The Future
of Model
Management

GHENT
UNIVERSITY

Prof. Dr. Frank De Jonghe

Themes

- The dot on the horizon, a.k.a. vision
- Two examples where classical MRM is challenged by AI
- The EU AI Act points a/the way
- There is no responsibility without causality
- MLOps thinking adds value

Dot on the horizon

- Automation & Standardisation for more speed and compliance by design

- Alignment between models increasing, with one's output feeding another

- The insight is in the data, more than in the model. Beyond a ML/AI capability, there is need for data hunting/curation/anonymisation skills

- Data can come from different sources (partners, other banks, ...). Data sharing infrastructure (data fabric)?

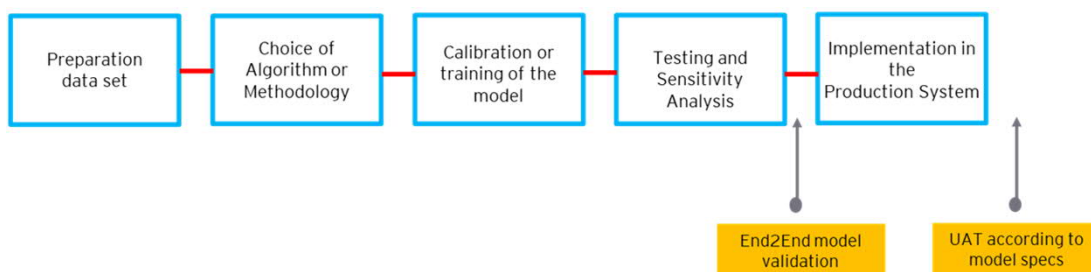- Regulators issuing guidance (EU AI Act, NIST 1270 on bias, EBA on BD&AA, ...)
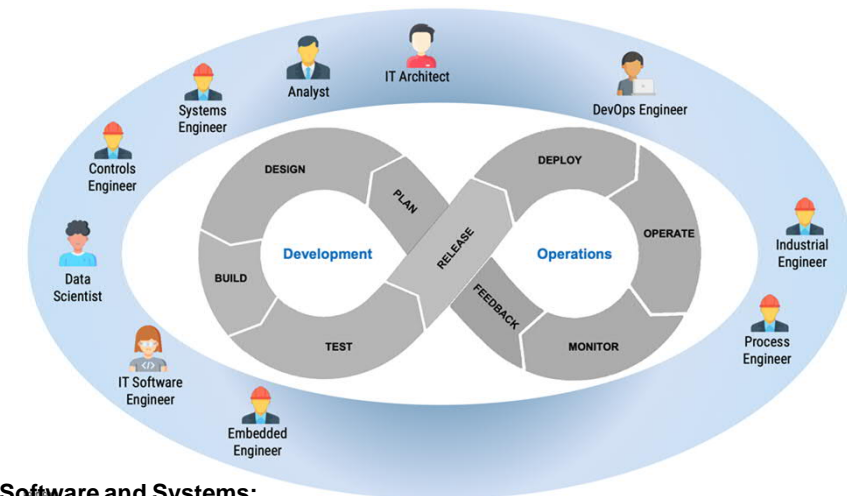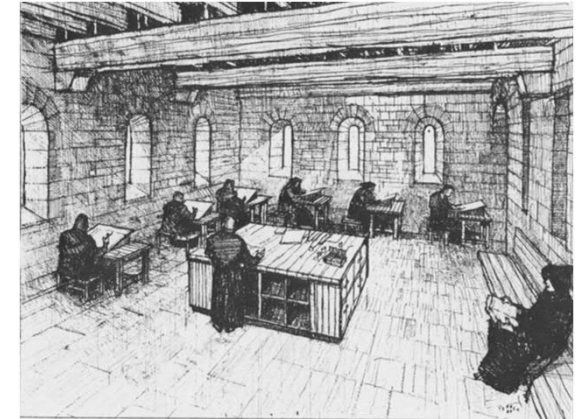
DevOps for Software and Systems:
Putting Algorithms and Models in Operation
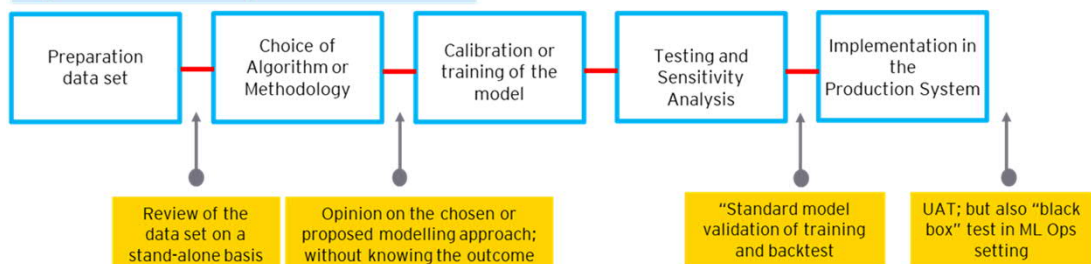*Peter Brady Martin Becker ; Matlab Expo 2021*

