PROPHECY LABS

AI4Business

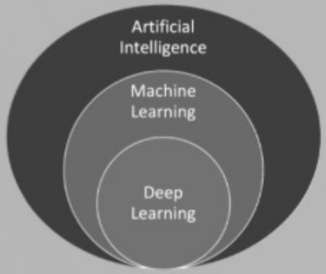
Practical AI Challenges for Business Managers

Past & Future AI4Business Webinars

Introduction to AI for Business Managers

Webinar: Introduction to AI for Business Managers

"For managers that want to start capturing business value with AI solutions"



When: Thursday, August 19 - 10AM CET

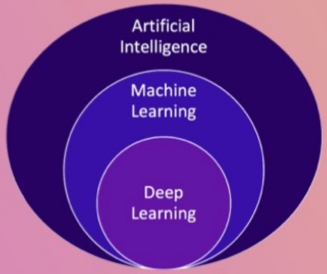
Duration: 1h (including Q&A)

Host: Roel Henckaerts

ML model lifecycle

Webinar: The ML Model Lifecycle

"We are going to explore the set of practices and principles necessary to deploy and maintain Machine Learning solutions in production"



When: Thursday, August 26 - 1 PM CET

Duration: 1h (including Q&A)

Host: Roel Henckaerts

Roadmap AI4Business Course

Part I	Part II	Part III	Part IV	Part V
Introduction to AI	Developing AI tools	Data and Value	Deploying AI	Monitoring

Table of contents

1 Real Life AI

AI is everywhere

Alphabet
Google

Self-driving cars
Play several Atari games and Go
Voice interface to make phone calls & schedule your appointments

IBM

Man vs. machine competitions:
Deep Blue (chess)
Watson (tv quiz Jeopardy)
Debater (professional debates)

Alibaba

Predict what you buy
Generate product descriptions
Reduce traffic jams in smart cities
Monitor farming crops

amazon

Digital assistant Alexa
Ship before you buy
Buy without checkout

NETFLIX

Content recommendations
Optimized streaming
Autogenerate personalized thumbnails

FaceID & Siri
Song recommendations
Navigation in Maps

If AI systems are everywhere…

- *How easy is it to build an AI solution?*
- *What are the requirements to build an AI solution?*
- *Is building an AI solution the same as building any piece of software?*
- *What are the challenges to make my AI solution work?*
- *What kind of special tools do I need to build an AI solution?*

Why are AI solutions so difficult?

Data Management

Configuration

Entanglement

Different expertise

Model Governance

2 AI Challenges

Challenges of building AI systems

Any software application comes with many challenges.

AI/ML brings around a couple of extra ones:

- Data is difficult to manage and resource consuming
- Iteration is necessary but slow
- The expertise needed is abundant and diverse
- Scaling quickly becomes an issue
- Maintenance becomes particularly difficult
- Selecting the right tool is not always so easy

Data is an investment

Having easily available and high quality data is expensive

Why invest in data?

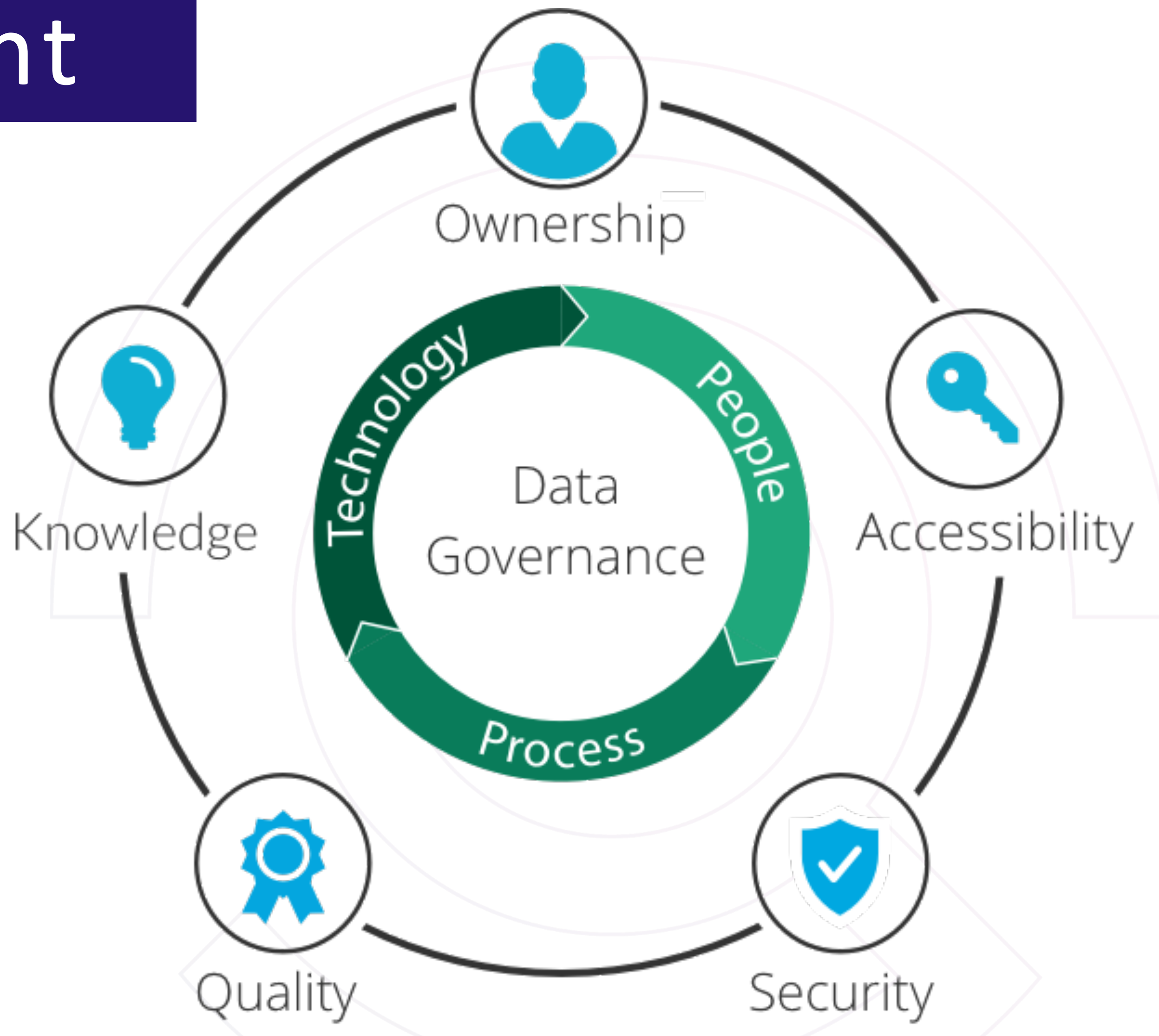
- Model quality depends on data quality
- Data is needed after deployment
- Data is worthless if not usable
- Data is at the core of the AI system

