

Hur AI påverkar utvecklingen inom Trafikverket

2023-05-02

Daniel Jakobsson, Strateg
Central Funktion Ekonomi & Styrning
Datadriven verksamhet och analys

Introduktion

- AI utvecklingen har startat i som av organisationen
 - Behovs driven utveckling med villighet till testa och försöka
- AI projekten i Trafikverket
 - Exempel inom Bildanalys, Tidsserier, chatbott,
 - Gjorda som forskningsprojekt, förvaltningsuppdrag och PoC
- Nästa steg i AI Utvecklingen
 - AI förmågorna under uppbyggnad
 - Etiska frågor

Bild analys: Personligintegritet för bil bilder

[PMSv4 - Start \(trafikverket.se\)](https://trafikverket.se)

- PMSv4 eng. Pavement Management System
- Mer än 20 miljoner unika bilder
- Mer än 3 Terabyte data för alla bilder
- 2018 enklare algoritm (Open CV, Haar-Cascade)
- 2020 mer avancerad algoritm (Yolo)



Tidsserie: Prediktion väg slitage

- Väg slitage på 98 500 km statliga asfalterade vägar
 - Mäter Spårdjup med laser
- Prediktion över slitage för koordinering av ombeläggning
 - Man tittar på prognoser spårdjup utveckling över tiden

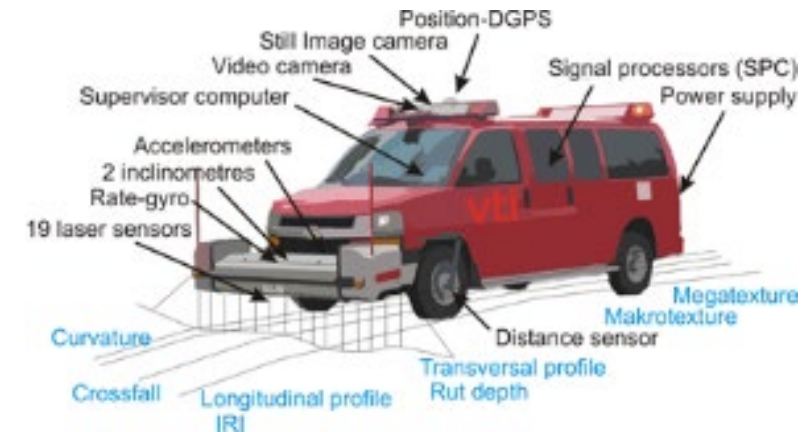
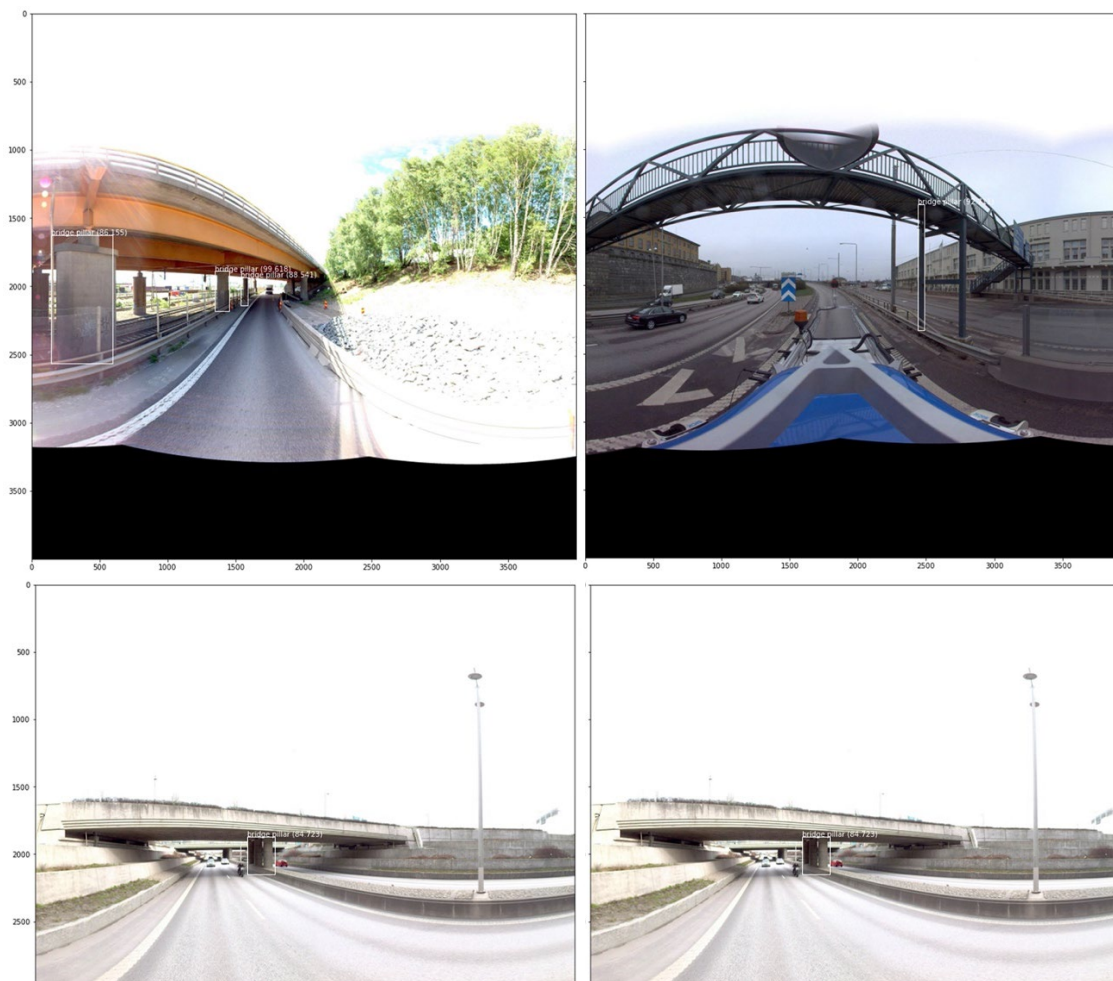


Bild analys: PoC Objekt klassificering i väg

- För att höja säkerheten är det viktigt att Trafikverket har koll på farliga objekt i säkerhetszonen bredvid vägen.
- Bropelare tränas för identifiering med 360 bilder.
- Litet data sätt totalt 314 bilder
 - Träningsdata 251 bilder 80%
 - Validering/Test 63 bilder 20%
- Algoritm: Yolo med egen tränade objekt