MRM is more than validation

- When it comes to AI/ML, the approach to validation still works, but needs to address the extra challenges (frequency, amount of data, open source, transparency of algorithm, ...)
- In function of risk triage, continuous monitoring of correct functioning may be adequate as risk mitigation
- Always focus on the domain of applicability
- Agile validation: effective challenge vs. independence

Standard view on independent validation

| Preparation data set | Choice of Algorithm or Methodology | Calibration or training of the model | Testing and Sensitivity Analysis | Implementation in the Production System |

End2End model validation

UAT according to model specs

Agile alternative algorithm validation

| Preparation data set | Choice of Algorithm or Methodology | Calibration or training of the model | Testing and Sensitivity Analysis | Implementation in the Production System |

Review of the data set on a stand-alone basis

Opinion on the chosen or proposed modelling approach; without knowing the outcome

"Standard model validation of training and backtest

UAT; but also "black box" test in ML Ops setting
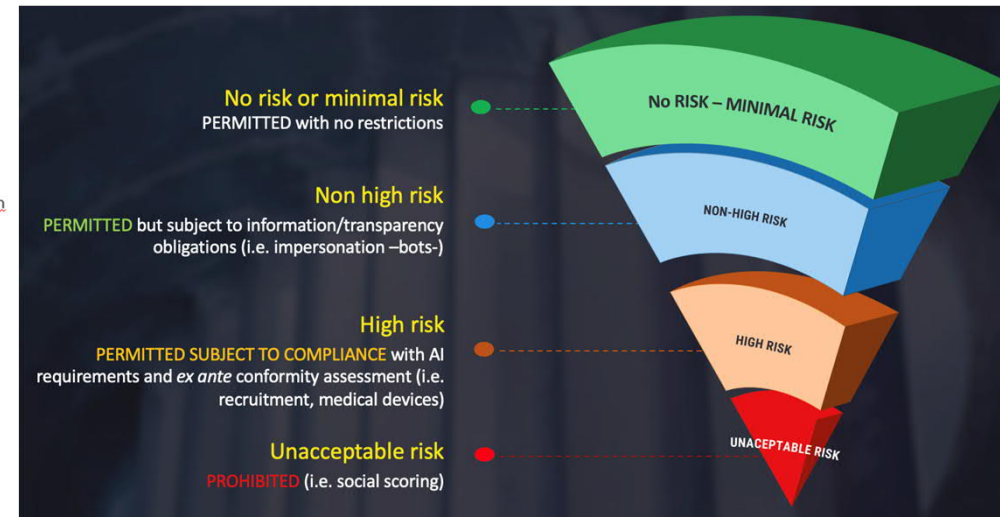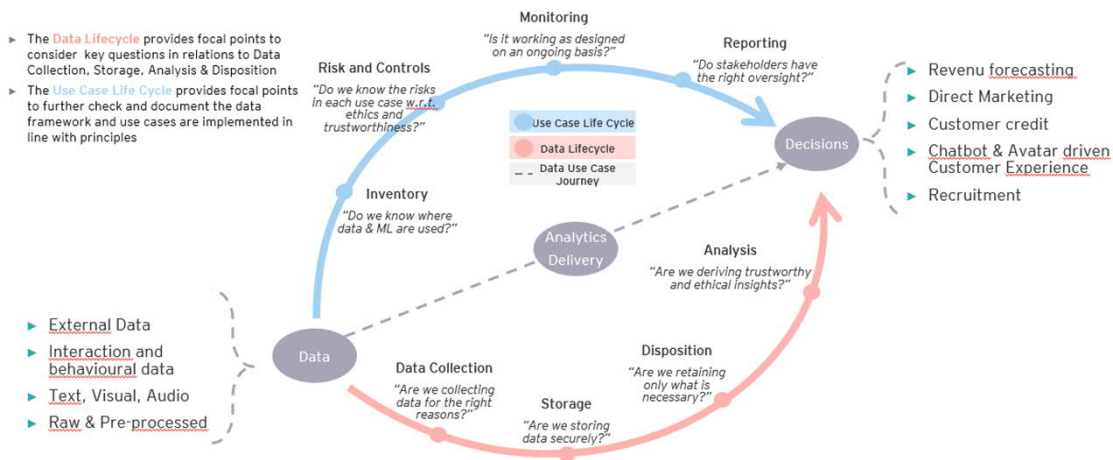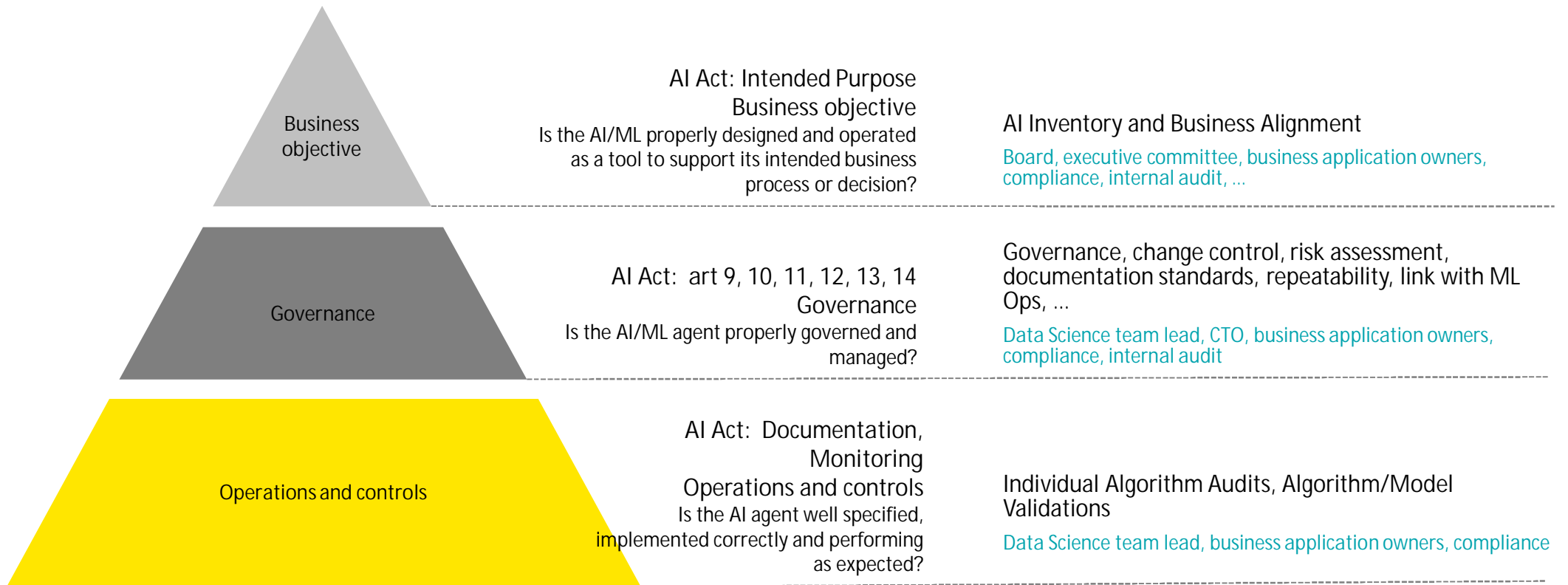
https://www.linkedin.com/pulse/agile-model-validation-threat-opportunity-frank-de-jonghe/?trackingId=hc9WuGDCqKKyDQ4fNjr%2FsA%3D%3D
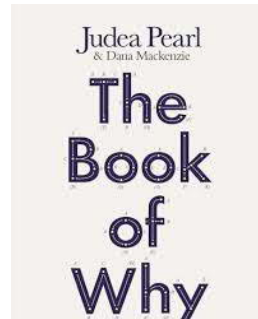
The EU AI Act – Need for different risk lenses

Monitoring
"Is it working as designed on an ongoing basis?"

Reporting
"Do stakeholders have the right oversight?"

Risk and Controls
"Do we know the risks in each use case w.r.t. ethics and trustworthiness?"

