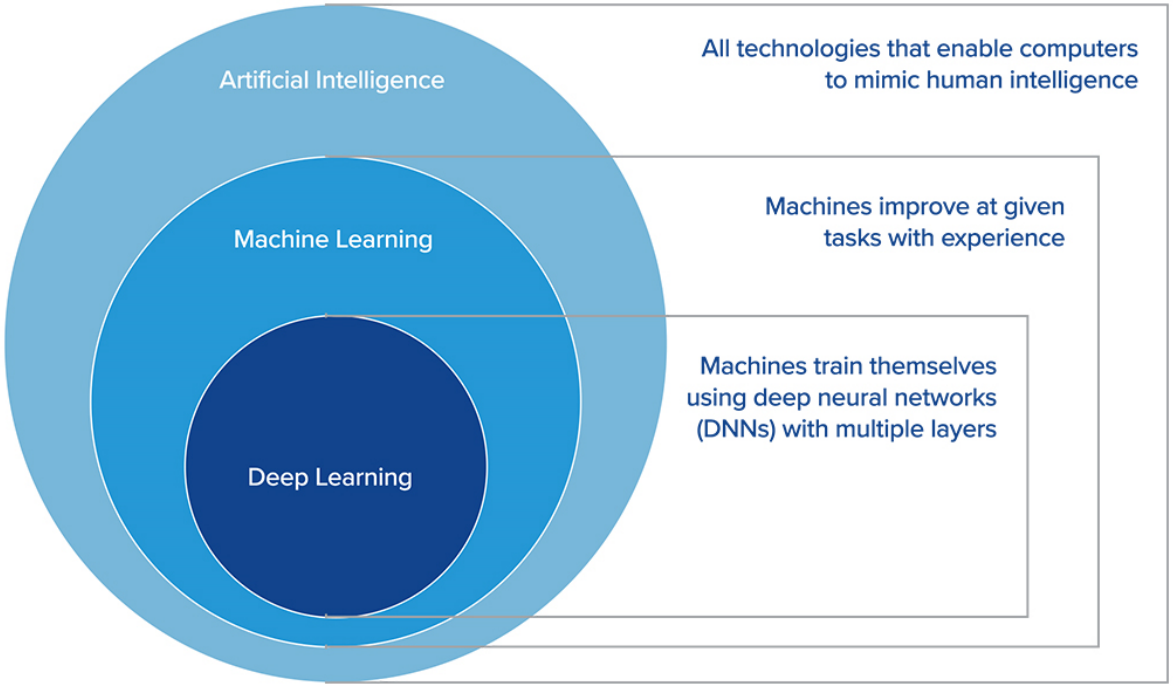
AI beyond AI

Guido Van Humbeeck – VDAB
guido.vanhumbeecl@vdab.be
June 2020

Themes

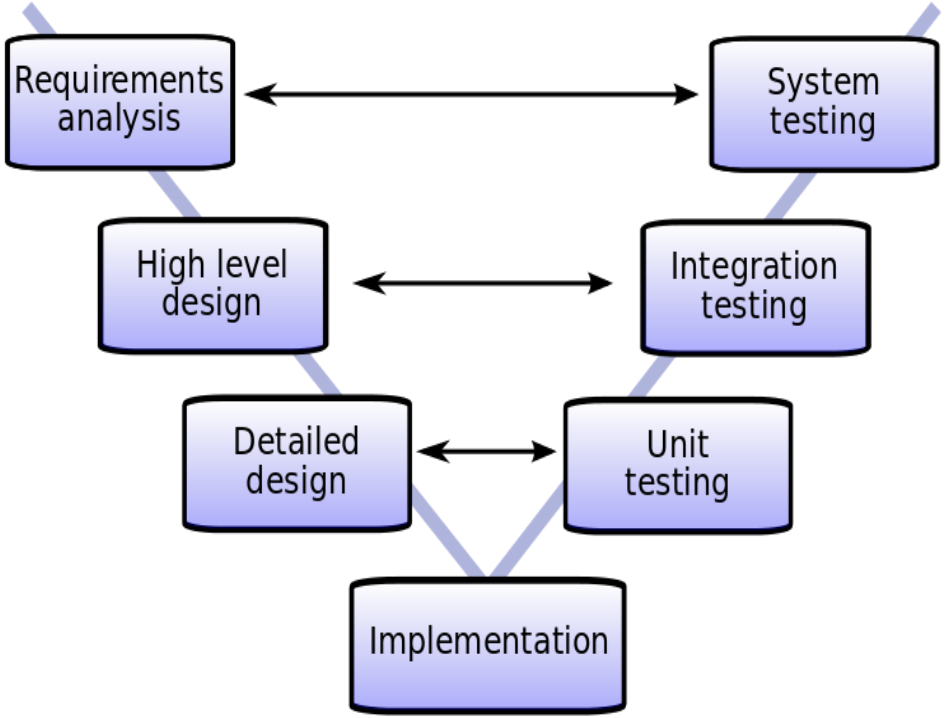
- Positioning AI in de IT landscape
- Building AI artefacts
- Operationalize AI Artefacts
- MLOps

