

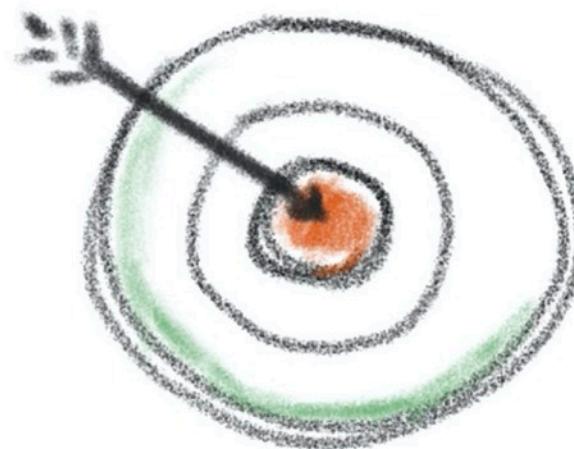


Machine Learning Use Cases with DDD and ML Design Canvas

Dr. Larysa Visengeriyeva

What kind of problems?

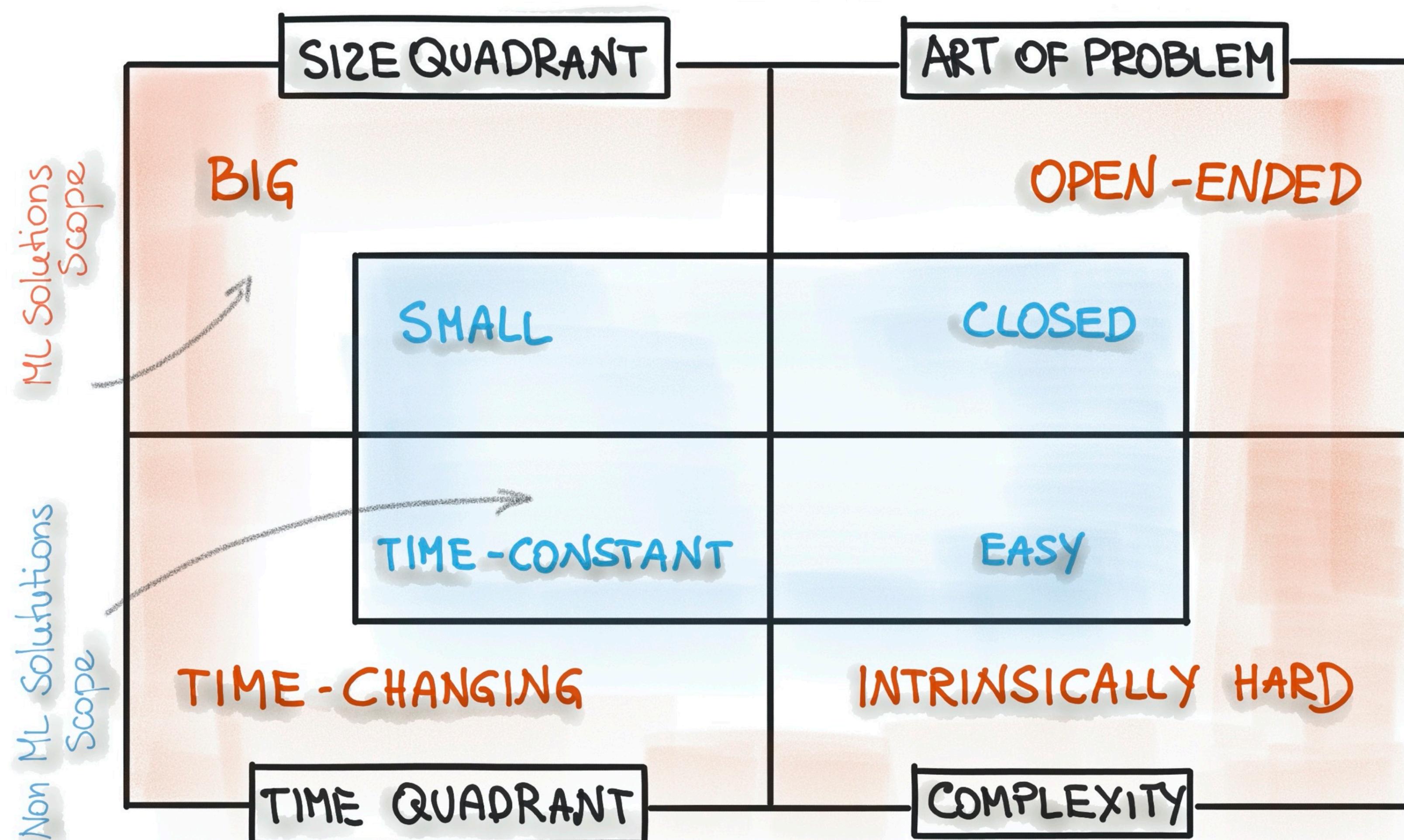
WHAT IS THE RIGHT
PROBLEM FOR AN
ML-SOLUTION ?



INNOQ

@visenger

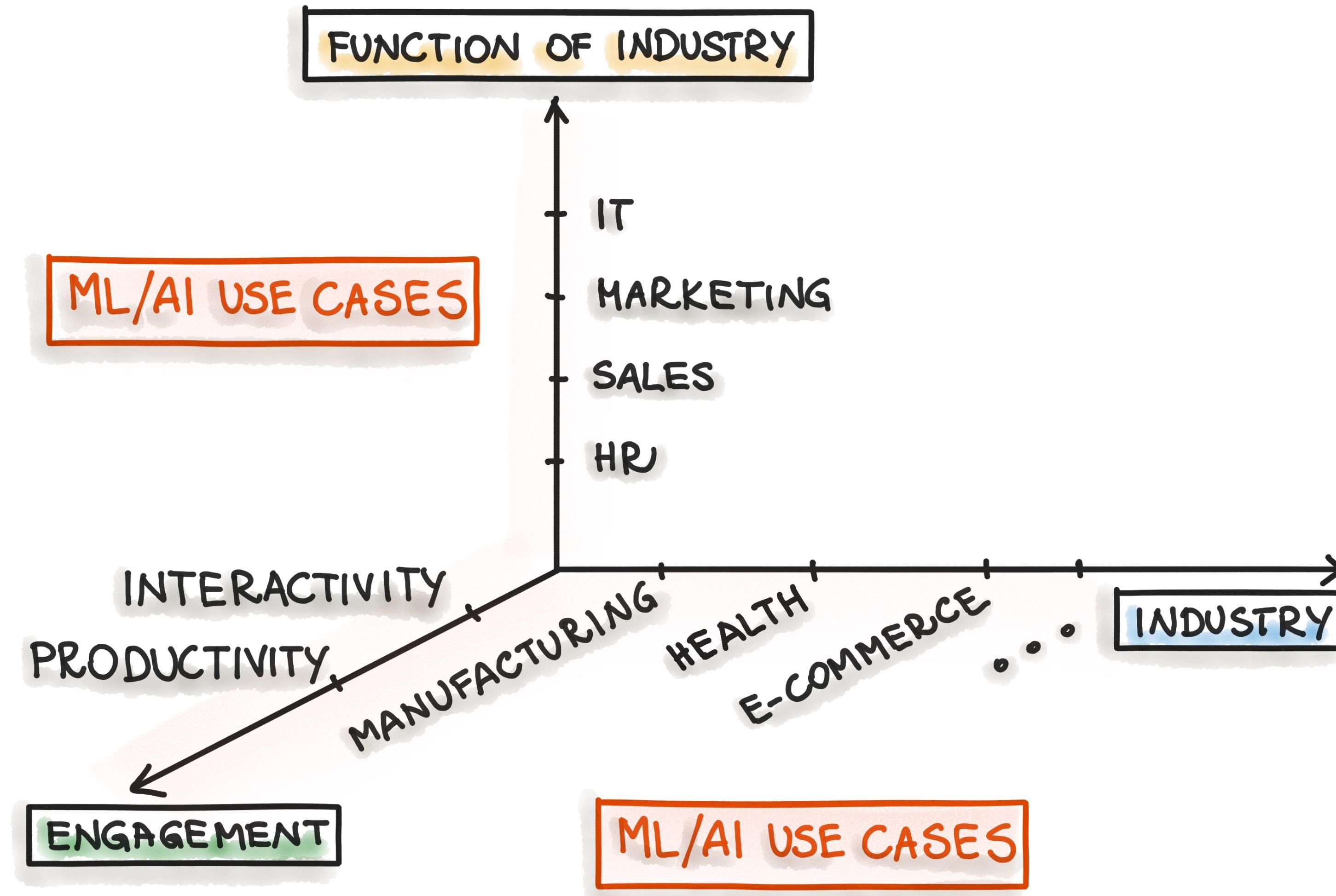
What kind of problems?