Träning

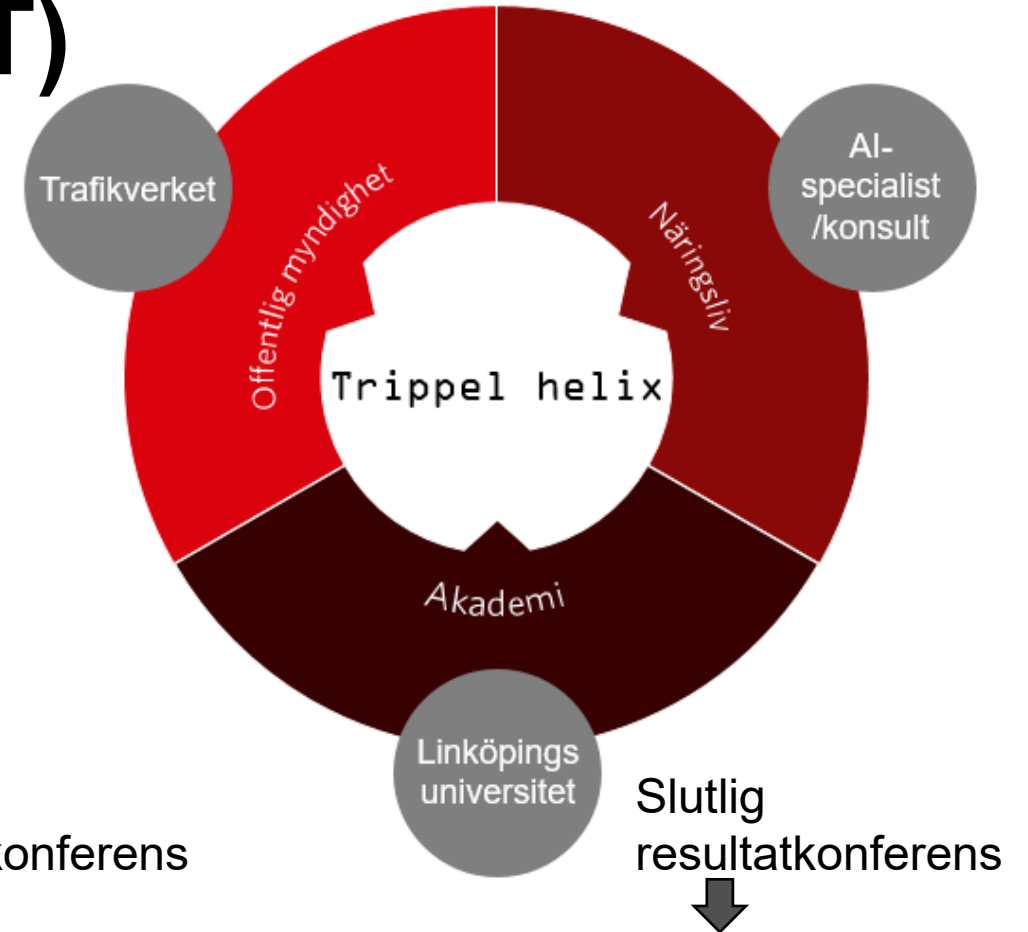
Test



Tidsserie analys:

Projekt AI baserad Realtidsprognostisering av Trafikinformation (AIRT)

- Forskningsportfölj: Möjliggöra
Målområde: Mer användbar trafikinformation och säkrare prognoser
Fokusområden:
 - Prediktion av tidsåtgång för trafikstörande händelse
 - Prediktion av ankomsttid av tåg till station
 - Vädrets påverkan på järnvägssystemet
 - Visualisering



