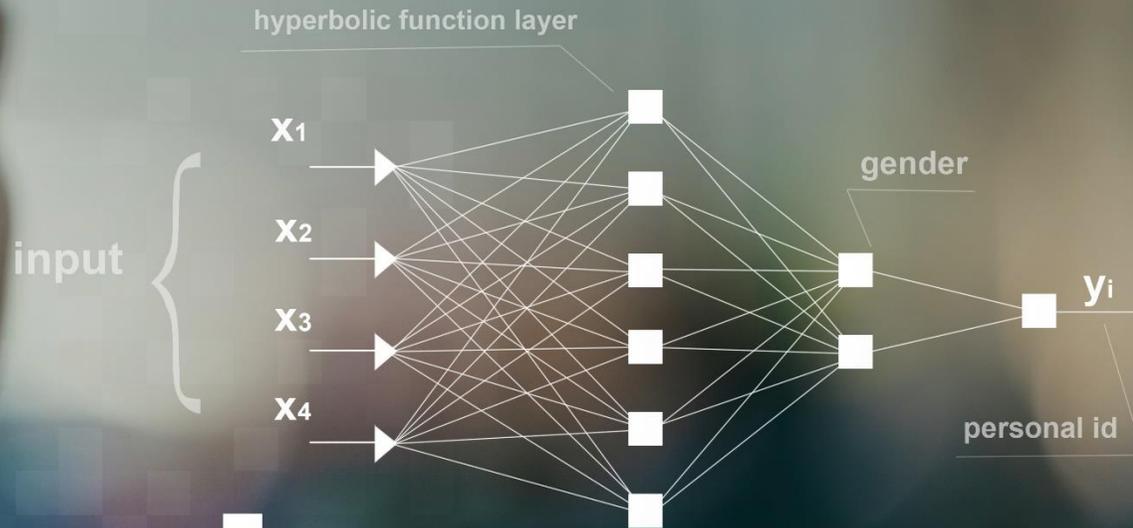




Algorithm Oversight

neural network

face recognition



The growing need of oversight on the use of algorithms in society

9th of June 2021

kpmg.nl

Introduction



Frank van Praat
Head of Trusted Analytics
KPMG Netherlands



Mourad Fakirou
Senior Consultant Trusted Analytics
KPMG Netherlands

Many incidents with algorithms have lead to an increasing need of oversight

 Euronews

Twitter is removing its photo crop algorithm that prefers white people and women

Euronews

 Insurance Journal

Retired Black Players Fight NFL Brain Injury Payout Algorithm Based on Race

Insurance journal

 Forbes

Data Dies In Darkness: Getting AI Algorithms To Think Outside The Black Box

Forbes

Faulty Aegon investment models see investors get \$97m refund

By Kirsten Hastings, 28 Aug 18

Work by an inexperienced junior analyst contained errors and 'didn't work as promised'

International-adviser

Apple's 'sexist' credit card investigated by US regulator

🕒 11 November 2019

BBC

 VICE

How a Discriminatory Algorithm Wrongly Accused Thousands of Families of Fraud

Vice

RETAIL OCTOBER 11, 2018 / 1:04 AM / UPDATED 2 YEARS AGO

Amazon scraps secret AI recruiting tool that showed bias against women

By Jeffrey Dastin

8 MIN READ



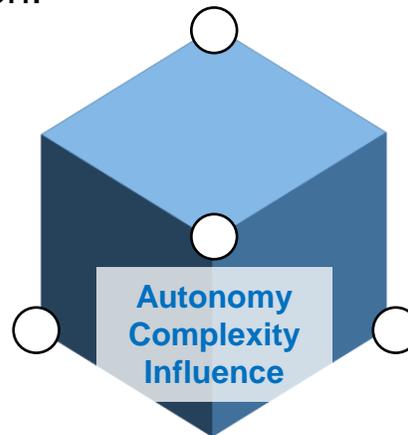
Reuters

Algorithms are a catch-all term, instead it is about automated decision making and its impact

All types of organizations can utilize algorithms' benefits. E.g. retailers for a better determination of prices, banks for better mortgage decisions and public organizations to automated decisions for citizens. Besides that, algorithms are a *means* in a broader process. In short, oversight is needed everywhere, and should have taken place already. Hence, the question for oversight is not on *organizational level*, but on *algorithm level* instead: the **impact** of the algorithm is key.