@visenger

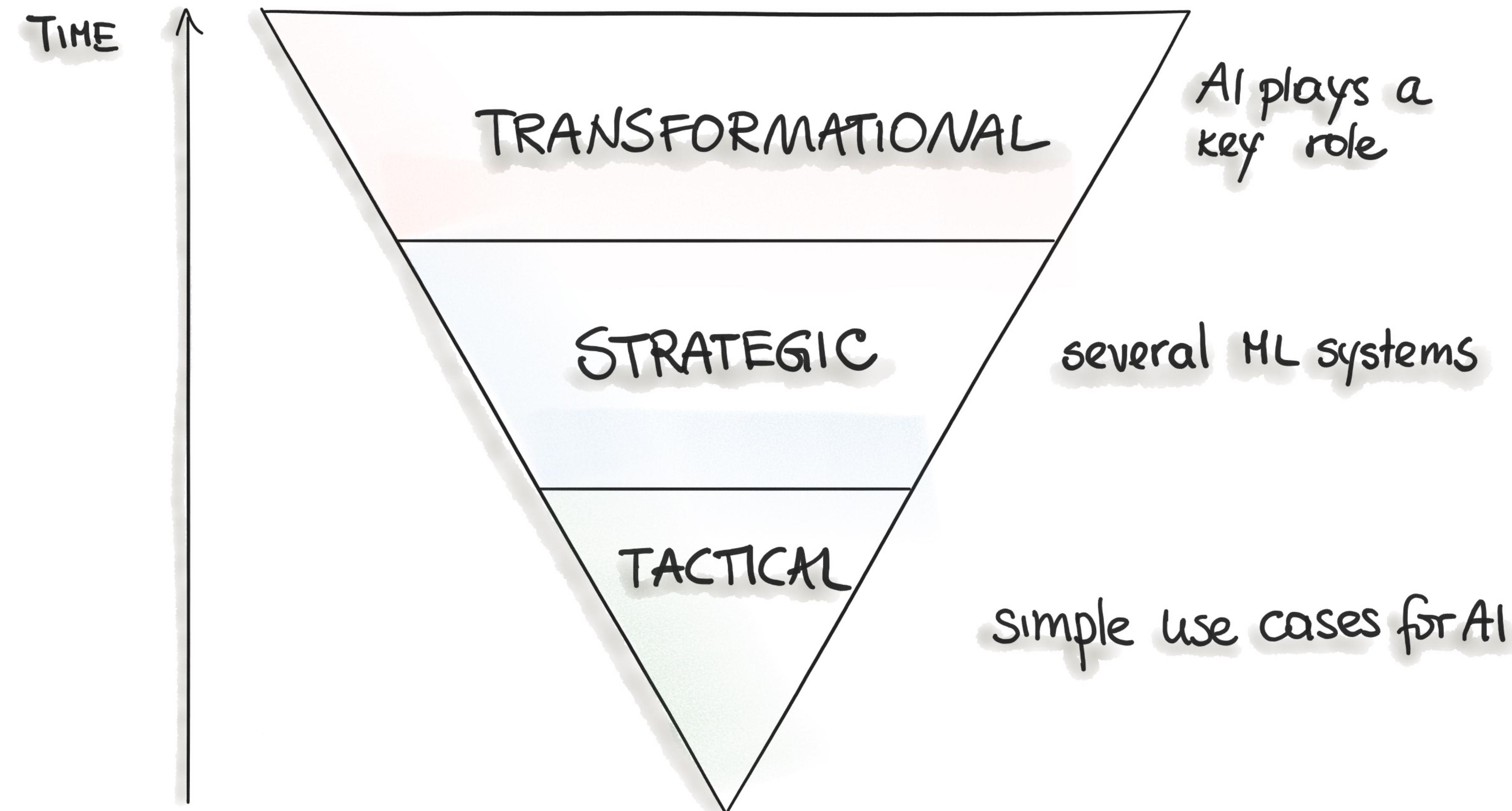
INNOQ

ML Use Cases Dimensions