Projektstart
05 01

Analysmiljön
tillgänglig
01 01

Interim
resultatkonferens
10 19

Slutlig
resultatkonferens

2020

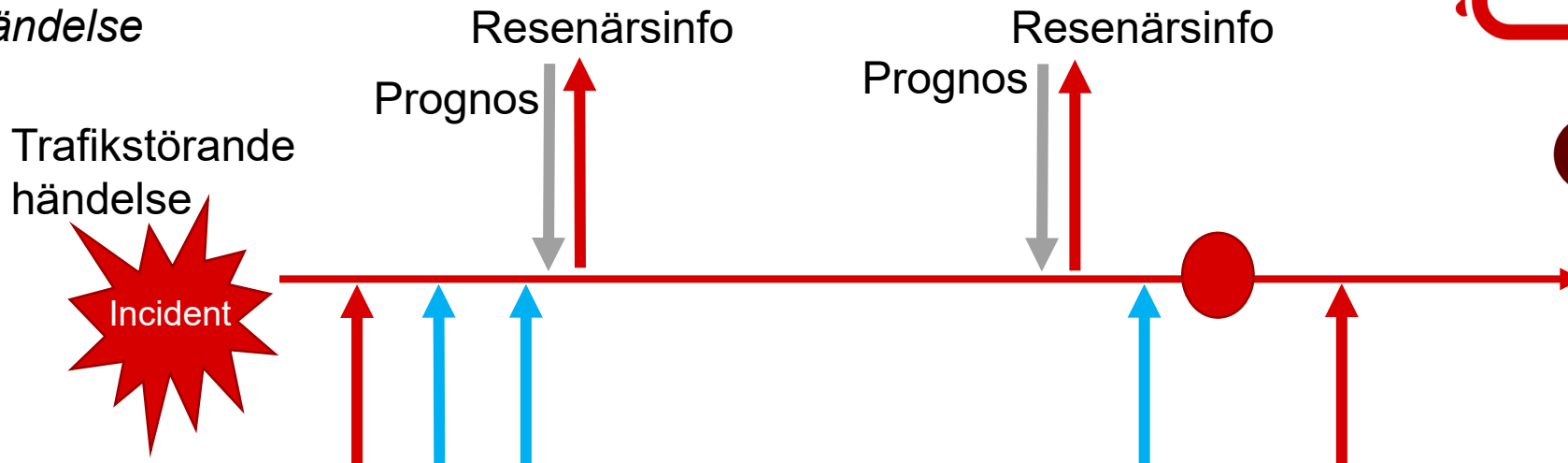
2021

2022

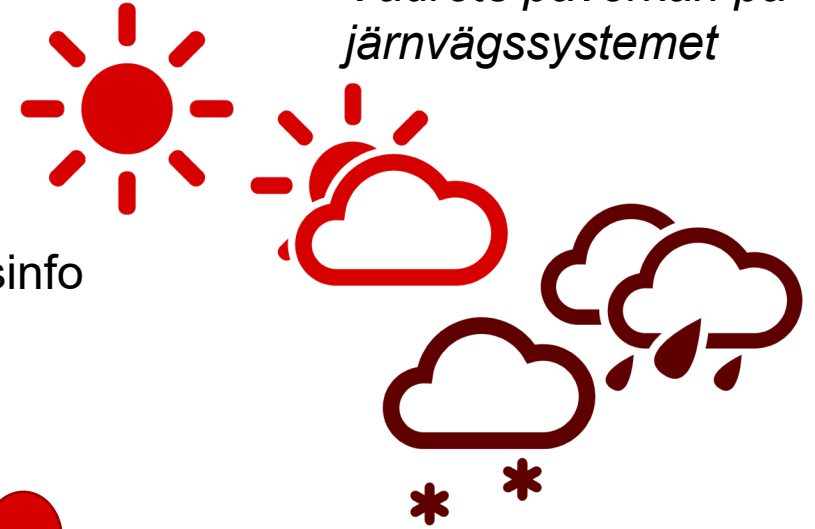
2023

Fokusområden

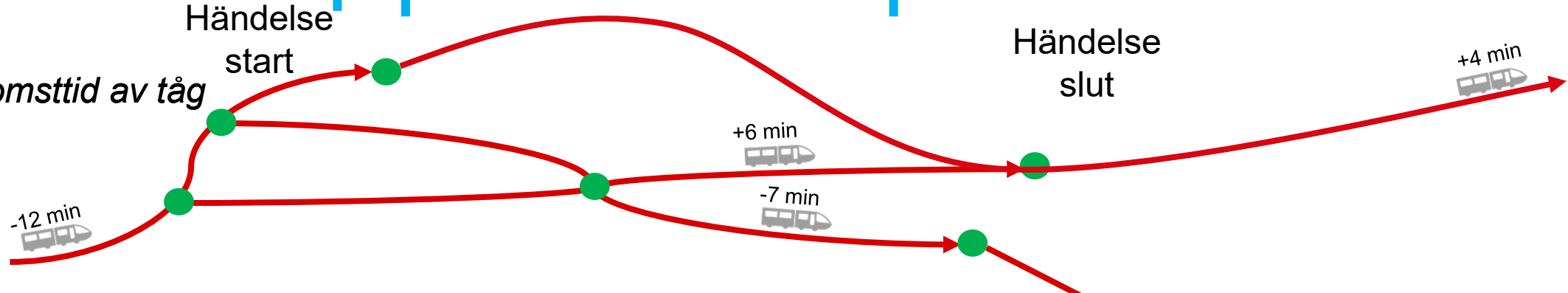
Prediktion av tidsåtgång för trafikstörande händelse