Autonomy: Decision making process based on automated processing of data, without an effective “human in the loop” making the decision.

Complexity: Advanced algorithm, based on a form of artificial intelligence or complex coherence with other algorithms.



Influence: Interaction with the algorithm impacts the rights of an individual, group or organizations.

The question for oversight is **not** depending on the type of organization, but on the **impact of the algorithm** instead

	Autonomy	Influence	Complexity
 High degree	<p>Decision making process based on automated processing of data, without an effective “human in the loop” making the decision</p> <ul style="list-style-type: none">• There is limited time/space to assess the algorithm’s outcomes• Based on the same information, an individual can never make such a decision (e.e. due to the time effort or lack of knowledge)• The “human in the loop” has few options to override the outcome of the algorithm	<p>Interaction with the algorithm impacts the rights of an individual, group of organizations</p> <ul style="list-style-type: none">• The results of the algorithm have a direct influence on an individual decision, with legal consequences or a determination that precedes it.• The results of the algorithm provide the possibility that individuals and companies feel more affected than others by discrimination or <i>singling out</i>.	<p>Advanced algorithm, based on a form of artificial intelligence or complex coherence with other algorithms</p> <ul style="list-style-type: none">• The application uses technology typically classified as artificial intelligence (/machine learning)• The application uses large amounts of data, often unstructured in nature.
 Low degree	<ul style="list-style-type: none">• There is a clear “human in the loop” with the time and space to judge each outcome of the algorithm• The “human in the loop” has the same information at his/her disposal as the algorithm• The “human in the loop” has the mandate to make decisions for themselves, even if they conflict with the algorithm’s outcomes	<ul style="list-style-type: none">• The results of the algorithm have no legal consequences for an individual and/or group• The outcomes of the algorithm have only a limited influence on a final decision	<ul style="list-style-type: none">• The application uses “traditional” data analysis techniques such as rule-based analyses.• The application uses clear data in a fixed structure.

In addition, the use of algorithms and AI has a number of specificities that play a role oversight



There is a fundamentally different design process



Explainability is needed due to the 'black box'



Continuous monitoring is required during the use



Ethics and compliance are important themes



Responsibilities around algorithms are different



Professional devaluation and *fool with a tool* are lurking

Who should supervise algorithms? In the current situation there is accumulation, and different ways of thinking

Centralized

One central body is responsible for all oversight of algorithms. This can possibly be divided per sector.



Pros
Centralized knowledge, one go-to-point, already existing. **One** centralized approach familiar for all organizations

Cons
Difficult to bring all knowledge together, requires considerable upscaling, specific domain knowledge is lacking. Difficult to scale in a general sense.

Pros
Centralized knowledge, one point of contact, already existing, can be built from the ground up with a clear approach.

Cons
Difficult to bring all knowledge together, requires considerable upscaling, specific domain knowledge is lacking. Difficult to scale in a general sense.

NEW!

Existing



Pros
Specific domain knowledge, existing structures, responsibilities and mandate. Easily scalable.

Cons
Technical knowledge decentralized, required to be built up per institution. Likely to lead to different forms of supervision/opinions/enforcement

New



Own responsibility
for example using three lines of defense and existing reporting structures (annual report, etc.)

Pros
Specific domain knowledge. Existing structures, responsibilities and mandate. Easily scalable. Responsibility lies with the organization itself.

Cons
Relatively free format, enforcement unclear.