- Bad data increases complexity → need easy access to high quality data

Data is an investment

Data governance is key

- Being data driven is more than just buying expensive data
- Having the processes is as valuable as having the right tool
- Having a cohesive data strategy is the key to success.
- Data governance is not the same as Data management



Inperva- Data Governance

Data can significantly change

Data distributions can shift.

Assumption that past data is representative of future data is broken.

Data Drift

- Distribution of the features or target changes
- Past performance does not guarantee future results
- Models are not ever lasting but speed of decay increases

Concept Drift

- Occurs when patterns learned by the model no longer holds
- It might happen over time or suddenly
- Is more difficult to correct as is related to fundamentals

Iteration is a must

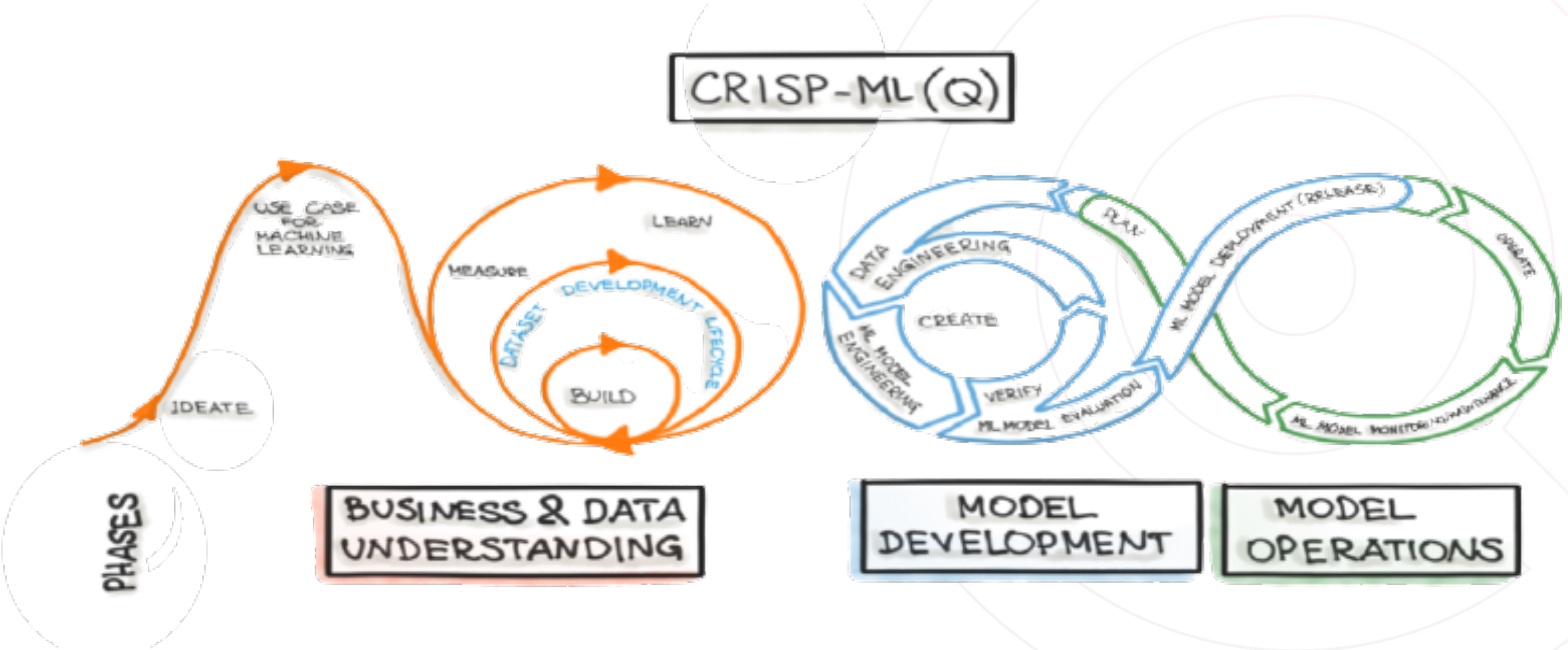
Be ready to iterate over:

- What does the business need?
- Do we have data as needed?
- Is data ready for modelling?
- What model should we build?
- Is our model good enough?
- How do we make results available?

CRISP-DM

Iteration is a must

As complexity of the environment increases, so does the workflow



[MLOps.org - CRISP-ML(Q)]

Business and Technical Leaders

Aligning business and technical leaders is not always easy.

But it is necessary to bridge the gap:

Business leaders

- Update their Data/AI literacy
- Understand the uncertainty around AI systems

Technical leader

- Set right expectations ahead of time
- Plan resources efficiently

Business and Technical Leaders

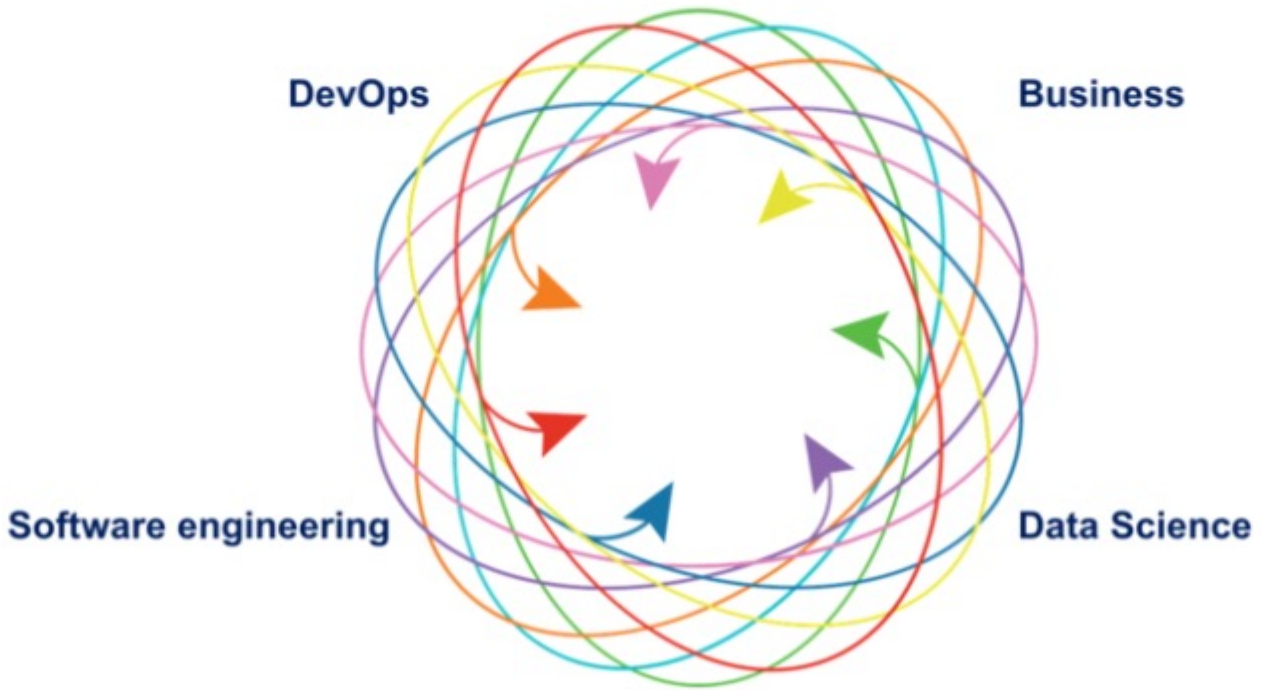
If deploying AI for the first time:

- Start small
- Look for low hanging fruits
- Look for problems with visible value

Remember: AI is not bulletproof, but when used correctly can be an extremely powerful tool

AI is not going to replace managers, but managers that use AI will replace those that do not