Types of Artificial Intelligence (AI)

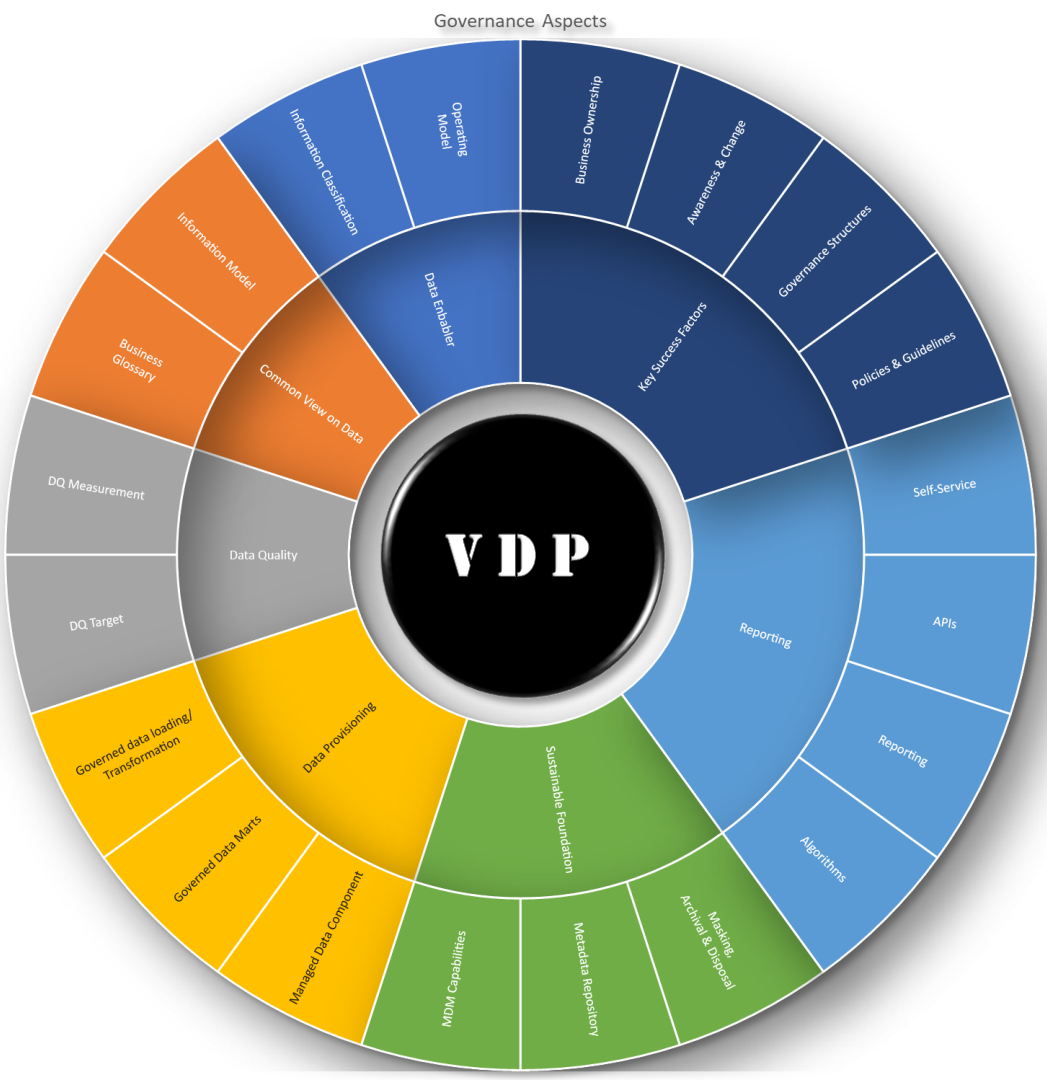


VDAB

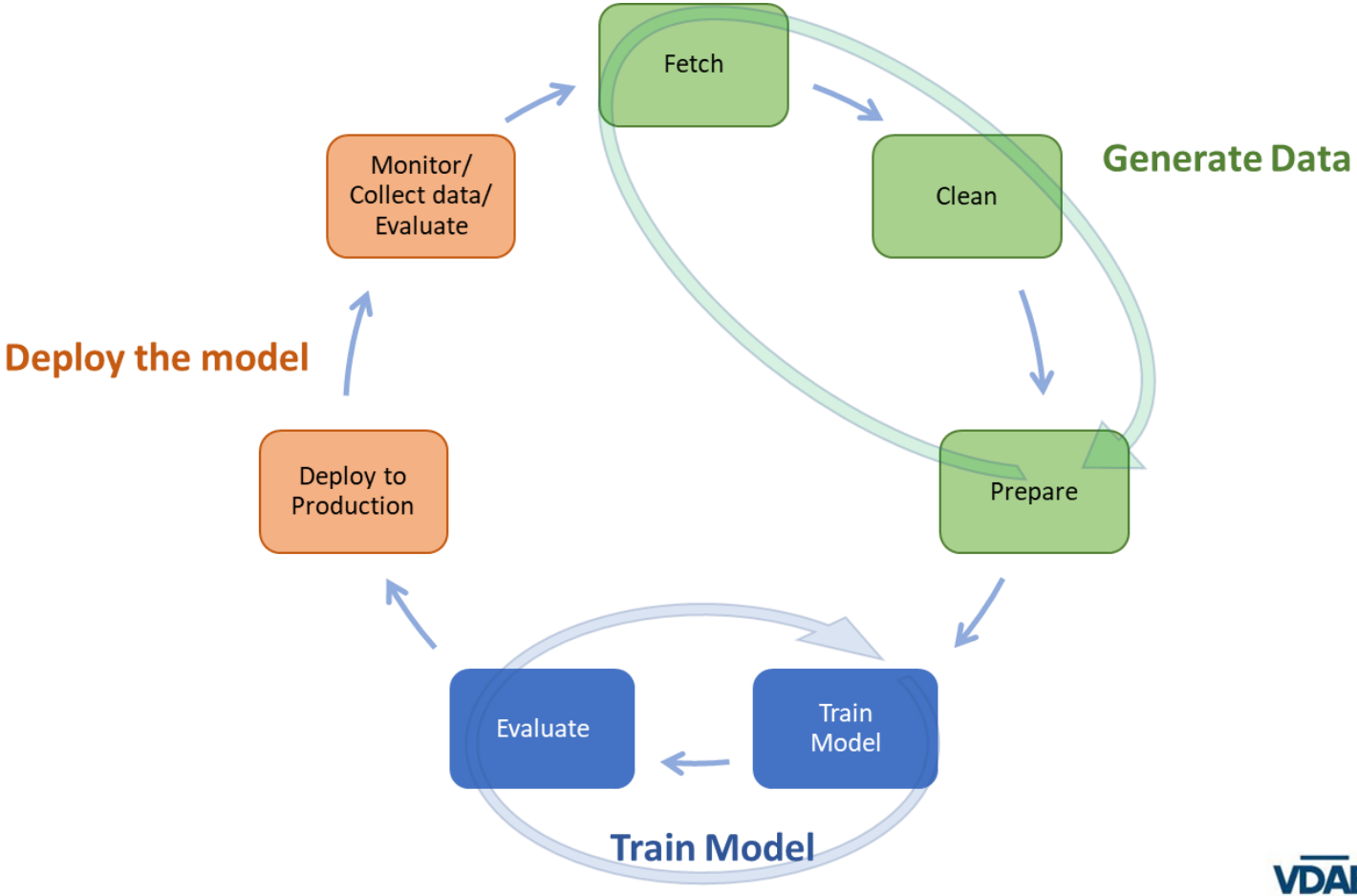
Application Development



Data Management

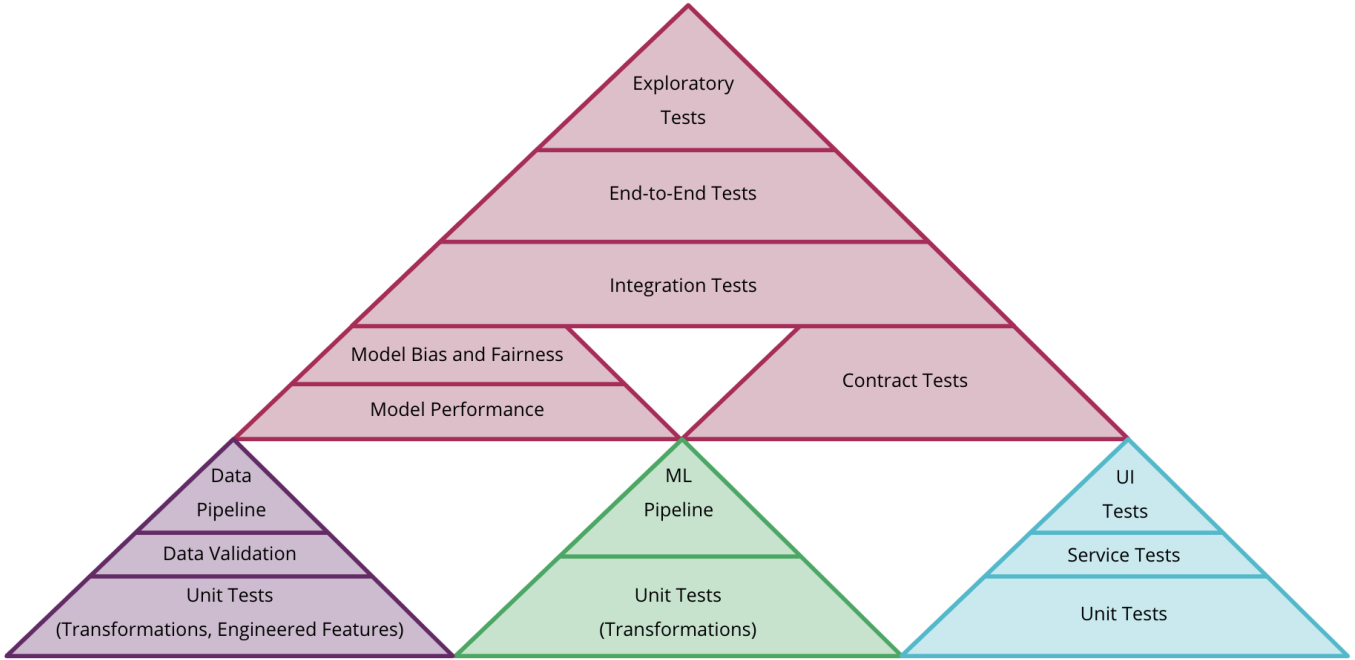


AI Artefact development



VDAB

Testing and Quality in Machine Learning



https://martinfowler.com/articles/cd4ml.html

VDAB

Model Monitoring *(Martin Fowler)*

Model inputs: ***what data is being fed to the models***, giving visibility into any training-serving skew. Model outputs: ***what predictions and recommendations*** *are the models making* ***from these inputs***, to understand how the model is performing with real data.

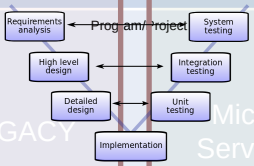
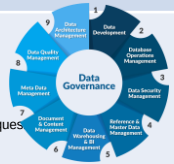
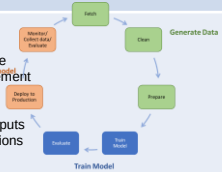
Model interpretability outputs: metrics such as model coefficients, ELI5, or LIME outputs that allow further investigation to understand ***how the models are making predictions*** *to identify potential overfit or bias that was not found during training.*

Model outputs and decisions: what predictions our models are making given the production input data, and also ***which decisions are being made with those predictions***. Sometimes the application might choose to ignore the model and make a decision based on pre-defined rules (or to avoid future bias).

User action and rewards: **based on** further **user action**, we can capture reward metrics to ***understand if the model is having the desired effect***. For example, if we display product recommendations, we can track when the user decides to purchase the recommended product as a reward.