Decentralized

Oversight is organized in a decentralized way and assigned to various organizations. This can be a distinction on sector but also on technological and functional level

Oversight helps pushing organizations to think of responsible implementation of algorithms

Oversight can be approached in two different ways : (1) preventive, also called: “ex ante” of (2) detective, also called “ex post”. Both approached have their pros and cons..

	Ex-ante oversight	Ex-post oversight
Pros	<p>This means that algorithms are required to be approved <i>prior to</i> deployment. This prevents major errors, but slows down innovation.</p> <ul style="list-style-type: none">• By checking algorithms in advance, incorrect algorithms are prevented and only reliable algorithms are actually used• There is no need to issue fines. After all, organizations will be corrected for the potential “offence”	<p>This means that algorithms are checked after deployment. This fosters innovations, but does not always prevent errors.</p> <ul style="list-style-type: none">• Puts the responsibility with the organization itself, so that they can reap the benefits, but also bear the risks themselves.• Encourages innovation. There is still room for speed and competitive advantage.• Scalable. Targeted (risk-based) extra checks can be made afterwards.
Cons	<ul style="list-style-type: none">• Scalability. It is virtually impossible to pre-assess all algorithms that are being developed. Even when the pre-check only applies to high-risk algorithms..• Slows down innovation and competitive advantage. Organizations have little chance to create speed and distinguish themselves from their competition.	<ul style="list-style-type: none">• Errors persist. Not all errors can be prevented by checking afterwards.• Requires a foundation of trust, which in some cases cannot be met..

Oversight requires insight on three levels of algorithm control in organizations

Most important themes and subjects



- Does the organization pay attention to the risks of using algorithms **at the highest level**? With clear **roles and responsibilities** and reporting structures?
- Does the organization have **sufficient** governance around the use of algorithms with functions such as privacy, security and data?



- Does the organization ensure **clear development and control/management processes** to ensure that algorithms are used sufficiently reliably, fairly, transparently and legally
- Is the organization **accessible** to citizens/consumers/customers to answer questions?
- Has the organization provided **the right people and (technical) resources** to carry out these processes?



- Are **the individual algorithms** used by the organization of sufficient quality and reliable?
- Has the organization made an estimate of the **potential negative consequences** for **each algorithm**?

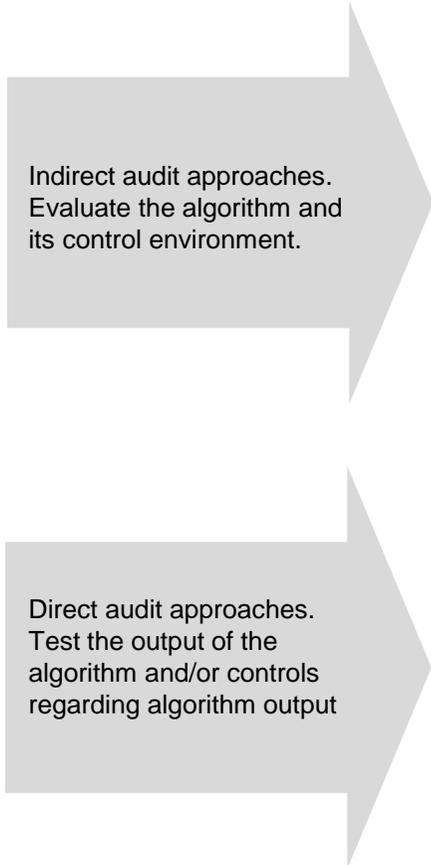
Four audit approaches to address algorithms

Object of investigation

Algorithm overall control environment

Algorithm design and maintenance

Algorithm output



Audit approach

 Evaluate algorithm entity level controls

 Test the model

 Test monitoring controls

 Substantive procedures

— Evaluate if entity level controls are in place to ensure algorithms are built in a controlled environment.

— Perform an in-depth assessment to determine if the algorithm performs in line with relevant criteria (including GITC when testing ToE).

— Test if internal controls are in place to monitor the transactions performed by the algorithm and mitigate the risks of algorithm failure.