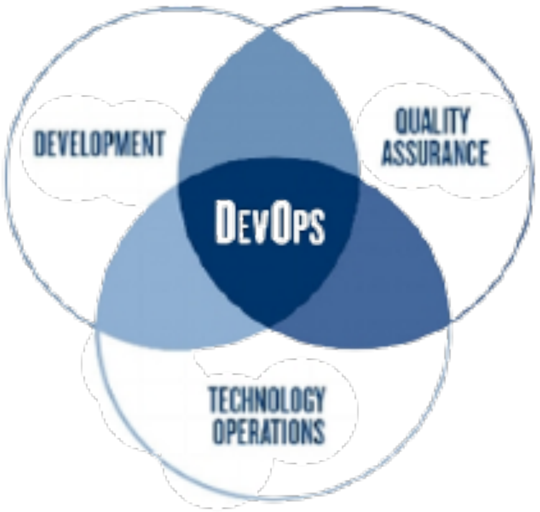
Technical teams working together



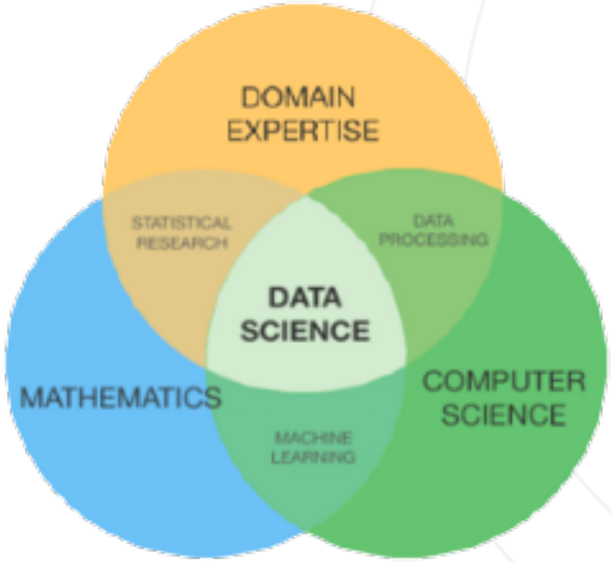
Technical teams working together

DevOps, IT and Data Scientist often organized in silos at organizations.

These silos must be connected* for AI. **Unless you found a unicorn that can do everything*