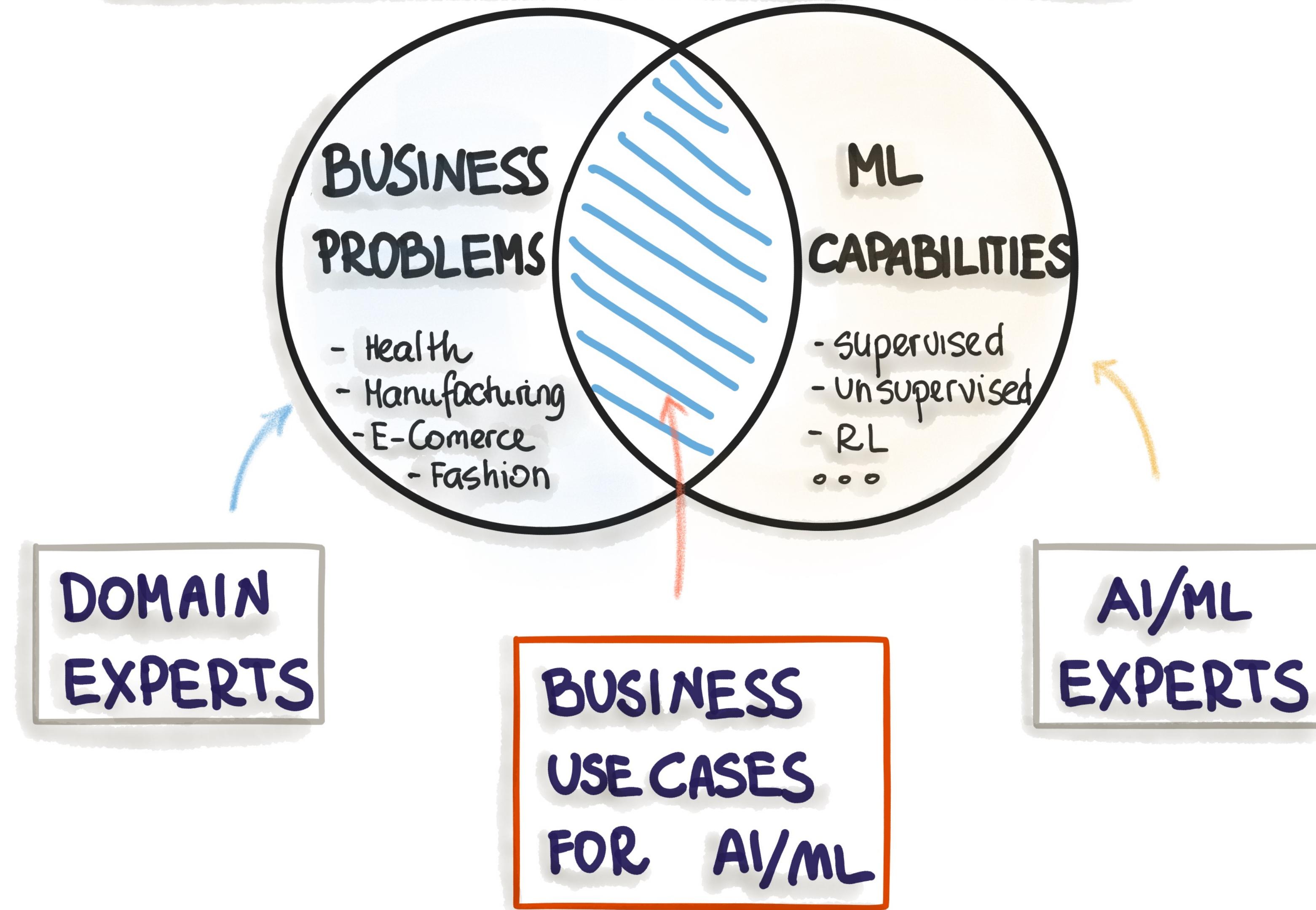
When?

AI READINESS



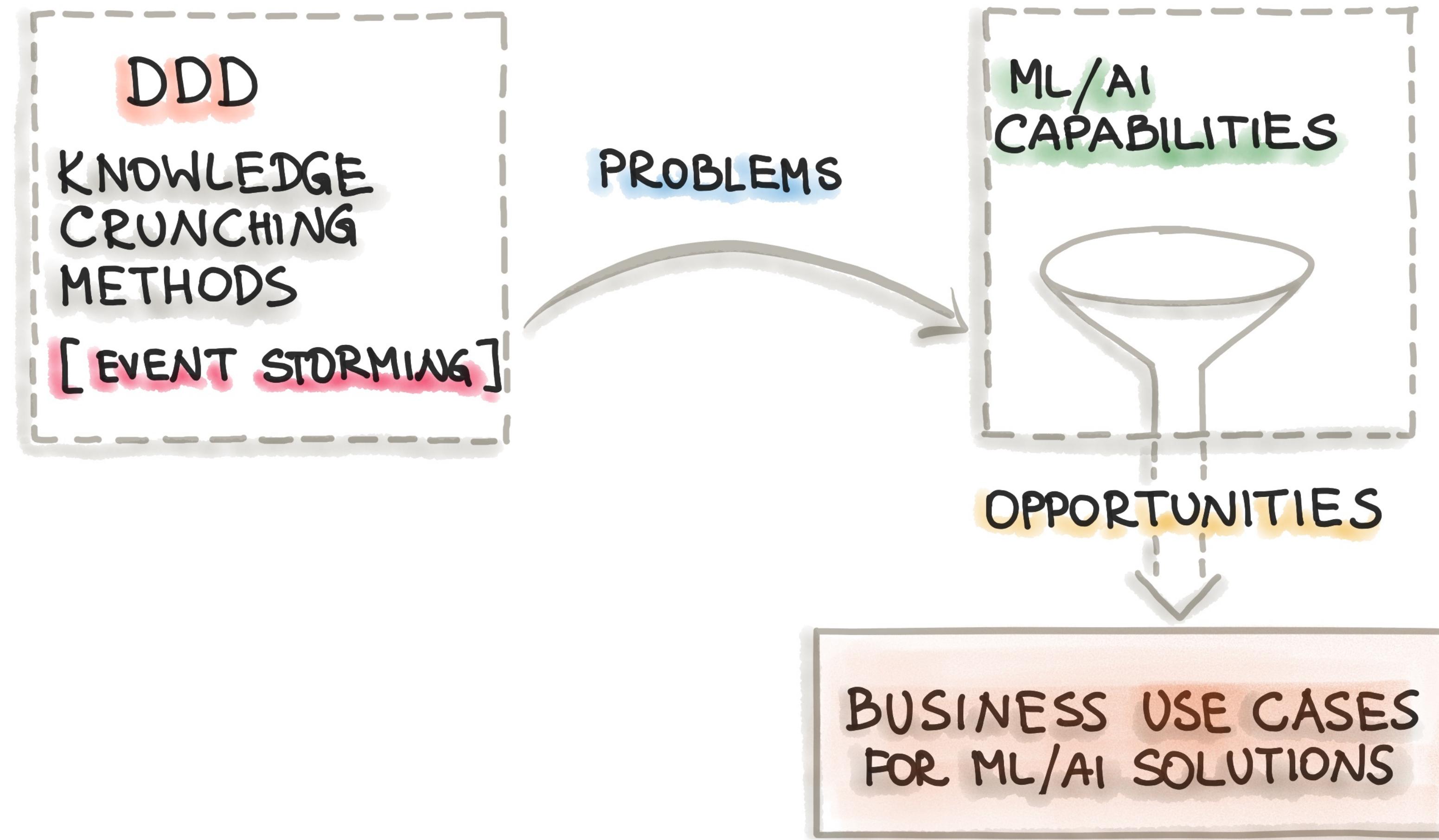
How?

