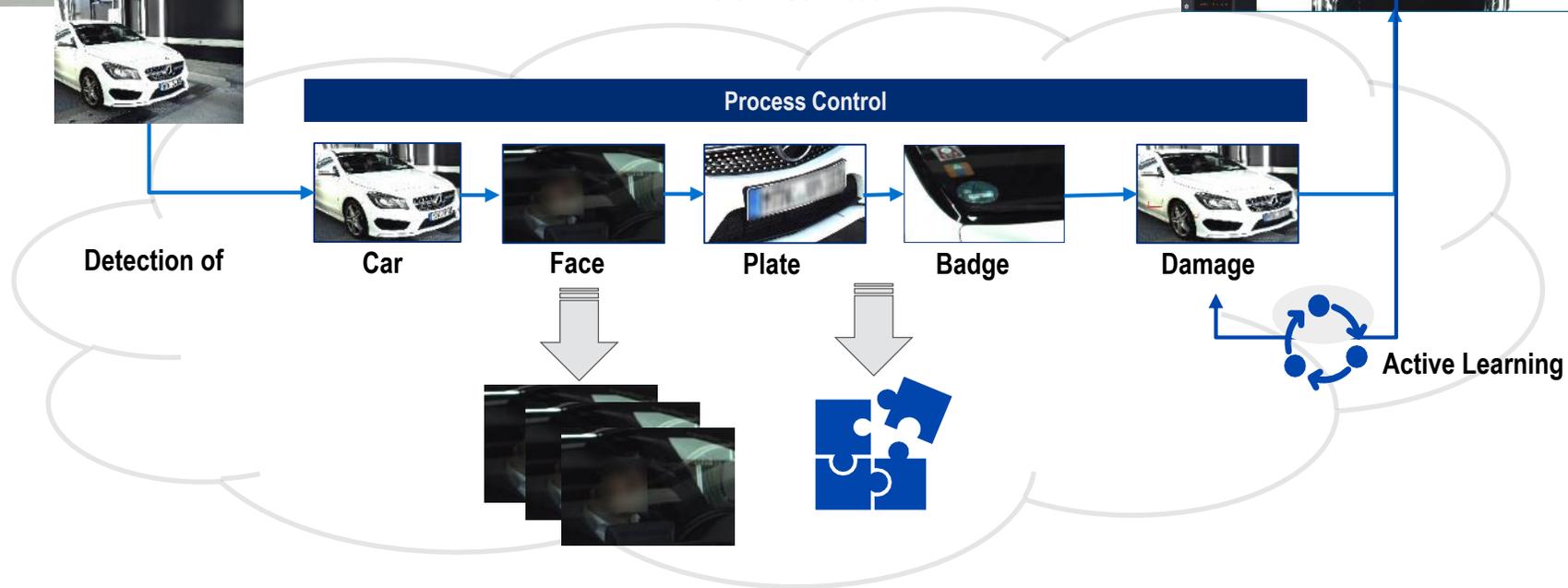
DVS



TÜV SÜD Backoffice

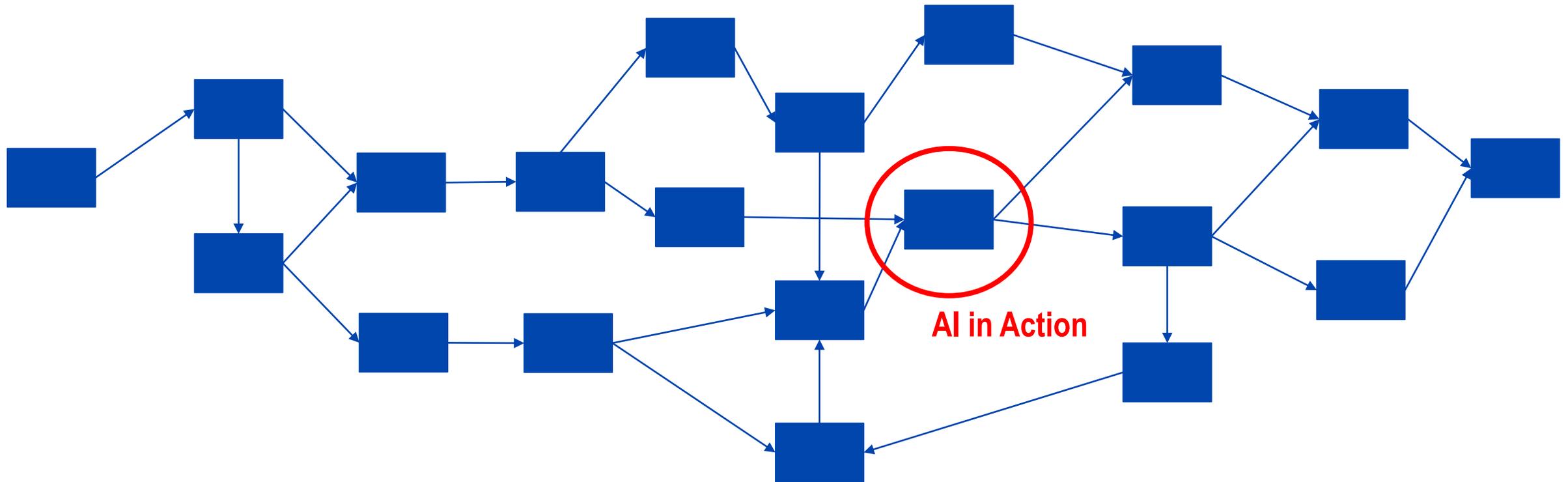


Cloud Services



Solution

The whole picture: AI alone is not the solution!



Solution

Key Learnings as a Data Scientist



The whole process matters.



Key question: Who is your user?



It's a team effort!



Keep it simple.



Jupyter notebooks don't scale very well.



Emphasise communication and explanation.

*So certain were you. “
Go back and closer you must look.”*

Master Yoda



Solution



**Mehr Wert.
Mehr Vertrauen.**

**Add value.
Inspire trust.**

Dr. Stephanie Thiemichen

stephanie.thiemichen@tuev-sued.de

TÜV SÜD Digital Services GmbH

Westendstr. 199

80686 München