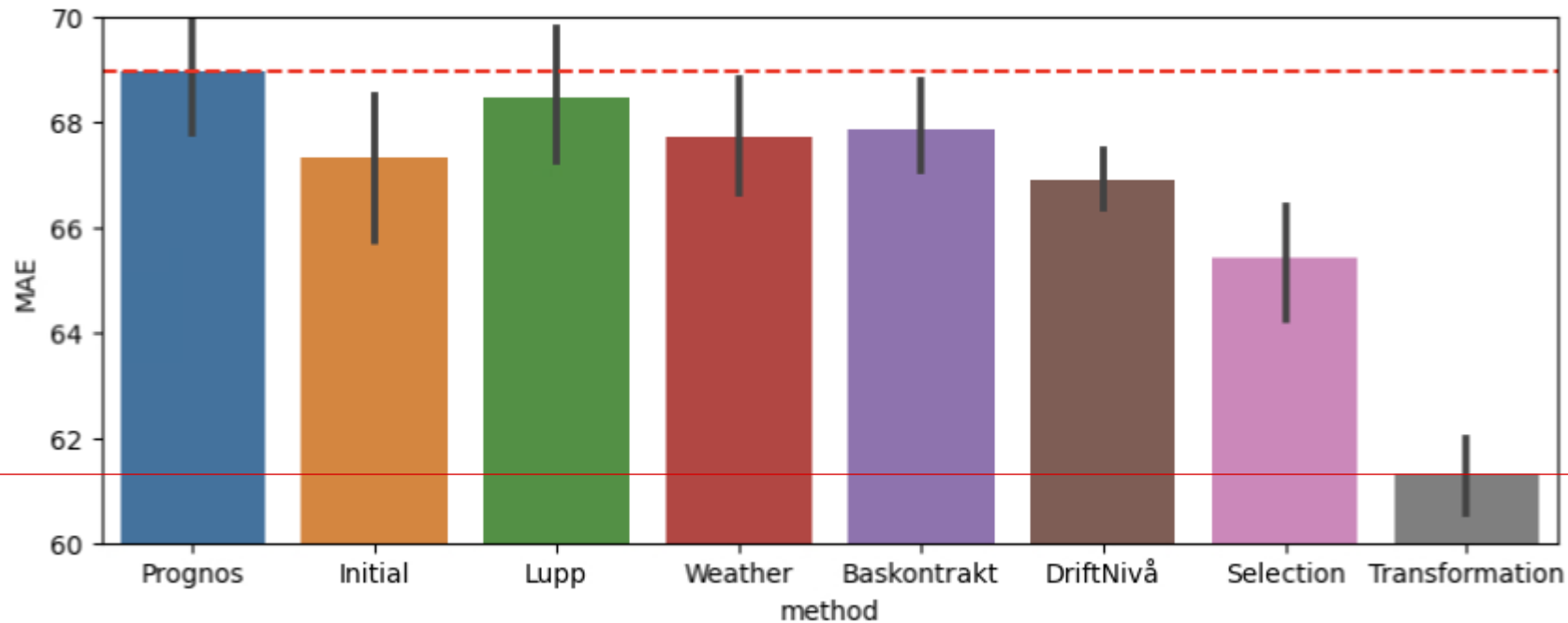
Vädrets påverkan på järnvägssystemet



Prediktion av ankomsttid av tåg till station

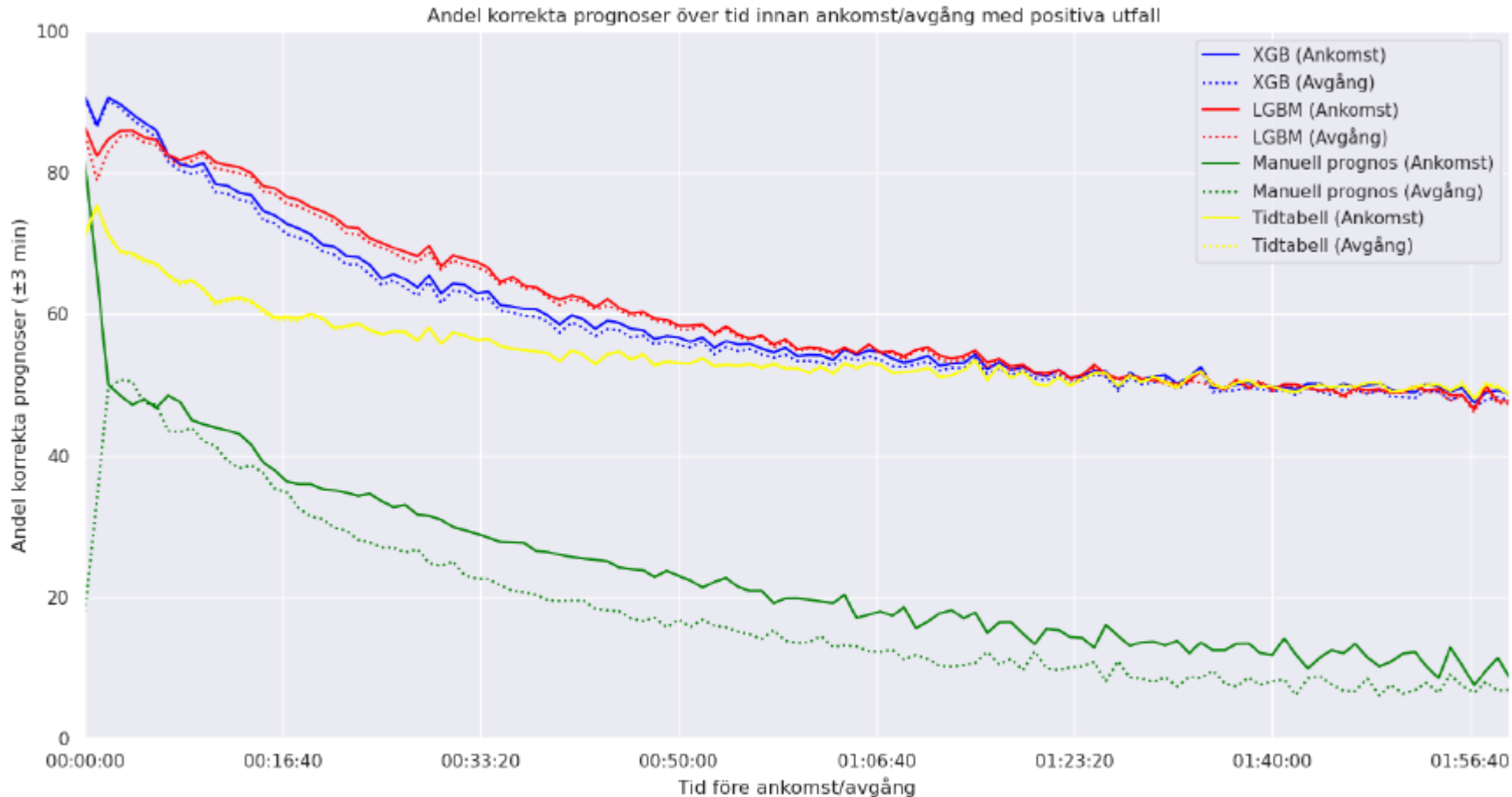


Resultat Trafikstörandehändelser



Trafikstörandehändelser har resulterat till i snitt ca 7 minuter (11%) bättre än manuella prognoser.

Tågprognos ankomst/avgång



Korrekt andel prognoser (± 3 min) över tid innan ankomst/avgång. Resultaten visar AI-modellerna och baselines.

- Tåg prognos ankomst har resulterat till i snitt ca 37 minuter (62%) bättre än manuella prognoser.
- Validerat vid 50 min innan slut station