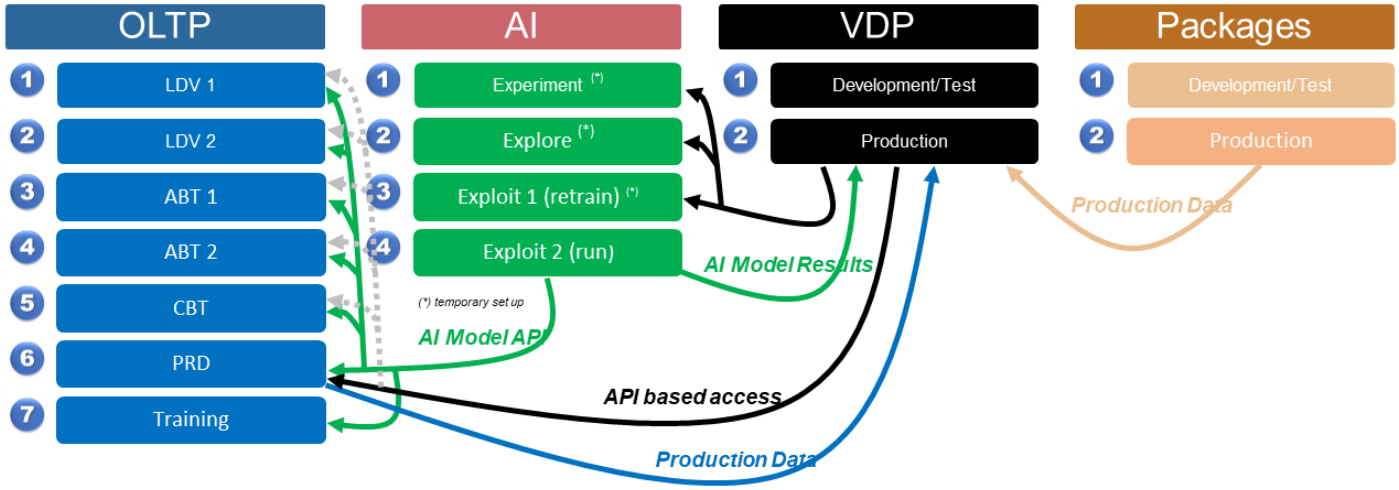
Model fairness: ***analysing input data and output predictions against known features*** that could bias, such as race, gender, age, income groups, etc

VDAB

	Dev(Sec)Ops	DataOps Data starts with the same letter as Discipline	MLOps *Creating value out of data* Data kwaliteit is functie van de "purpose"
Goal	**Starts from Processes & Requirements** Design, build, integrate, run in house developed applications and of the shelf packages in the most (cost) efficient way. CI/CD/CT (continuous testing) using mainstrain development tools & environments (Java, C#, frameworks, …) whereby Development & Operations operate as one team with maximal automation of repetitive SDLC tasks.	**Starts from Data & Information needs** Deliver (real-time) qualitative data, fit for different purposes, through governed and industrialized data pipelines starting from different sources like operational data, events, external data, open data, …	**Starts from challenges and data: what if we could … ?** Design, build, run (retrain) and deploy AI-models using • the data services of DataOps. • Specific tools and frameworks (e.g. Python, Tensorflow, BERT, Python Frameworks, …) • CI/CD/CT (Continuous Training) Every AI-model is accessible through API's.
Roles & Skills Generic: Project leaders	Business Architect, Solutioner, Technical Architect, Developer, Database Developer, Data Modelling, Operations Engineer	**Data Guard** Data Solutioner, Data Architect, Data Engineer, Data Modeller, DataOps Engineer	Data scientist, ML Expert, Data Engineer, MLOps Engineer
QA	Automated testing Security testing Sonarcube	Test Data Pipelines Check Data Quality Test VDP Data Provisioning tp consumers	Check Data Quality – Check Bias in Data – Check Model quality – Check Bias in Model + robustness (monitoring)
Governance		- **Data Guard role** - Data Architectural Rules & Policies - No metadata = no data - No Swamp rule - 1 month rule - Data solutioning - Central management of business requests - Service Catalogue - Self service: controlled roll out of self-service tools - Focus on Data Literacy - Data Solutioning 	• Led by Digital Ethics • Ethical Board • Legal Constraints on data usage • Automated Model Risk Management • Model Monitoring • Model inputs • Model interpretability outputs • Model outputs and decisions • User action and rewards • Model fairness 
SDLC	LDV(2)-ABT(2)-CBT-PRD ??? Automated QA CI/CD	2-layer model (Test/Development and Production) Industrialize Data Pipelines	Data Management Phase-Model Training Phase-Serving Management Experiment-Explore-Exploit(2) 2-layer model (Test/Development and 2 Production) Continuous Delivery for Machine Learning (CD4ML) Industrialize Data Pipelines
Development Stack & Tooling	Technical Reference Architecture Coding guidelines (enforced) Java, Spring, Axon, …	Data Vault, Python, S3, Glue, Dremio, Snowflake, Redshift, …	R, Python, TensorFlow, Bert, Sagemaker, Java, …
Deploy	**On premise:** VMs, **Cloud :** Containers/Kubernetes Containers CD/CI	On premise + Cloud Containers (sandboxes) Usage Monitoring	Deployes As A Service : MaaS (API driven) On premise + cloud: Containers CD4ML Model Technical Monitoring
Infrastructure	VMs, Containers	Cloud	High performance (CPU, Memory) (bullions, OBDA, Cloud)

CLOUD OPERATIONS

Environments: Interoperability. Who uses what?