► The Data Lifecycle provides focal points to consider key questions in relations to Data Collection, Storage, Analysis & Disposition
► The Use Case Life Cycle provides focal points to further check and document the data framework and use cases are implemented in line with principles

Use Case Life Cycle
Data Lifecycle
Data Use Case Journey

Decisions

► Revenu forecasting
► Direct Marketing
► Customer credit
► Chatbot & Avatar driven Customer Experience
► Recruitment

Inventory
"Do we know where data & ML are used?"

Analytics Delivery

Analysis
"Are we deriving trustworthy and ethical insights?"

Disposition
"Are we retaining only what is necessary?"

Data

Data Collection
"Are we collecting data for the right reasons?"

Storage
"Are we storing data securely?"

► External Data
► Interaction and behavioural data
► Text, Visual, Audio
► Raw & Pre-processed

No risk or minimal risk
PERMITTED with no restrictions

Non high risk
PERMITTED but subject to information/transparency obligations (i.e. impersonation –bots-)

High risk
PERMITTED SUBJECT TO COMPLIANCE with AI requirements and *ex ante* conformity assessment (i.e. recruitment, medical devices)

Unacceptable risk
PROHIBITED (i.e. social scoring)

No RISK – MINIMAL RISK

NON-HIGH RISK

HIGH RISK

UNACCEPTABLE RISK

► The AIA extends the GDPR initiated governance on data, to the use of this data in an AI/ML context. It is use cases that are regulated.

► All actors in the value chain of an AI use case, have their responsibility, including public sector.

► The approach to the AI applications is risk based. However, the risk appetite is skewed towards human rights and human agency. From an internal company perspective, other use cases may be considered high risk too, requiring similar standards for development and putting in production.

The EU AI Act – Not rocket science, but still onerous

Business objective

AI Act: Intended Purpose
Business objective
Is the AI/ML properly designed and operated as a tool to support its intended business process or decision?

AI Inventory and Business Alignment

Board, executive committee, business application owners, compliance, internal audit, ...

Governance

AI Act: art 9, 10, 11, 12, 13, 14
Governance
Is the AI/ML agent properly governed and managed?

Governance, change control, risk assessment, documentation standards, repeatability, link with ML Ops, ...

Data Science team lead, CTO, business application owners, compliance, internal audit

Operations and controls

AI Act: Documentation, Monitoring
Operations and controls
Is the AI agent well specified, implemented correctly and performing as expected?

Individual Algorithm Audits, Algorithm/Model Validations

Data Science team lead, business application owners, compliance

There is no responsibility without causality



- All modelling starts from the business, and relies on the business knowledge

- Fairness/bias are quickly emerging as modelling topics, but

- ownership/governance questions make progress (too) slow

- Bridge between the ethical principles and the model what-to-do not always clear in literature

- Different steps in the modelling chain where one can intervene (training data, algorithm, post-calibration, …)