MACHINE LEARNING USE CASES

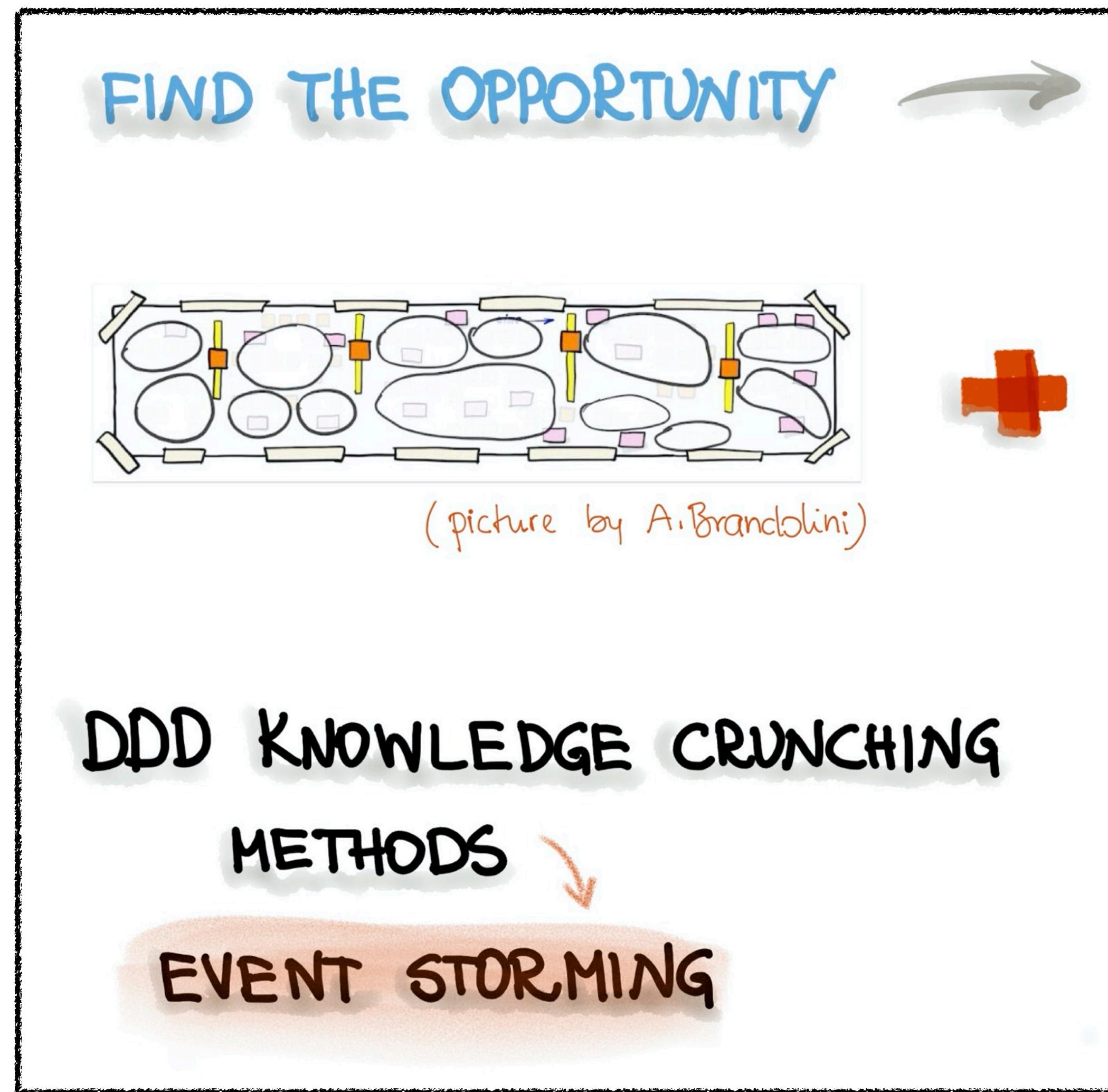


Knowledge Crunching

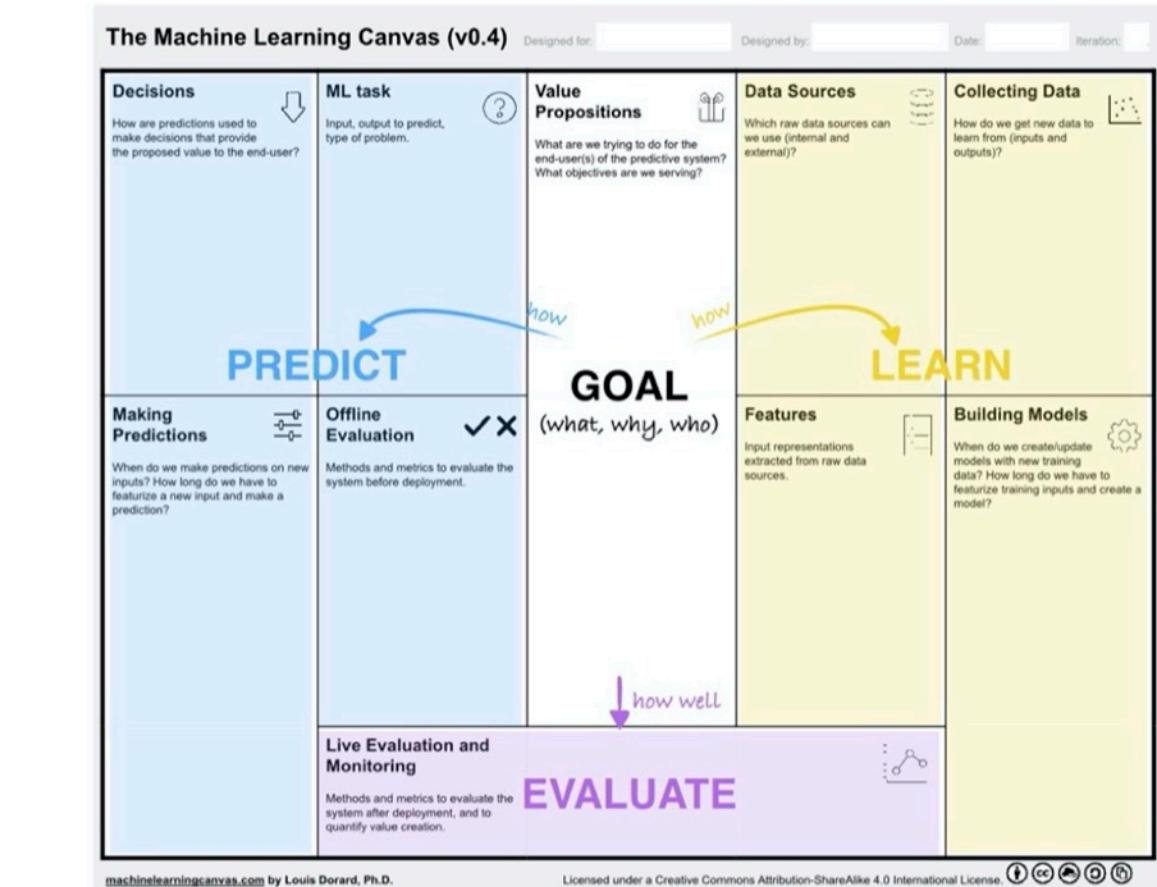
Event Storming For Finding AI Opportunities