C-level support	Experiment	Explore	Exploit
Phase 1 • BIAS Awareness • Business Informed • No business case	• Purpose (look for challenges not use-cases) • Data • People (competences) • Tools (e.g. Python, Neo4J, TensorFlow, …) • Computing power • Data exploration	X	X
Phase 2 • Keep experimenting • Start working on Data & AI Literacy • Set up Digital Ethics/Ethical AI • Business Consulted	• Multiple simultaneous experiments • Enrich internal data with external data • People must spend 20% of their time in experimenting and exploring new technologies & data	• API first • Manually Set up (realtime) data pipelines • Choice of tooling • Manual Model Risk Mgmt • First steps in MLOps	• Integrate in operational systems • Integrate with operational databases • APIs published • Data Pipelines run manually • Manual deployment of model • **Manual Model Risk Management**

Phase 3
• Keep experimenting
• Data & Literacy rolled out
• Ethical AI in place
	• Find solutions for new problems/challenges
	• Identify new and better solutions for existing problems
	• Keep looking for new ideas, technologies for the
		• Project oriented approach
		• Several parallel projects
		• Model Risk management set up for every new model
		  • Model inputs
		  • Model interpretability outputs
			• API first -versioned
			• Industrialized Data Pipelines
			• **Regular retraining of AI-models**
			Industrialized Model Risk Mgmt
			  • Version mgmt of data,

VDAB Ethical AI programma

Trust – Transparency – Benefit

“ Technology is neither good nor bad; nor is it neutral.”

First Law of Technology, Melvin Kranzberg (1917-1995), *Professor, and Co-Founder, Society for the History of Technology*

Mensen

- Inclusie & diversiteit in het team
- Multi-disciplinaire teams

Social Bias

- Gedrag in de markt
- Gedrag binnen bepaalde groepen
- Sociale realiteit

Data

- Verantwoorde keuze van de variabelen
- Kritisch kijken samen met business experts
- Checks op de data
- Data cleaning

Model

- Impact analyse op het model van de gebruikte variabelen
- Geautomatiseerde BIAS detectie middels tools
- Peer review van de gebruikte methodes

ETHICS GUIDELINES FOR TRUSTWORTHY AI

INDEPENDENT HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE
SET UP BY THE EUROPEAN COMMISSION

Ethical Principles in the Context of AI Systems
(i) Respect for human autonomy
(ii) Prevention of harm
(iii) Fairness
(iv) Explicability

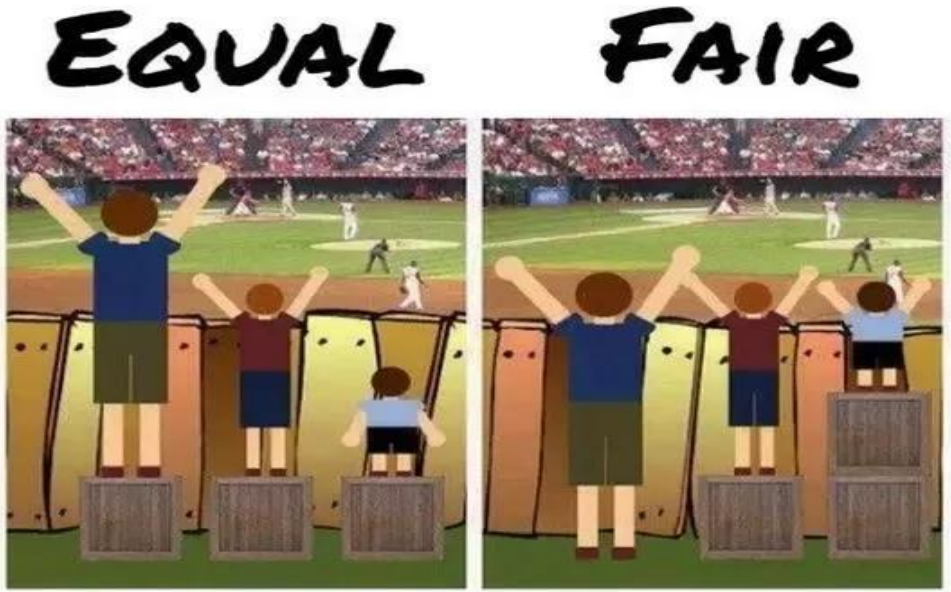
Trustworthy AI has three components, which should be met throughout the system's entire life cycle:
1. it should be **lawful**, complying with all applicable laws and regulations;
2. it should be **ethical**, ensuring adherence to ethical principles and values; and
3. it should be **robust, both from a technical and social perspective,** since, even with good intentions, AI systems can cause unintentional harm.