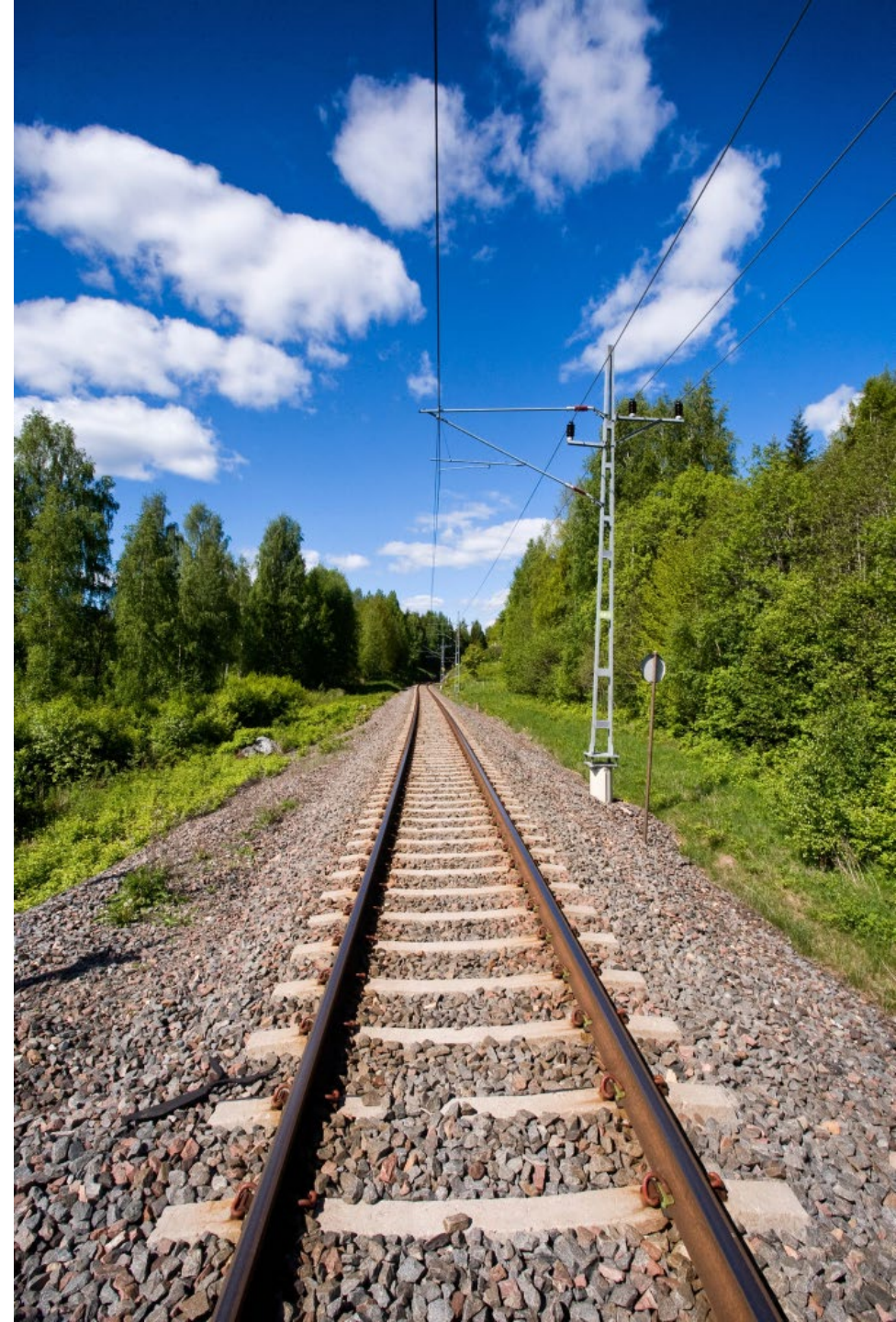
AIRT 2020-2022, Exempel på enkel start

- AI en resa från första steget:
 - PC på steroider (dubbla GPU, vatten kylning av CPU)
- Sandbox utveckling
 - Multi user med OS Ubuntu
- Virtual Environment, Python
- JupyterLab, VSC
- Tensorflow, Keras, Scikit-learn
- XGBoost, LightGBM, NN



Bild analys: Defekta Järnväg cement slipers

- Digital Besiktning
 - Mål att minska antalet besiktnings personer i järnvägs spår
 - Tåg fotograferar banan alla delar automatisk
 - En slipers leverantör av cement slipers får sprick bildningar
 - Vi använder AI för att identifiera sprickorna
 - Moment:
 - Annotering (Skapa träningsdata)
 - Kroppa (att välja ut slipersdelen)
 - Sprick detektering



Annotering

- Bildannotering – ge etiketter till bilder i ett dataset för att skapa träningsdata till ML modellerna
 - Olika sätt att annotera beroende på uppgift (detektion, segmentering, klassificering, ...)
- LabelMe
 - Verktyg skrivet i Python
 - Polygon, rektanglar, cirklar, linjer, punkter
 - Saknar viss annoteringstyp som vi behöver
- CVAT (Computer Vision Annotation Tool)
 - Har fler annoteringstyper som t.ex pensel (sprickannotering)
 - Bra integration med FiftyOne
 - Kan sättas upp som en egenhostad lösning