Event Storming + ML Canvas



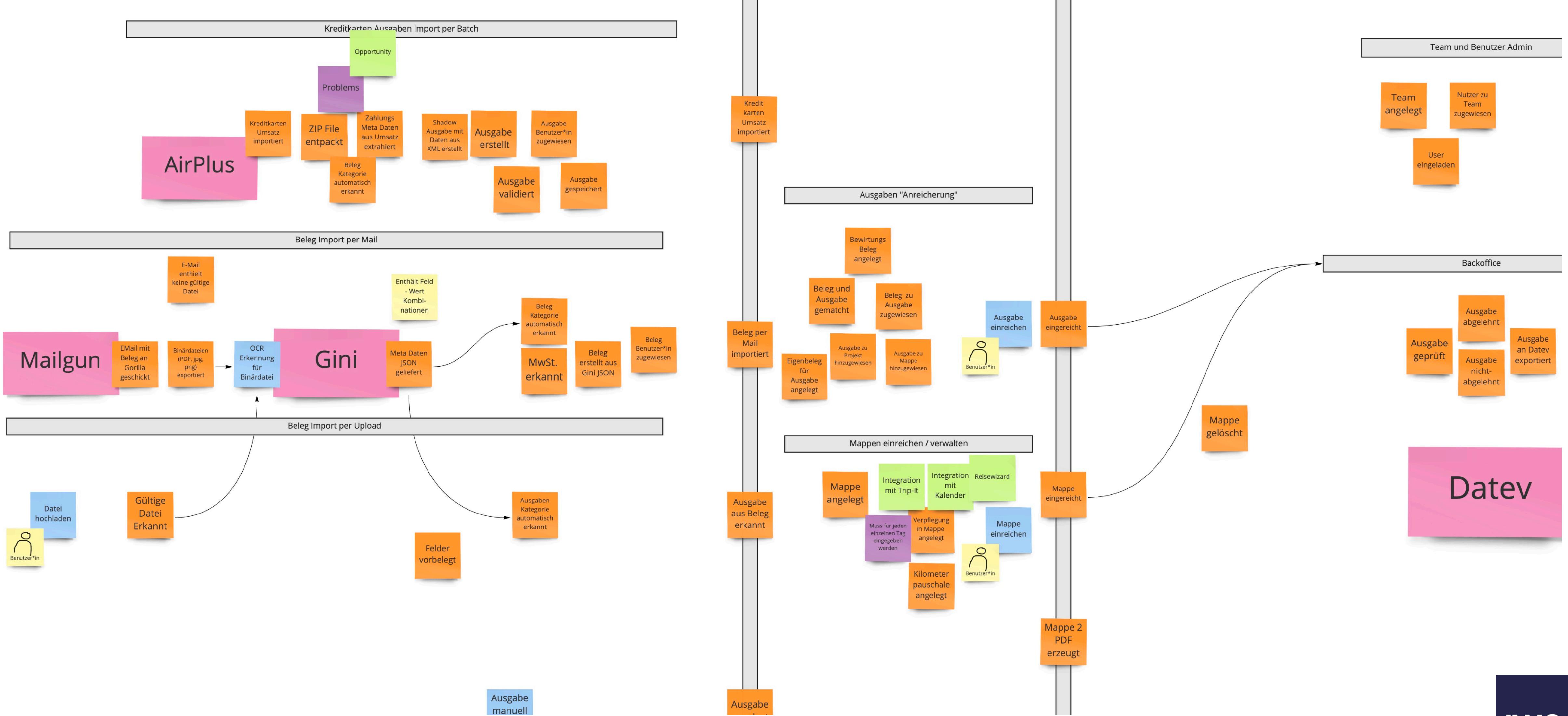
STRUCTURE PROJECT



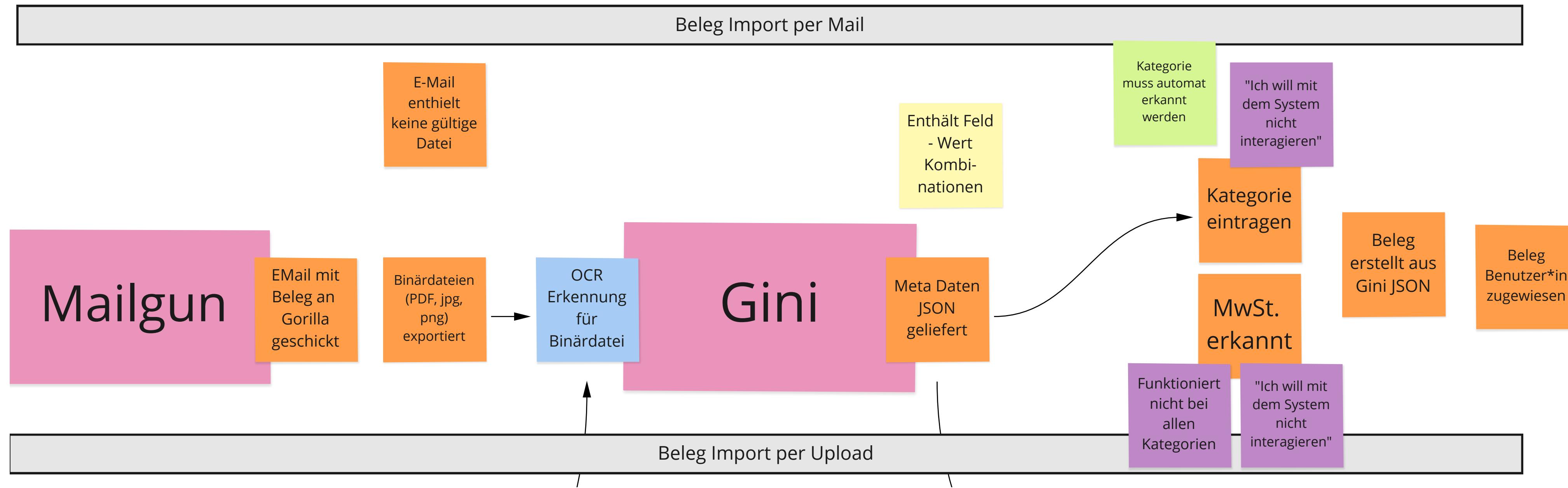
MACHINE
LEARNING
DESIGN
CANVAS

@visenger

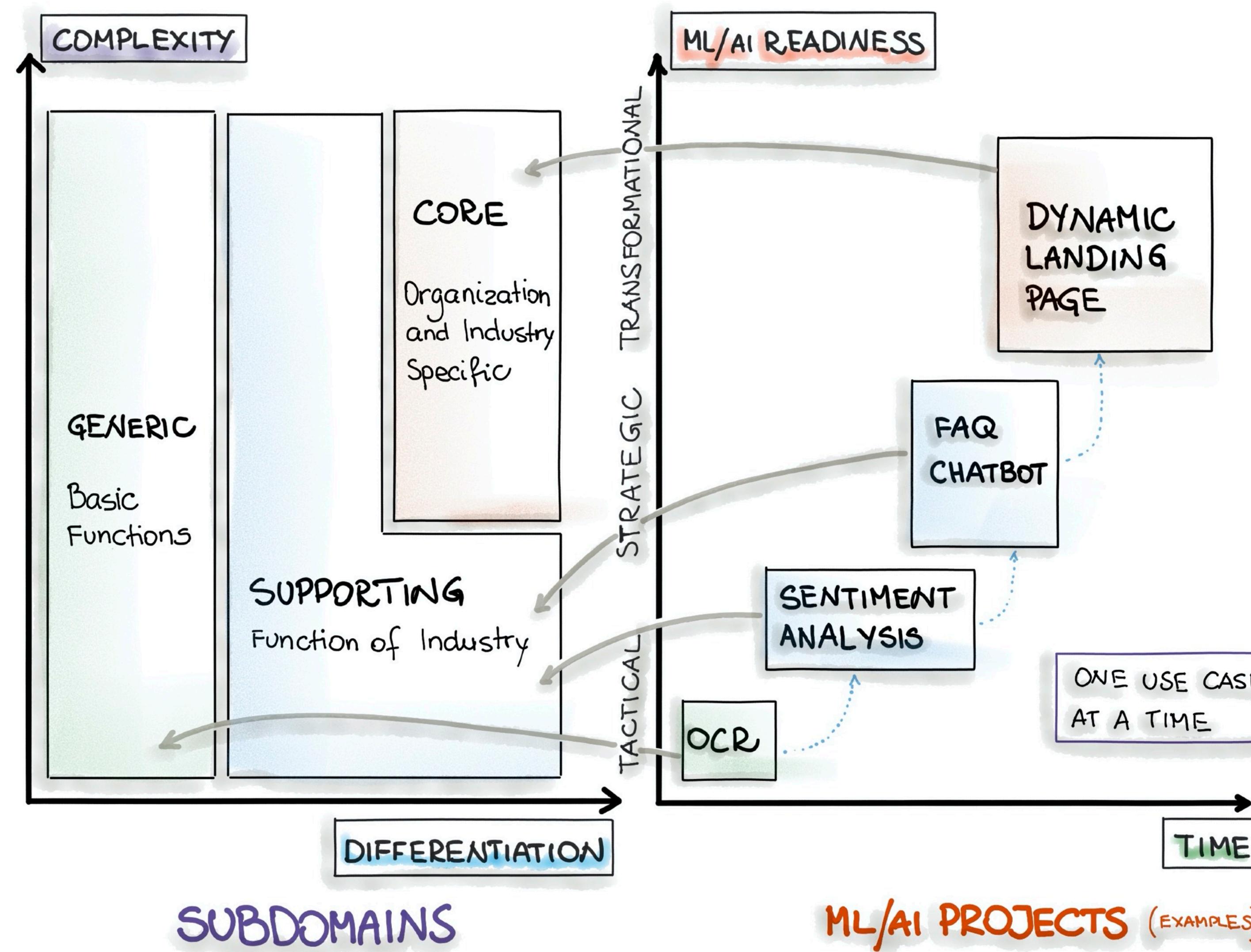