Requirements of Trustworthy AI

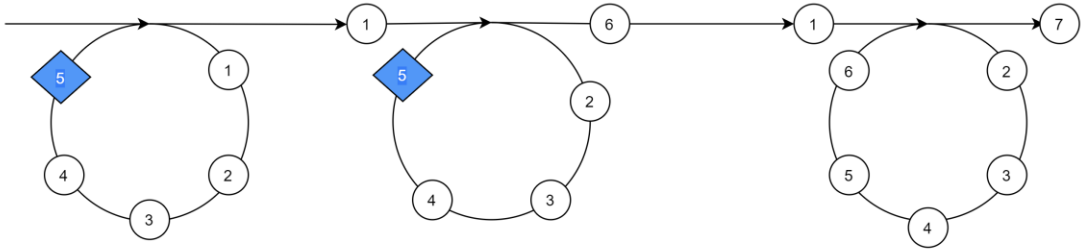
1. **Human agency and oversight**
   Including human control and human supervision
2. **Technical robustness and safety**
   Including resilience to attack and security, fall back plan and general safety, accuracy, reliability and reproducibility
3. **Privacy and data governance**
   Including respect for privacy, quality and integrity of data, and access to data
4. **Transparency**
   Including traceability, explainability and communication
5. **Diversity, non-discrimination and fairness**
   Including the avoidance of unfair bias, accessibility and universal design, and stakeholder participation
6. **Societal and environmental wellbeing**
   Including sustainability and environmental friendliness, social impact, society and democracy
7. **Accountability**
   Including auditability, minimisation and reporting of negative impact, trade-offs and redress.

VDAB

Fair or Equal



Experiment-Explore-Exploit



EXPERIMENT

1. Define AI use case
2. Initiate experiment and iterate
3. Define the sensitive variables
4. Perform qualitative assessment
5. /* **Milestone**: Provide feedback to move to next phase

EXPLORE

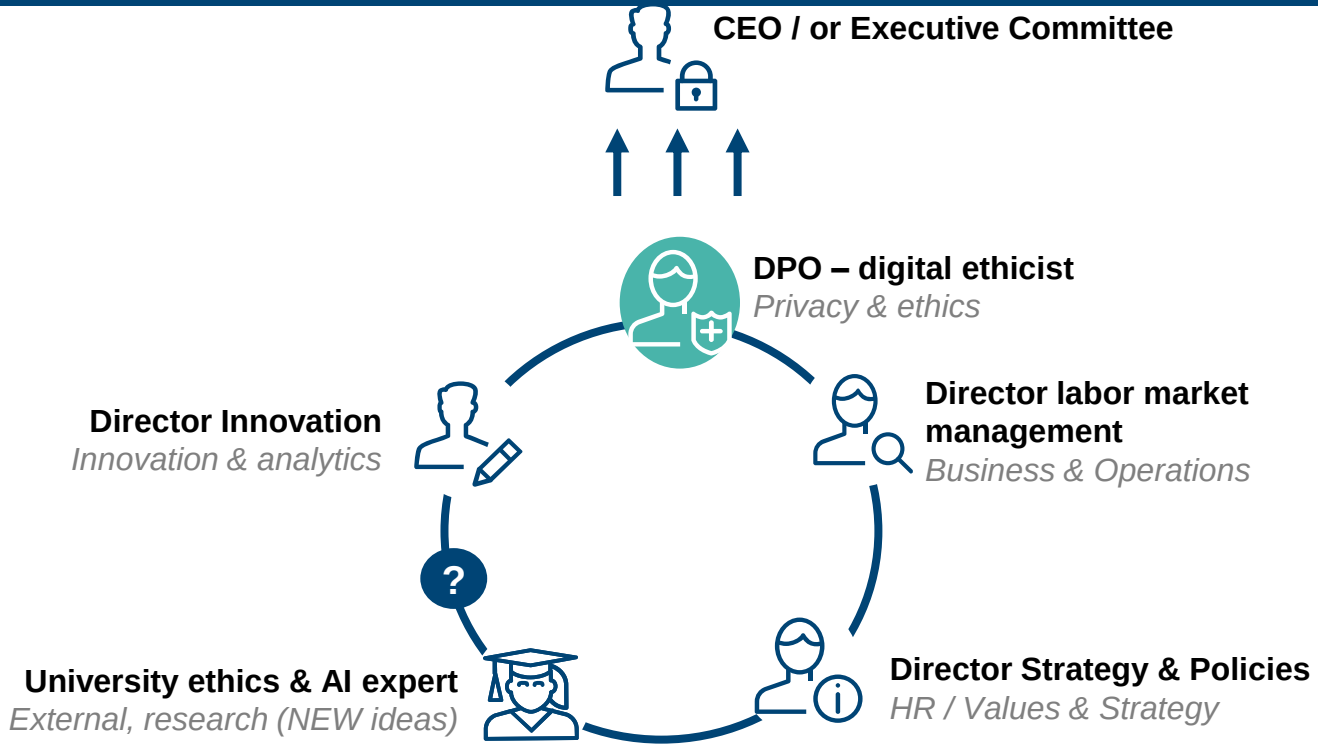
1. Select fairness metrics & define fairness for the use case
2. Perform qualitative assessment
3. Perform quantitative assessment
4. Create & share assessment results
5. /* **Milestone:** Review assessment results & provide feedback
6. Prepare for internal communication