● Unsaved view

☰ Samples

+



🗑 LABEL TAGS +

METADATA -

metadata.size_bytes

metadata.mime_type

metadata.width

metadata.height

metadata.num_channels

LABELS 3 ✓ -

prediction_v2

prediction_v3

fastening

PRIMITIVES 🖱 3 ✓ -

id

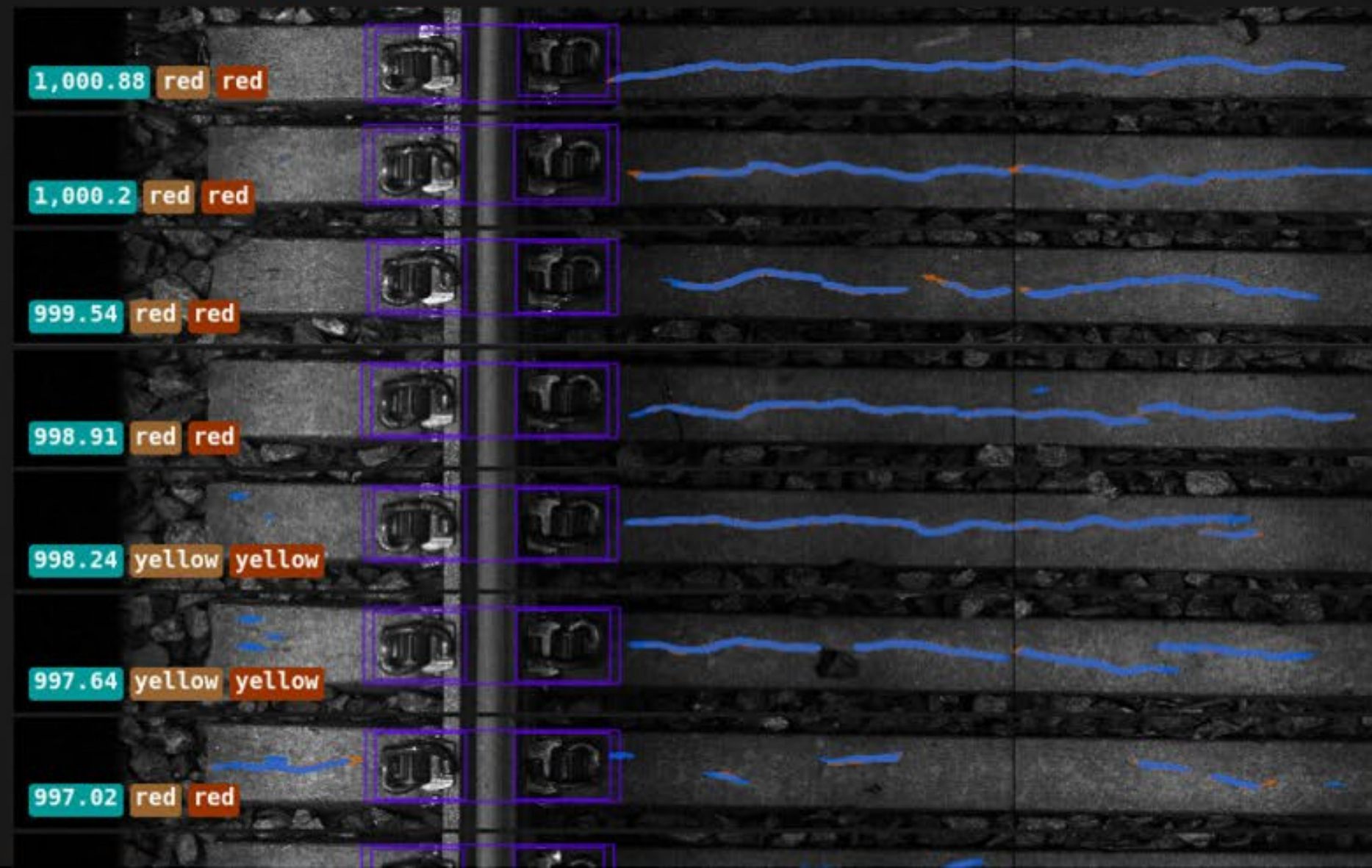
filepath

def_class_v2

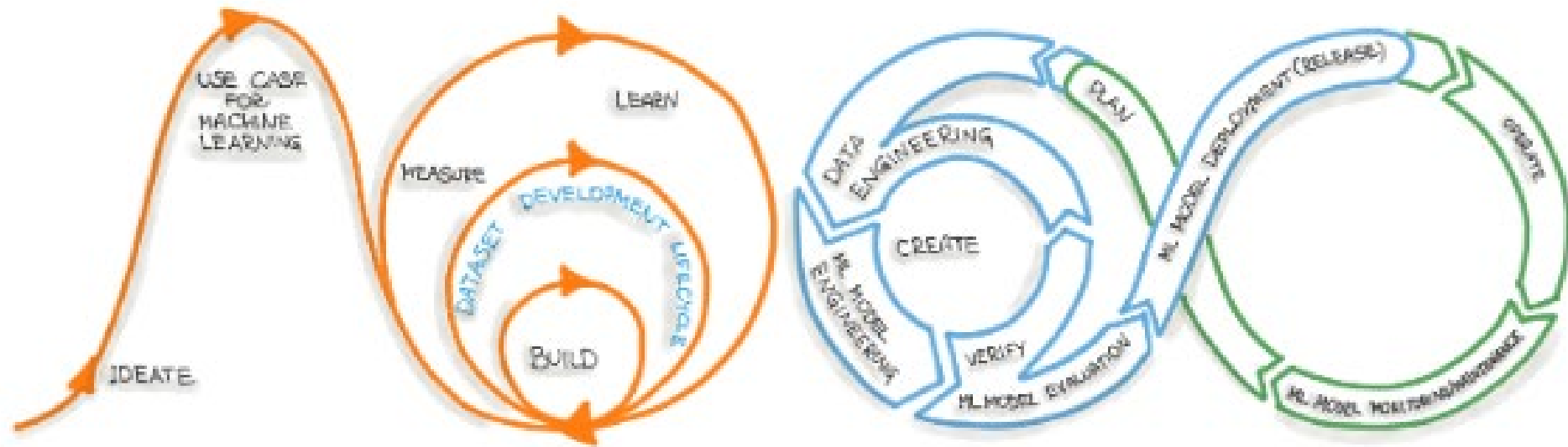
def_class_v3

def_class_truth

seq



CRISP-ML(Q)



PHASES

BUSINESS & DATA UNDERSTANDING

MODEL DEVELOPMENT

MODEL OPERATIONS

@visonger

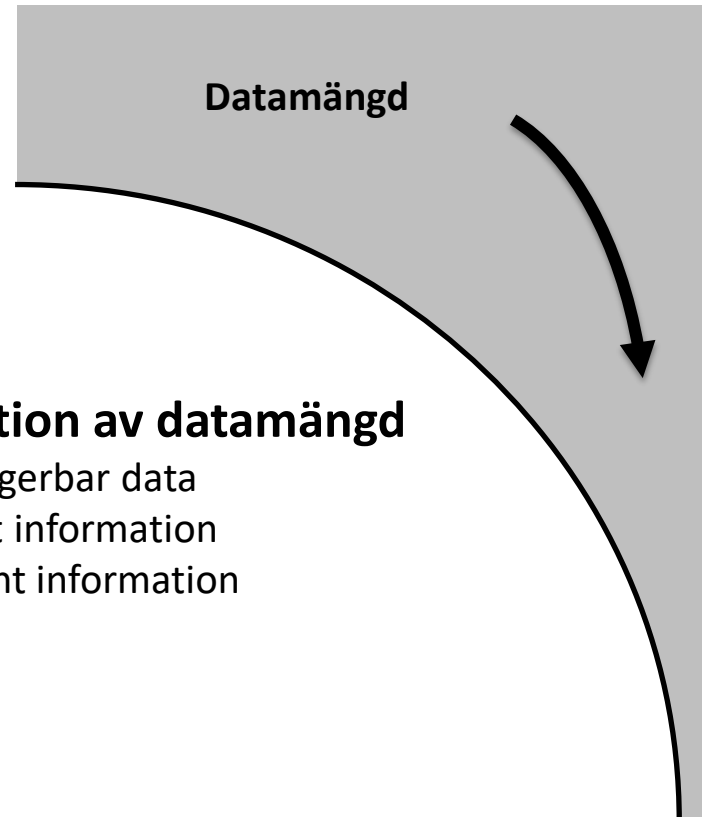


Data Readyness

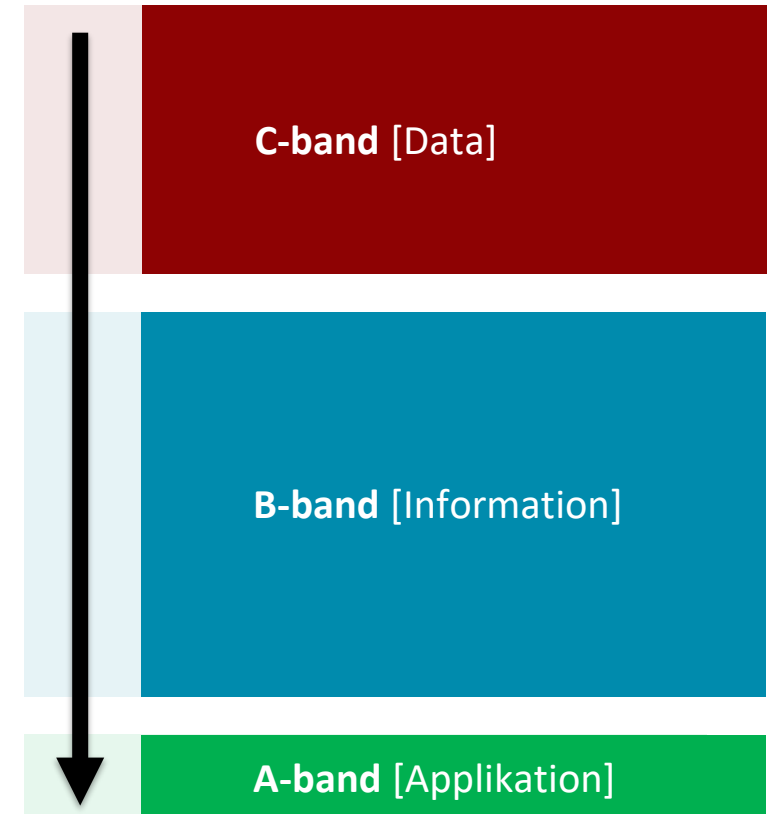
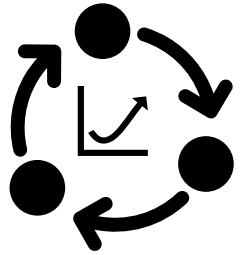
- Mål
- **Data**
- Kompetens
- Verktyg
- Process