INNOQ



Finding Pain Points: Candidates for ML-Use Cases?

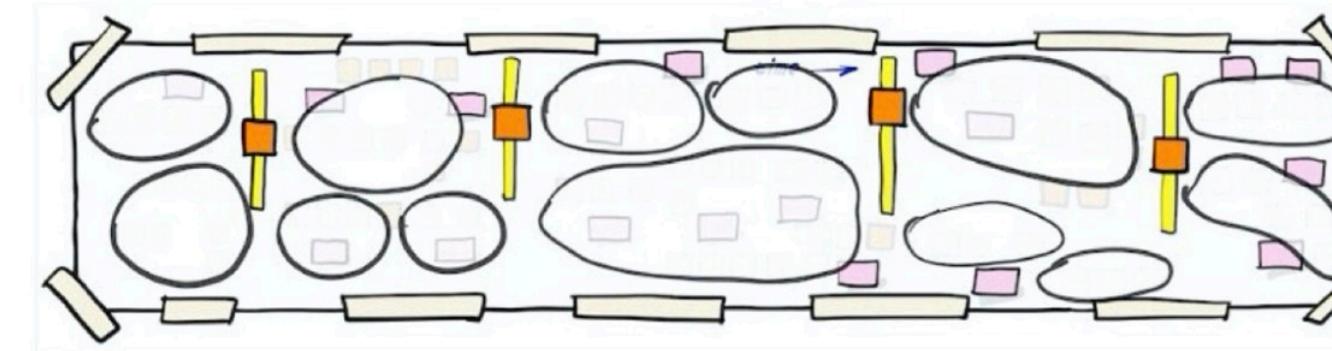


Using Domains For Prioritization



How to Structure ML Projects

FIND THE OPPORTUNITY



(picture by A. Brandolini)