EXPLOIT

1. Handovers (1. production ready AI systems to software factory 2. Usecase to Business Owner)
2. Share / update internal communication
3. Perform qualitative assessment checks
4. Monitor and assess data & model checks using the quantitative assessment asset
5. Create & Share assessment results
6. Review assessment results & provide feedback
7. Create & Share external communication material

Model Risk Mgmt Industrialization

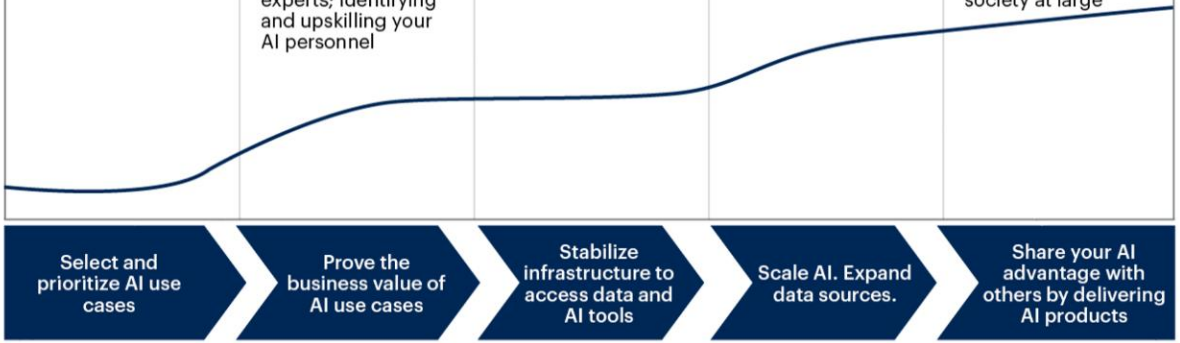
A sustainable operating model ‘Ethical board structure’
There will be decisions to be made …



Gartner AI Maturity Model

AI Maturity Model

Level 1 Planning	Level 2 Experimentation	Level 3 Stabilization	Level 4 Expansion	Level 5 Transformation
• Early AI interest with risk of overhyping • Pioneers explore available AI techniques and solutions • First AI (speculative) use cases identified • Agree upon use case success criteria, such as KPIs, with stakeholders • Determine what AI means for your organization and define AI accordingly	• AI experimentation: in a data science context, or with cloud AI APIs or with intelligent applications for predictive analytics • Successful AI POCs, proving that a technology (e.g. computer vision, speech recognition, ML) can deliver business value • First pilots, typically on revenue increase use cases (price optimization, product recommendations) • AI lab. First recruitment of AI experts; identifying and upskilling your AI personnel	• Several AI use cases in production, creating value with cost reduction use cases (e.g., process optimization or product/service innovations) • Budget is assigned to corporate function(s) • Mastering end-to-end AI development and implementation culminating in standardization on a platform for AI • AI community of practice or (virtual) center of excellence	• AI is pervasively used across the organization • C-level ownership of AI • AI ethics, governance and risk management • Hybrid centralized/ decentralized AI organization • Strategies and ongoing innovations with AI as a given • Significant impact on workforce, roles and skills	• AI is seamlessly integrated for digital process and chain transformation, and disruptive new business models • AI is part of business DNA • Shift from AI projects to AI products • Synergy of human intelligence and AI in augmented intelligence • A combination of in-house and external hiring • Profound impact on people, culture and society at large



Source: Gartner

VDAB

Thank you