You might find that all variables are correlated …
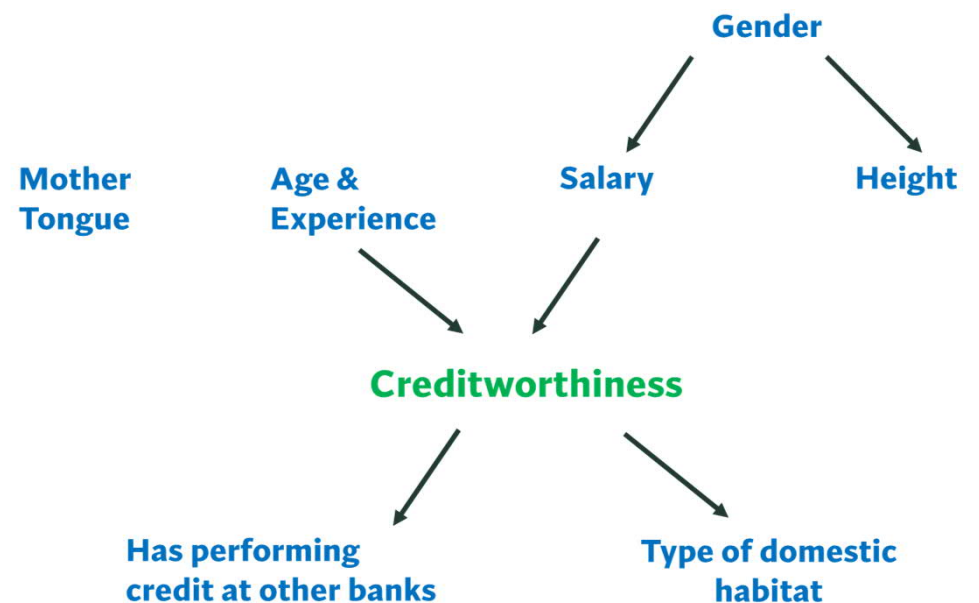
ML Ops can teach a lot to classical MRM ...
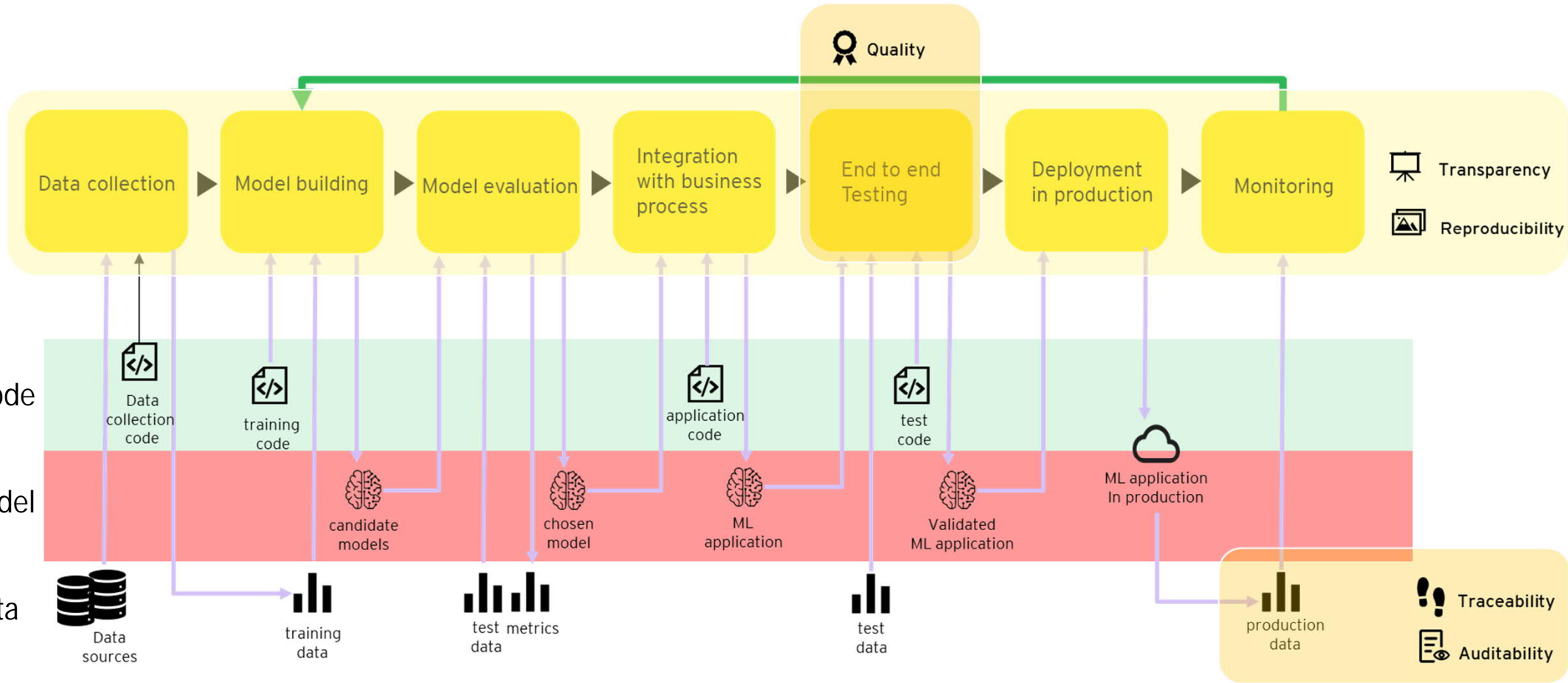
Example: EBA Guidelines on PD/LGD modeling benefit from translation in Data Science jargon

Data Drift

If a modelled relation fails to hold, the model loses its ability to rank risks and we encounter concept drift