Stor reduktion av datamängd

- Icke-korrigerbar data
- Irrelevant information
- Redundant information

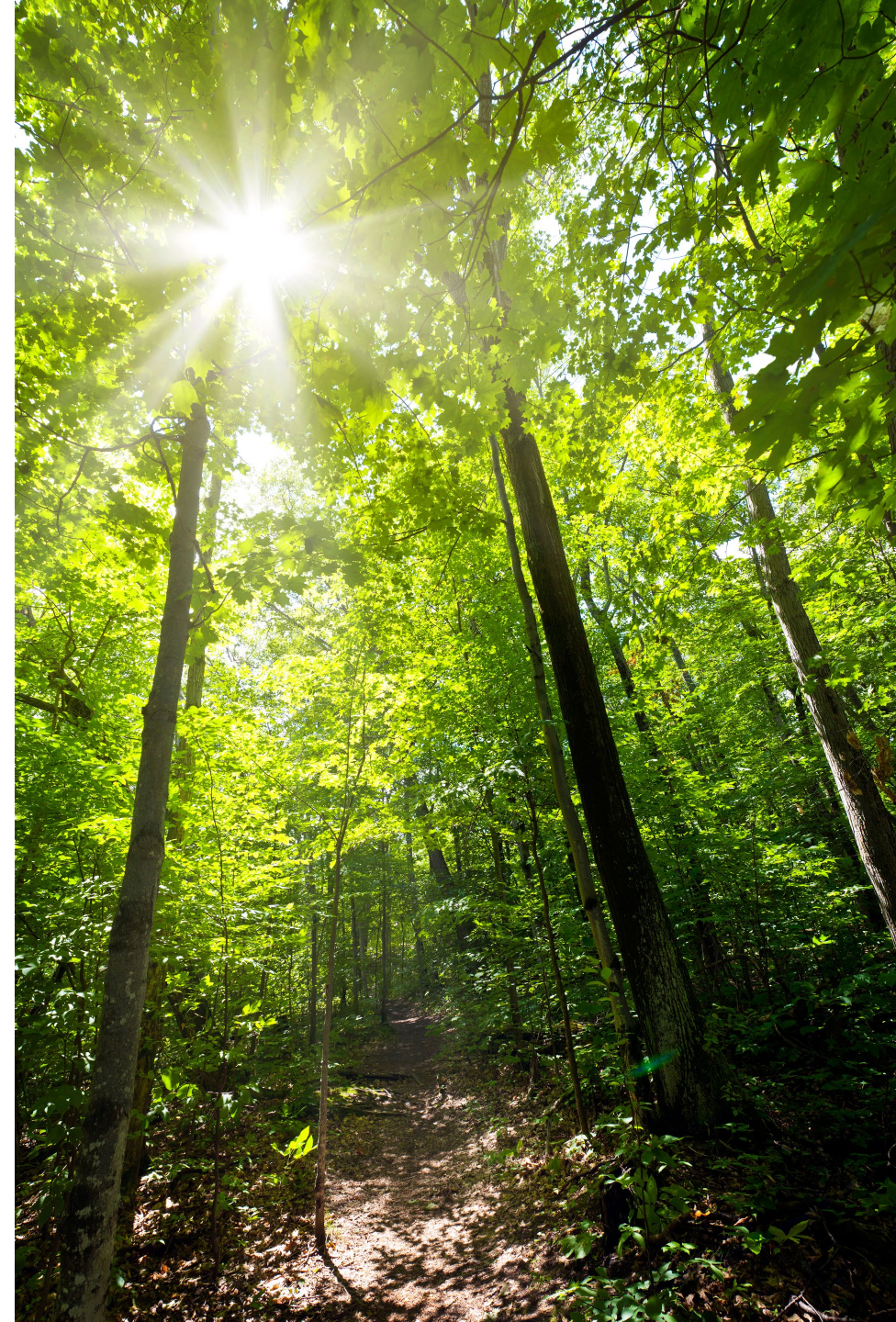


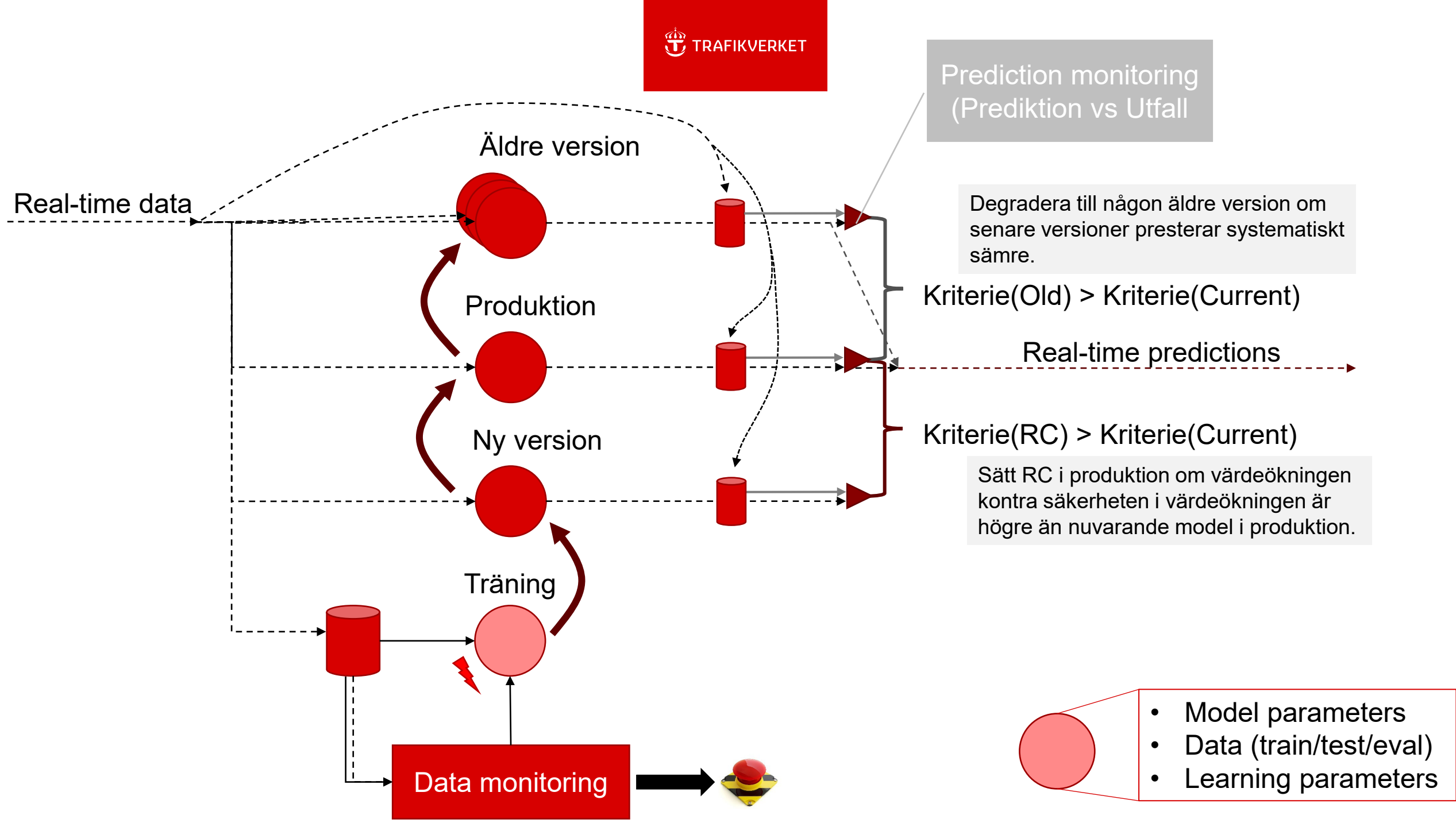
Data Readiness Levels,
N. D. Lawrence 2017