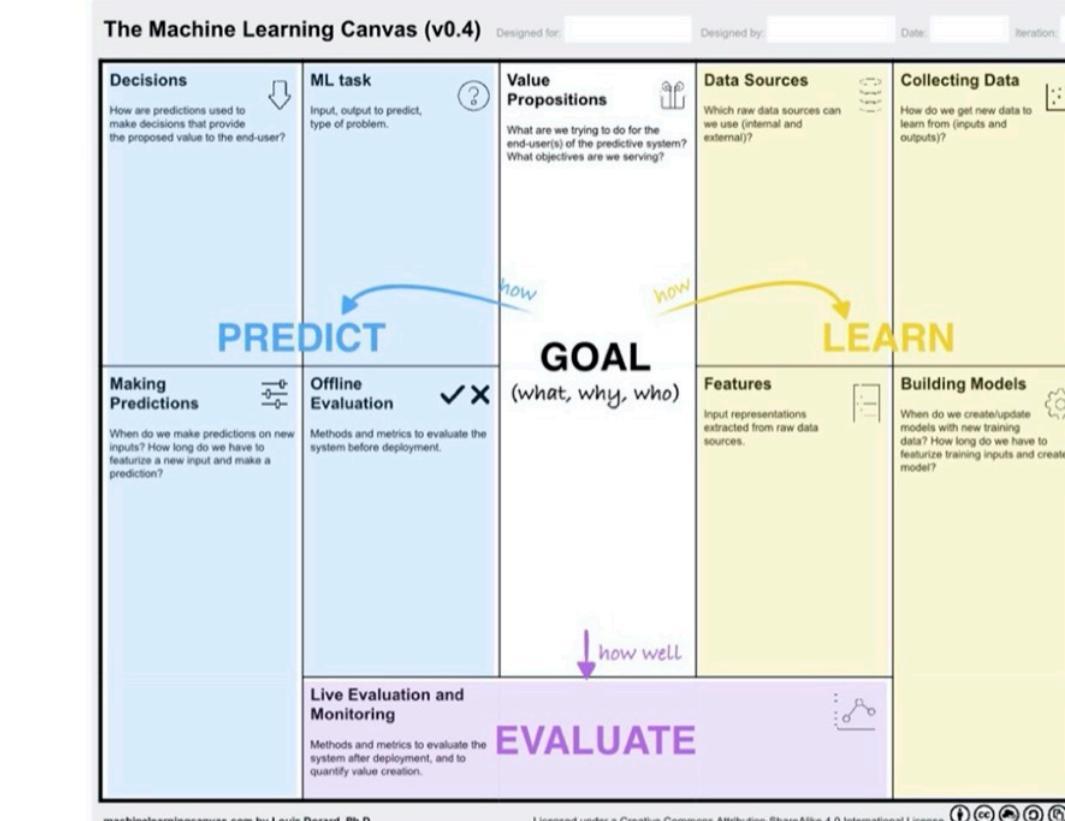
DDD KNOWLEDGE CRUNCHING

METHODS

EVENT STORMING



STRUCTURE PROJECT

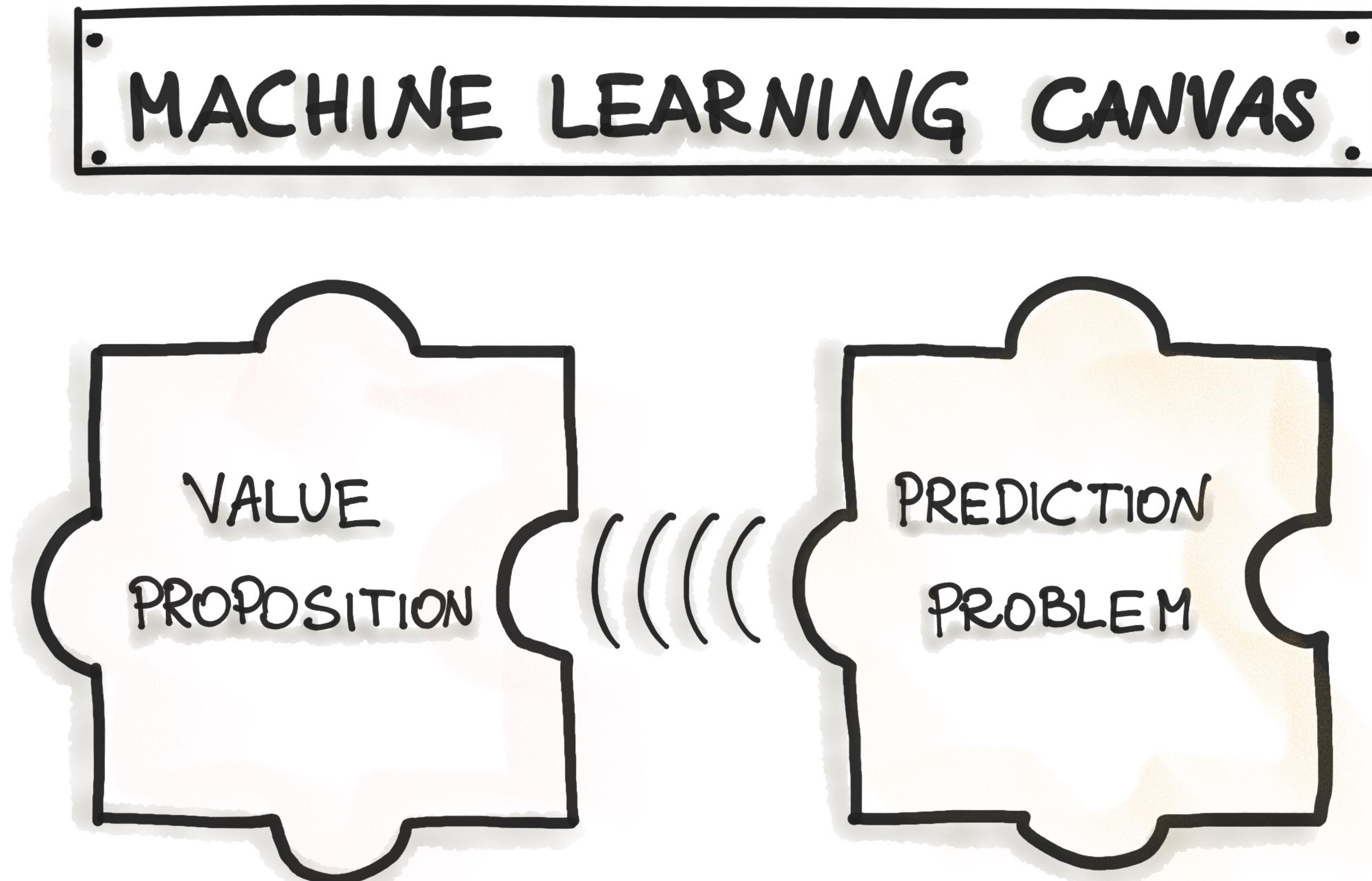


MACHINE
LEARNING
DESIGN
CANVAS

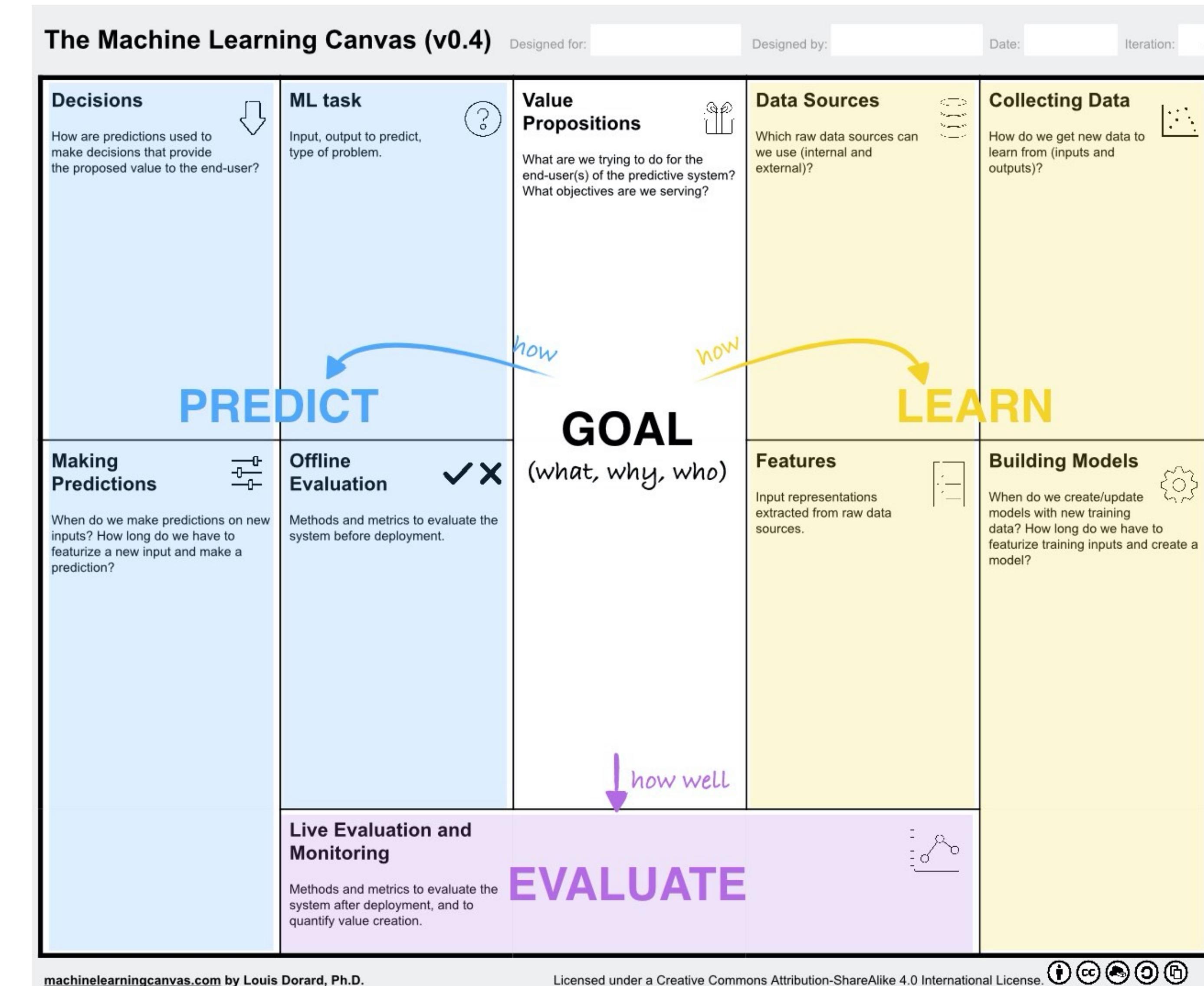
@visenger

INNOQ

Structuring ML Projects



Structuring ML Projects



Value Propositions

Value Propositions



1. What is the problem? What objective are we serving?
2. What are we trying to do for the end-user?
3. Why is it important?
4. Who is the end-user? Can we specify the persona?