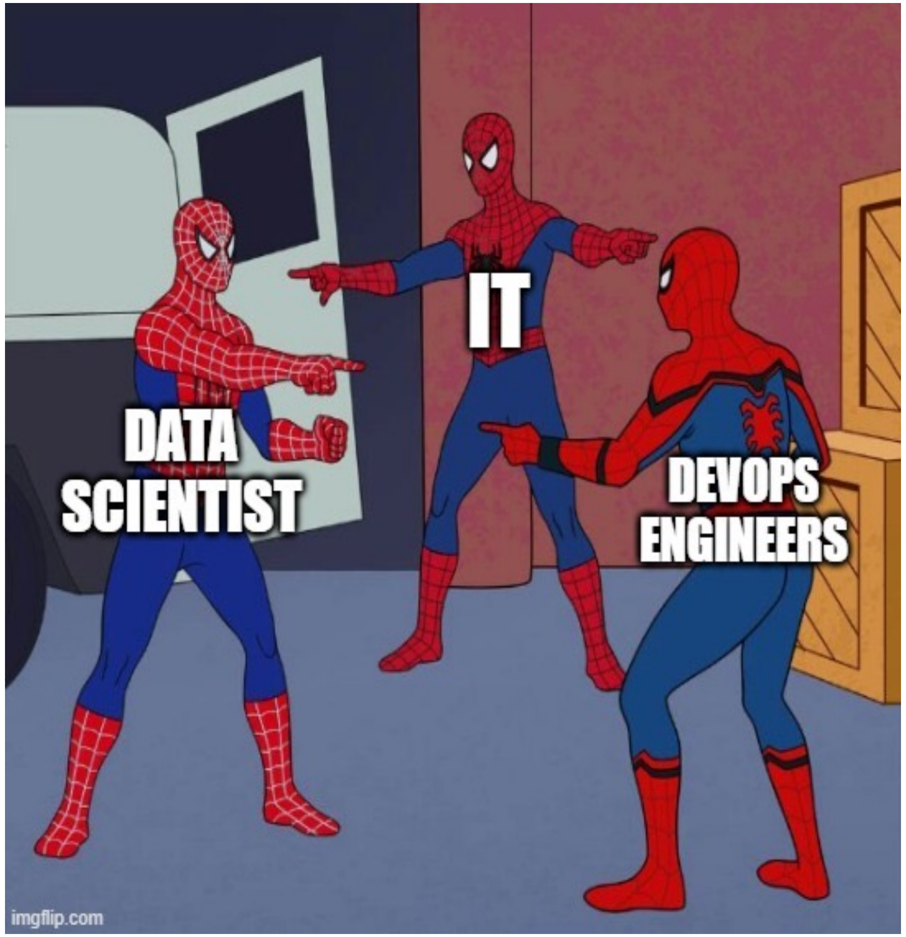


+



=

Technical teams working together

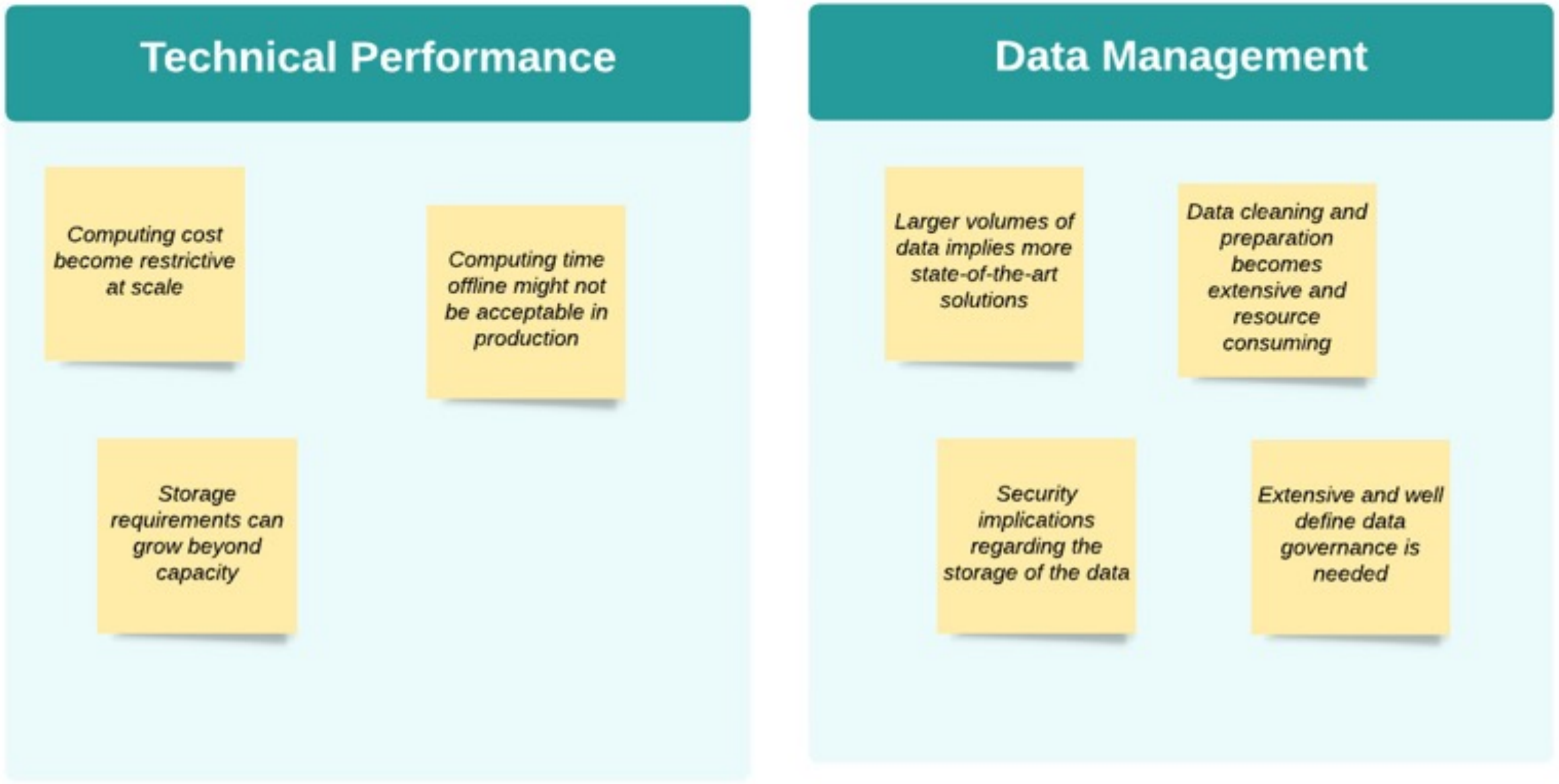


When working in silos:

- Impossible to have a high-level overview of the solution
- Constant blaming across teams
- Can’t tackle complex problems (e.g. real time applications)
- Maintenance rapidly becomes a nightmare

Think about scaling

Scaling AI solutions is not easy nor cheap



Maintaining AI solutions

"As the machine learning (ML) community continues to accumulate years of experience with live systems, a wide-spread and uncomfortable trend has emerged: developing and deploying ML systems is relatively fast and cheap, but maintaining them over time is difficult and expensive."

Sculley et al.

Selecting right technology

Selecting the right tool for the problem at hand is not always simple, as the technology supporting AI is

- Diverse
- Fast growing
- Tailored

Remember: Don’t marry yourself to a tool. Tools are just means to an end.

Selecting right technology

Some general tips on selecting the right technology

Integration should be easy

- You are already on an ecosystem, new tools need to be easily integrable.