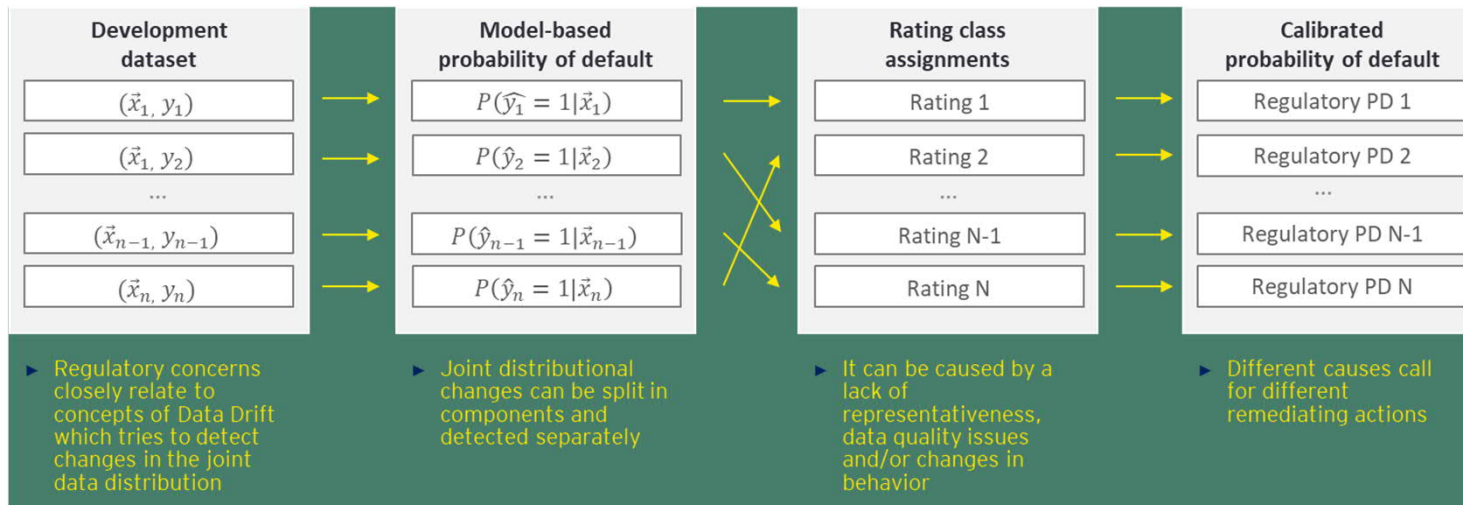$$P_A(y \mid \vec{x}) \neq P_B(y \mid \vec{x})$$

$$P(\vec{x}, y) = P(y \mid \vec{x})\, P(\vec{x})$$

$$P_A(y; \vec{x}) \neq P_B(y; \vec{x})$$

Training on a non-representative dataset causes data drift

If the active portfolio has different characteristics compared to the training data, we can get covariate drift

$$P_A(\vec{x}) \neq P_B(\vec{x})$$

Development dataset	Model-based probability of default	Rating class assignments	Calibrated probability of default
$(\vec{x}_1, y_1)$	$P(\widehat{y_1} = 1 \mid \vec{x}_1)$	Rating 1	Regulatory PD 1
$(\vec{x}_1, y_2)$	$P(\hat{y}_2 = 1 \mid \vec{x}_2)$	Rating 2	Regulatory PD 2
...	...	...	...
$(\vec{x}_{n-1}, y_{n-1})$	$P(\hat{y}_{n-1} = 1 \mid \vec{x}_{n-1})$	Rating N-1	Regulatory PD N-1
$(\vec{x}_n, y_n)$	$P(\hat{y}_n = 1 \mid \vec{x}_n)$	Rating N	Regulatory PD N

▶ Regulatory concerns closely relate to concepts of Data Drift which tries to detect changes in the joint data distribution

▶ Joint distributional changes can be split in components and detected separately

▶ It can be caused by a lack of representativeness, data quality issues and/or changes in behavior

▶ Different causes call for different remediating actions

$$P(\hat{y}) = \sum_{\vec{x}} P(\hat{y}; \vec{x}) = \sum_{\vec{x}} P(\hat{y} \mid \vec{x}) P(\vec{x})$$

If the active portfolio has different characteristics in terms of overall default rate, we encounter target drift

$$P_A(y) \neq P_B(y)$$

Wrap-up

- Multi-year transformation journey that
- Needs to build on a robust technology for repeatability
- Requires different skills to work together
- Will be shaped by evolving regulatory expectations

Thank You

frank.dejonghe@ugent.be