— Test if (a sample of) the transactions were processed by the algorithm in line with relevant criteria.

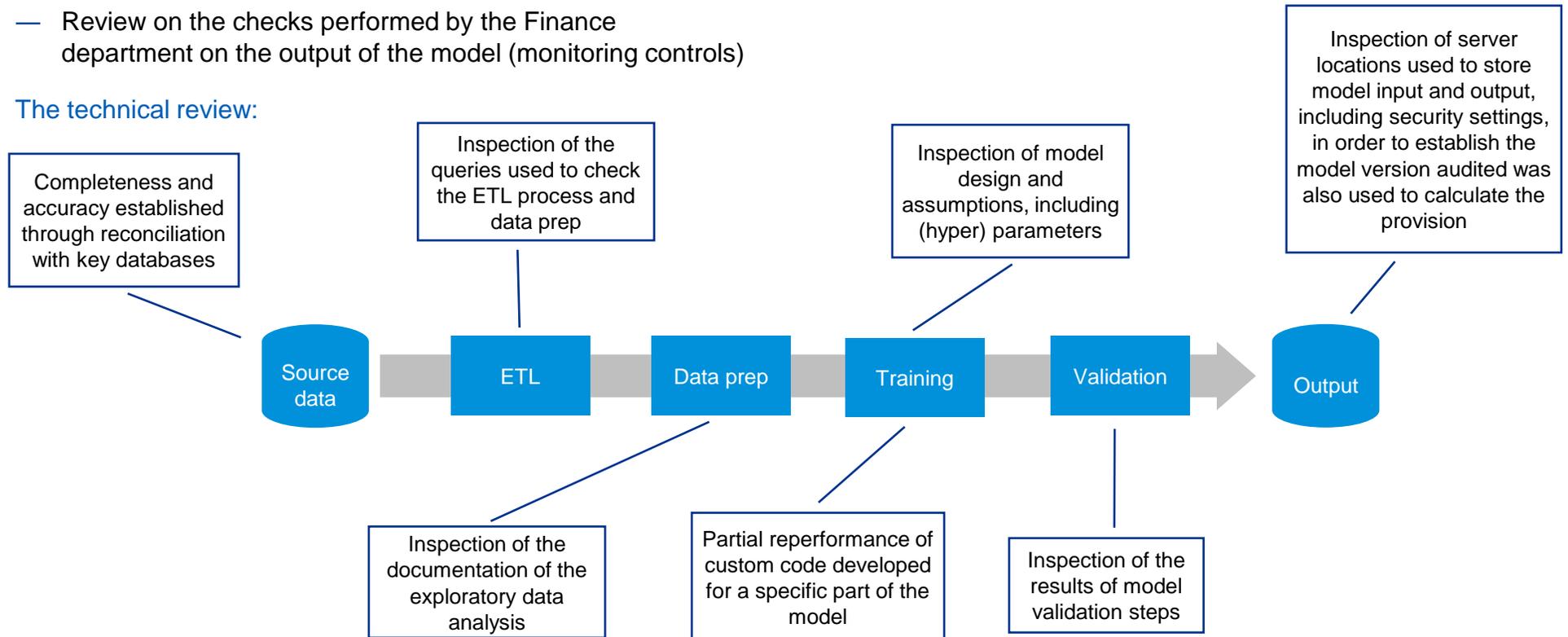
...role of the algorithm in a process drives the selection of audit approaches

An example of a detailed approach of “test the model”

Combined approach:

- Technical review on the design of the algorithm (testing the model)
- Review on the checks performed by the Finance department on the output of the model (monitoring controls)

The technical review:



ID 20190609/007.1215.6

like\followers\subscriptions

Contacts

Frank van Praat, MA MSc RE

Trusted Analytics

vanpraat.frank@kpmg.nl

+31 (0) 6 512 06 152

Mourad Fakirou, MSc

Trusted Analytics

fakirou.mourad@kpmg.nl

+31 (0) 6 223 59 407