Flexibility is key

- Tools should easy to use and flexible to customization

Scalability is your friend

- Not all tools scale well for all problems

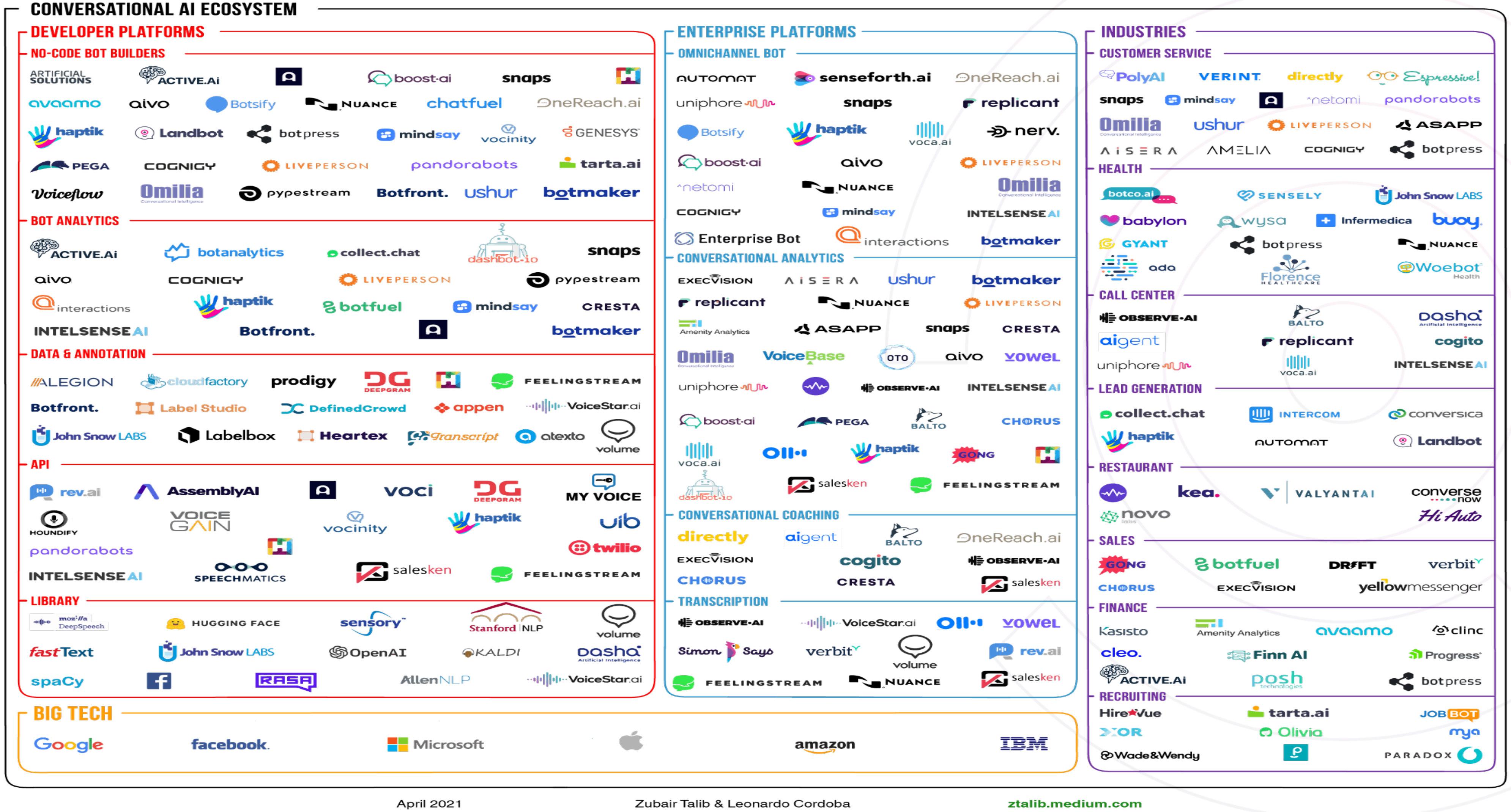
Right tool for the right job

- There are many tools available, no need to overspend

Support is important

- Having good support and a large community behind is key

AI Ecosystem is crowded



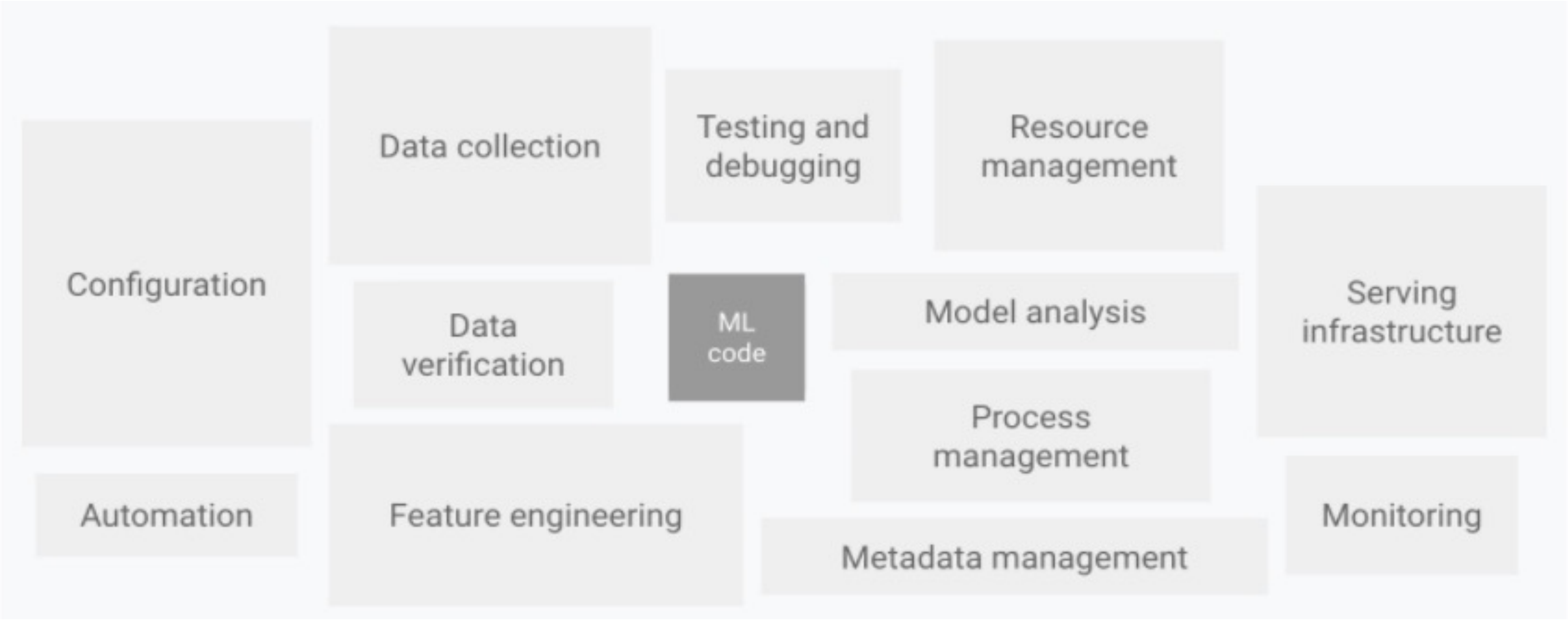
3 PoC to Production Gap

PoC versus Production

"All of AI, .., has a proof-of-concept-to-production gap. The full cycle of a machine learning project is not just modelling. It is finding the right data, deploying it, monitoring it, feeding data back [into the model], showing safety—doing all the things that need to be done [for a model] to be deployed. [That goes] beyond doing well on the test set, which fortunately or unfortunately is what we in machine learning are great at."

- Andrew Ng

The big picture



[Google - CD and automation pipelines in ML]

Basic ML building blocks

Data Management	Experimentation	Production
Process and govern the data used by models: • Usually large data sets • Should be of high quality • Should be compliant with legislation • Should be tracked	Build a model based on business requirements, after