Problem ""Ich will mit dem System nicht interagieren""

What we are trying to do: "Die RKG Kategorien automatisch erkennen"

Kategorie ist ein wichtiger Teil des Beleges.

Manuelle Arbeit

End User:
RKG - User

ML Task

ML Task



Input, output to predict, type of problem.
(Start with a heuristic)

Multiclass
Classification

Clustering
kNN

topic
modelling

Input:
OCR
Text

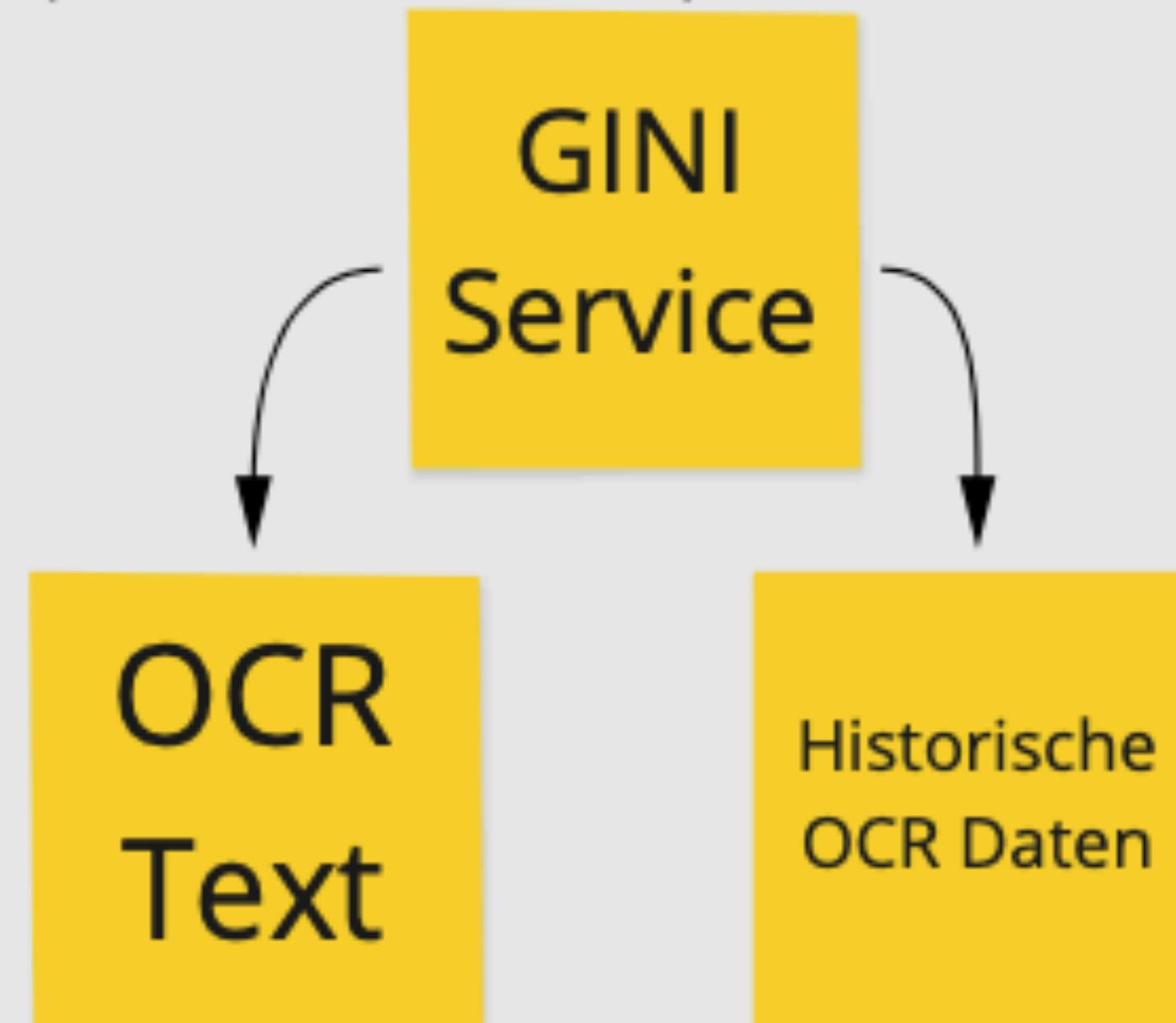
Output:
Eine
Kategorie

Data Sources

Data Sources



Which raw data sources can we use (internal and external)?



Features

Features

Input representations extracted
from raw data sources

TfidfVectorizer
auf dem OCR
Text



Offline Evaluation

Offline Evaluation



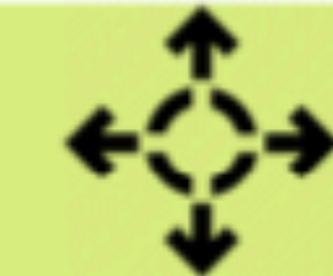
Methods and metrics to evaluate the system before deployment.

Accuracy

F1
Precision
Recall

Decisions

Decisions



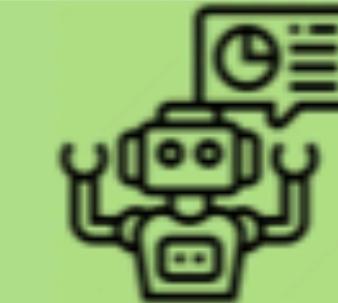
How are predictions used to make decisions that provide the proposed value to the end-user?

Kategorie wird automatisch im Formular eingetragen (incl. Mwst.)

Kategorie-erkennung muss vom User kontrolliert werden

Making Predictions

Making Predictions



When do we make predictions on new inputs?
How long do we have to featurize a new
input and make a prediction?

Jedesmal, wenn
ein Dokumen
im System
hochgelader
wurde.

Gini Service
stellt OCR Text
von dem neuen
Dokument zur
Verfügung

TfidfVectorizer
für Feature
Extraction

Klassifizierung
on-the-fly

Collecting Data

Collecting Data



How do we get new data to learn from
(inputs and outputs)?

3-
Monate
Batch

Vom Backoffice
gelabelt/geprüft
(Human)

Gini/RKG-
Server

Building Models

Building Models



When do we create/update models with new training data? How long do we have to featurize training inputs and create a model?

Alle 3 Monate

Jeden Monat:
Chron Job für
Batch holen
und Model re-
training

S3 für
Model
Repository

Heroku
Instanz

Weil wenig Daten
und einfaches
Model: -> Model
re-training dauert
ca. 1-2 Min

Live Evaluation and Monitoring

Live Evaluation and Monitoring

Methods and metrics to evaluate the system after deployment, and to quantify value creation.

Weniger
Ausfüllen
bei
Formular

Kürzere Zeiten bei
RK-
Belegeabrechnung

MLOps @ INNOQ

ML Ops

ml-ops.org

...

∞ MLOps

Machine Learning Operations

With Machine Learning Model Operationalization Management (MLOps), we want to provide an end-to-end machine learning development process to design, build and manage reproducible, testable, and evolvable ML-powered software.

```
graph TD; Design((Design)) --- ModelDevelopment((Model Development)); ModelDevelopment --- Operations((Operations)); Operations --- Design;
```