Utmaningarna framåt

- Vi behöver bygga förmåga MLops i Organisation
 - Vi har behov av stöd för AI i realtids applikationer inklusive monitorering, skugg model utveckling i en support struktur
 - Utveckling, Produktions test och Produktion
- Data utbytes plattform delar publicerad data var än behov och källsystem finns i domän arkitekturen
- Informationshantering 2.0
 - Mål förbättrad informationshantering och hanterar information som en tillgång. Strukturerad och semistrukturerad information beskrivs tillsammans genom en uppdaterad verksamhetsbeskrivning med en tydlig koppling mellan process, information och IT.
- Skala upp AI Kapaciteten i verksamheten
- Individuella kompetenslyft för att tillgodo se behoven





Prediction monitoring
(Prediktion vs Utfall)

Degradera till någon äldre version om senare versioner presterar systematiskt sämre.

$Kriterie(Old) > Kriterie(Current)$

Real-time predictions

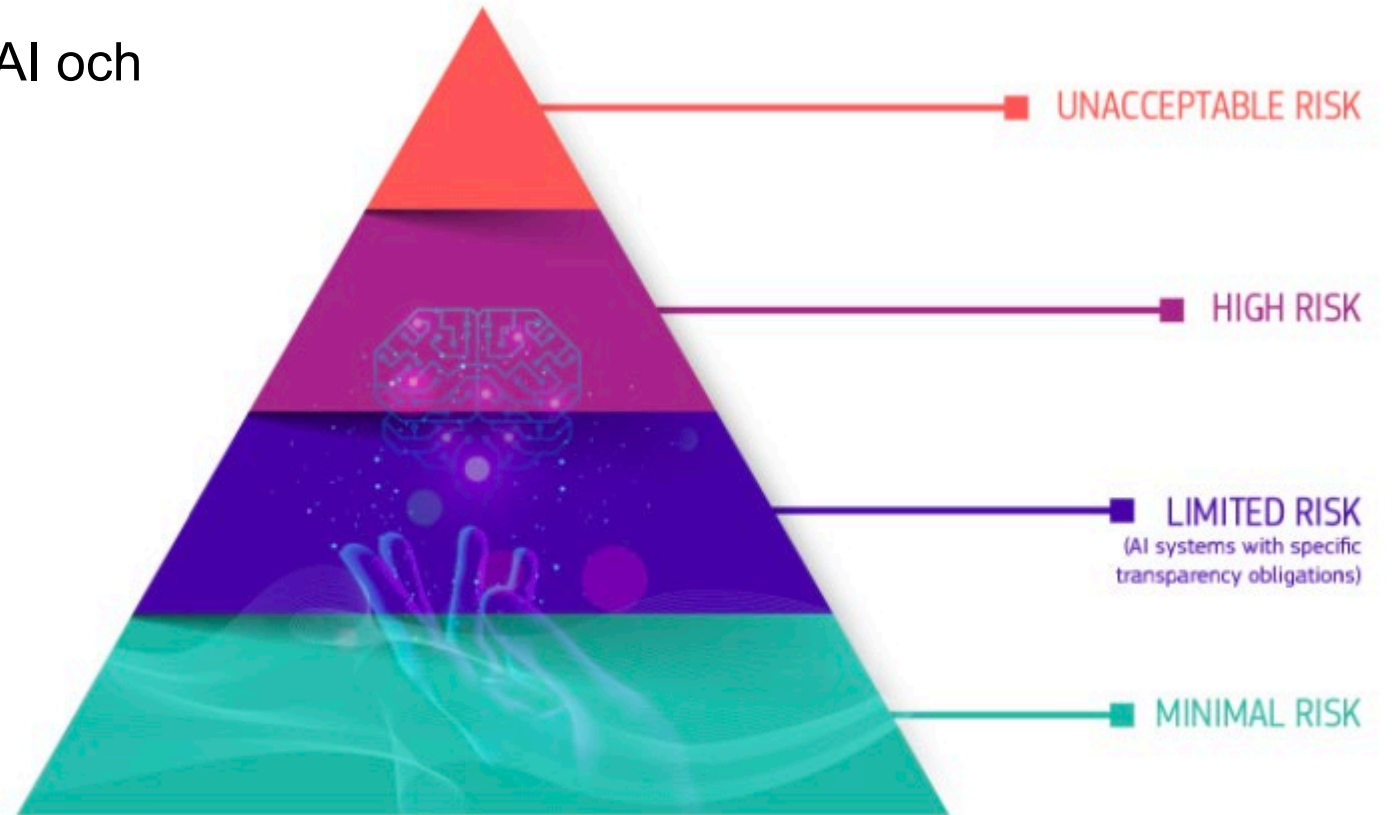
$Kriterie(RC) > Kriterie(Current)$

Sätt RC i produktion om värdeökningen kontra säkerheten i värdeökningen är högre än nuvarande model i produktion.

- Model parameters
- Data (train/test/eval)
- Learning parameters

Trafikverket startar Data och AI Etikråd

- Bereda Etiska frågeställningar inom AI och Data området
- Säkerställa att vi följer lagstiftning
 - Dataskyddsförordning, GDPR
 - The Artificial Intelligence Act





TRAFIKVERKET