iteration of experimentation: • Workflow is iterative • Experiment should be tracked • Code should have standards • Accuracy metrics should be tracked • Retraining should be possible • Requires specific infrastructure	Integrate prediction into production and business processes: • Generate systematic predictions • Track performance across time • Follow best engineering practices

Moving to production is hard

(Not so) Fun fact

According to VentureBeat, roughly 1 out of 10 Machine Learning models actually makes it into production. But why?

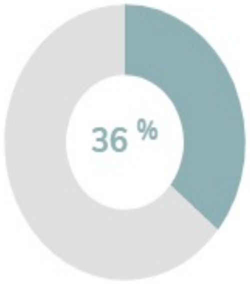
The Set up is not right

- Bad infrastructure
- Disconnect between the relevant parties
- Poor data management
- Leadership doesn't understand

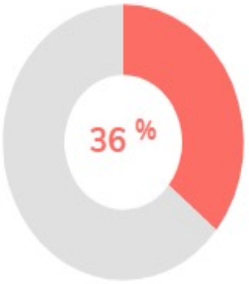
ML has its own difficulties

- Scaling is not easy
- Duplication is widespread
- Management not on board
- Lack of Reproducibility
- Support across technologies

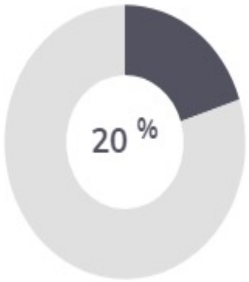
Deploying models takes time



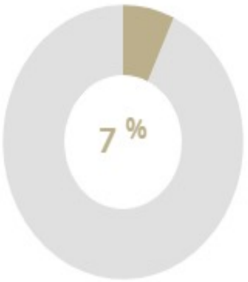
36% of survey participants said their data scientists spend **a quarter** of their time deploying ML models



36% of survey participants said their data scientists spend **a quarter to half** of their time deploying ML models



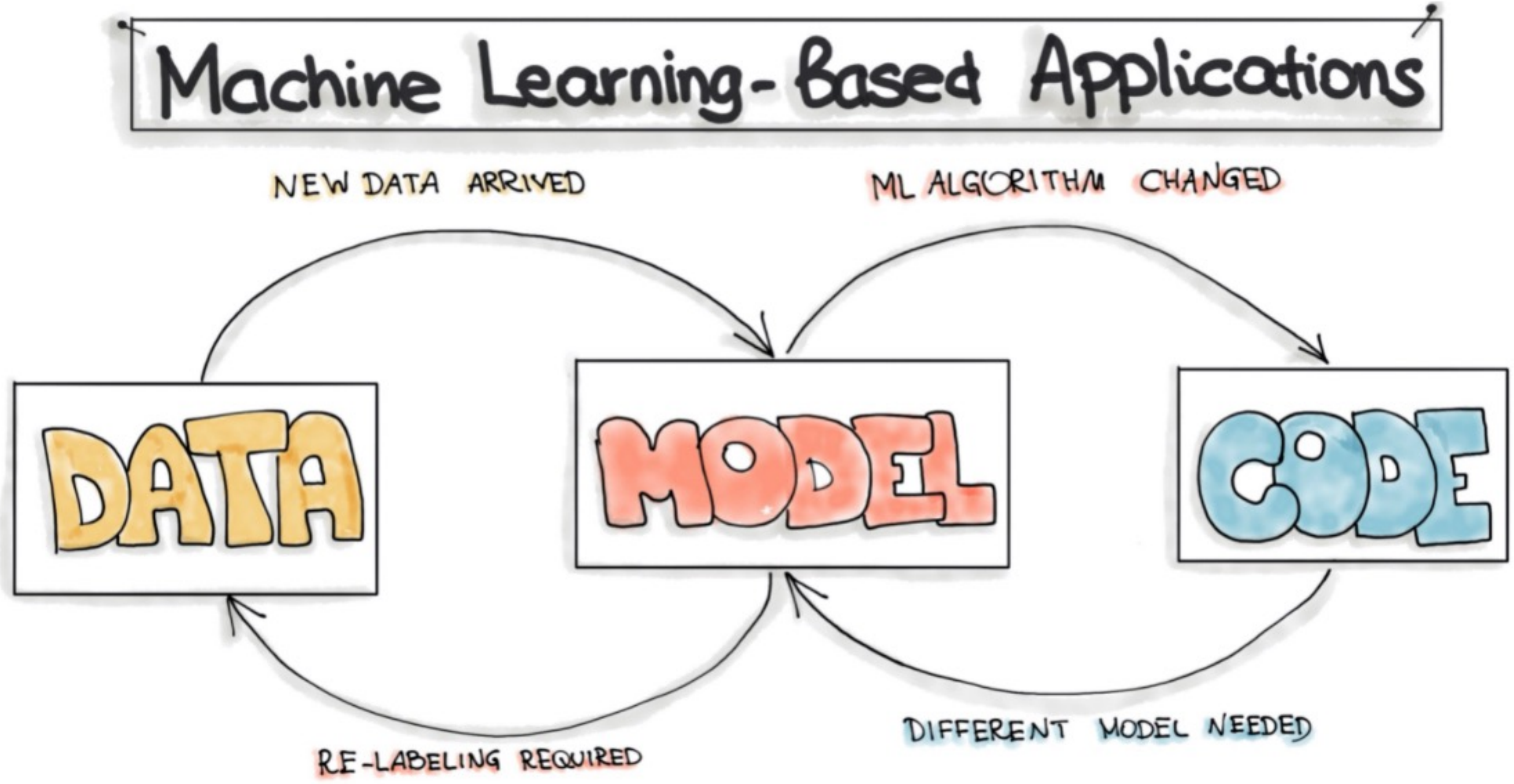
20% of survey participants said their data scientists spend **half to three-quarters** of their time deploying ML models



7% of survey participants said their data scientists spend **more than three-quarters** of their time deploying ML models

Algorithmia - State of Enterprise ML

Changing anything changes all



[MLOps.org - Motivation for MLOps]

Hidden technical debt

Developing and deploying ML systems is relatively fast and cheap, but maintaining them over time is difficult and expensive. Some of the reasons for this are:

- Data dependencies cost more than code dependencies
- Feedback Loops
- ML-Systems anti-patterns
- Configuration debts
- Always changing external world
- Other ML related debt (e.g Data testing, Reproducibility debt)

Other production issues

- Data quality:
  - ML models reflect the data they are build on, so they are very dependent on its size and quality
- Model decay:
  - As times goes by, there might be changes in behavior that the original data would not necessarily reflect causing the quality of the model to drop
- Locality:
  - The quality of the performance of ML model does not always translates completely to production

4 Trusted AI & Ethics

Trusting AI systems

- Any practical AI system in production needs to be:

- Fair
  - Not allowing for any bias or discrimination
- Robust
  - Not able to be manipulated from the outside
- Explainable
  - Able to understand the internal decision process

- Need for AI governance and responsible AI
  - Technical solutions exist, but at some costs (e.g., slower execution)

Fairness

- No discrimination against minorities or bias in decisions
- Bias is often present in data and transferred into models
  - Toxic effects of reinforcing existing unhealthy stereotypes
- Some recent examples
  - Facial recognition worked better for light-skinned males (Buolamwini)
  - Man is to computer programmer as women is to homemaker? (Bolukbasi)
  - Amazon’s hiring tool discriminated against women (Reuters)

Robustness

- Not able to be manipulated by a third party via adversarial attacks
  - Deliberately force to make a wrong prediction and trying to fool the AI
- Make the system do something else than it is intended to do:
  - Stickers on stop sign confuse the AI
  - Patch that tricks AI into thinking a banana is a toaster
  - Glasses make facial recognition AI think you're actress Milla Jovovich
- Adversarial use of AI
  - Obama Deep Fake video

Explainability

- Understand why a specific decision is made
  - User has the “right to an explanation” (GDPR)
  - Especially important for high-stakes decisions with a big impact on lives
- Wolf vs husky experiment (Ribeiro et al.)
  - Snow in the background? → Husky
- Two options to guarantee explainability
  - Tranparent models
  - Ex-post interpretation techniques of black box models (many exist)

AI4Business

PROPHECY LABS

ctg

nannyML

ESF
INVESTEERT IN JOUW TOEKOMST

Europese Unie

Vlaanderen is werk