SILO.AI

THE LARGEST PRIVATE AI LAB IN THE NORDICS

AI AND **TRADITIONAL ANALYTICS**

This white paper outlines a definition of artificial intelligence (AI) and machine learning and how these differ from traditional analytics. After reading this white paper, you will understand:

1) What AI and machine learning means in practice

2) How using AI is different from rule-based systems, optimisation, simulation and statistical analysis

3) How to implement data driven decision making with machine learning, and

4) What type of use cases can be served with machine learning.

SILO.AI

TABLE OF **CONTENTS**

EXECUTIVE **SUMMARY**

Artificial intelligence (AI) and machine learning have evolved from traditional analytics. The technologies often referred to as AI, such as machine learning, computer vision, and natural language processing, can complement existing analytical methods based on rules, optimisation, simulation, and statistical analysis. In order for companies to leverage machine learning and AI, it is, however, crucial to understand where and when it makes sense to implement AI solutions, and when to leverage traditional analytics. In addition to stand-alone applications, AI can augment existing analytical methods, and vice versa.

This white paper aims to explain when, where and how you should utilise machine learning and AI to support your business processes.

1 WHAT IS AI

AND HOW DOES IT DIFFER FROM TRADITIONAL ANALYTICS

The definitions of artificial intelligence (AI) seem to be almost as plentiful as there are practitioners. With the ongoing hype around AI there seems to be a perceived value in promoting anything as "AI" or machine learning, as these technologies represent "the latest and greatest" in automation and data-driven decision support. This is not helpful when trying to implement these technologies.

AI is not necessarily always better than the pre-existing "traditional methods" for data driven decision support, such as rule-based systems, optimisation, simulation, or statistical analysis, nor will AI displace these methods. AI can complement existing methods and co-exist with them. Therefore, it is important to understand when and where to use any given method to solve the problem at hand.

The goal of this white paper is to provide a definition of AI that puts AI in the context of a bigger set of technologies for data-driven decision support. This will help you understand what AI means for your organisation.

> **"AI can complement traditional analytical methods and co-exist with them.**

DEFINITIONS

Artificial intelligence (AI)
An umbrella term for many technologies for data-driven decision making

Machine learning (ML)
A system that learns without being explicitly programmed

Machine learning model
Set of algorithms, such as neural networks or tree-based models, that permit the learning and adjusting to happen

What is: analytics

To understand AI and machine learning, it is useful to go back to a definition of analytics: a process to find patterns in data to support better decision making.

A further commonly used distinction of analytics is that of descriptive, predictive, and prescriptive analytics, summarised in Table 1 below. The principal difference between descriptive and predictive analytics can be summarised as “what happened?” and “what will happen?”, whereas prescriptive analytics is attempting to provide a course of action to facilitate a desired outcome. On a general level, predictive analytics often makes use of statistical modelling and forecasting techniques, whereas prescriptive analytics is frequently based on optimisation, heuristics (a trial-and-error approach to finding a solution), and rule-based systems (“if this, then that”). Simulation is frequently used for both predictive and prescriptive analytics.

TYPE OF ANALYTICS	DESCRIPTIVE & PREDICTIVE	PRESCRIPTIVE
Question to answer	What happened? What will happen?	How to get to a desired outcome?
Tools	Statistical modelling Forecasting Simulation	Optimisation Heuristics Rule-based systems Simulation

TABLE 1. Different types of traditional analytics

Relationship between analytics and AI

One way to think of AI is to see it as a set of four activities to achieve an outcome. The outcome is typically solving the problem at hand, e.g., recommending a course of action, finding a pattern in data, automating a task, or winning a game.

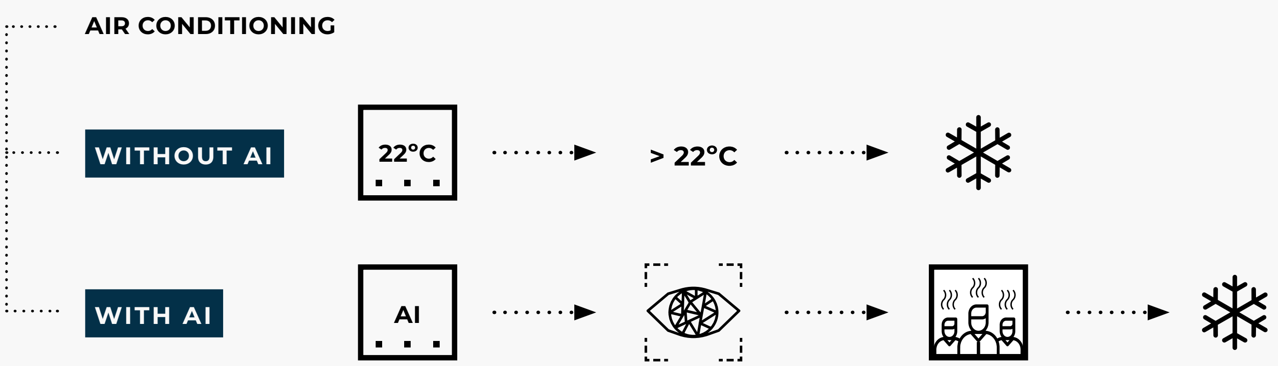
THE FOUR ACTIVITIES OF ARTIFICIAL INTELLIGENCE ARE:

Sensing, Reasoning, Reacting, Learning

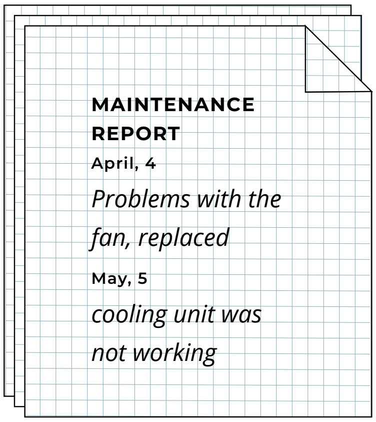
When looking at AI from this perspective, the first activity, **sensing**, and the fourth activity, **learning**, are of particular interest. It is these two activities that differentiate AI from “traditional analytics”, as the second and third activity, **reasoning** and **reacting**, are essentially the outcome of all the previously mentioned types of analytics: descriptive, predictive, and prescriptive analytics (see Table 1 on the previous page).

One simple example of this type of reasoning and reacting could be a rule-based system to regulate indoor climate. A rule requires that air conditioning should be switched on when certain temperature thresholds are exceeded. This is not necessarily very sophisticated reasoning and reacting, but it illustrates a principle that can be extended to any analytics approach.

However, modern AI technologies allow for entirely new ways of sensing, that were unavailable to us until fairly recently. The two most prominent categories of “AI sensing” are computer vision (CV) and natural language processing (NLP). In our climate control example, the rule that turns on the air conditioning relies on sensors that detect a temperature. However, a computer vision based “AI climate control system” could detect who, and how many people are in the room, and adjust cooling accordingly.



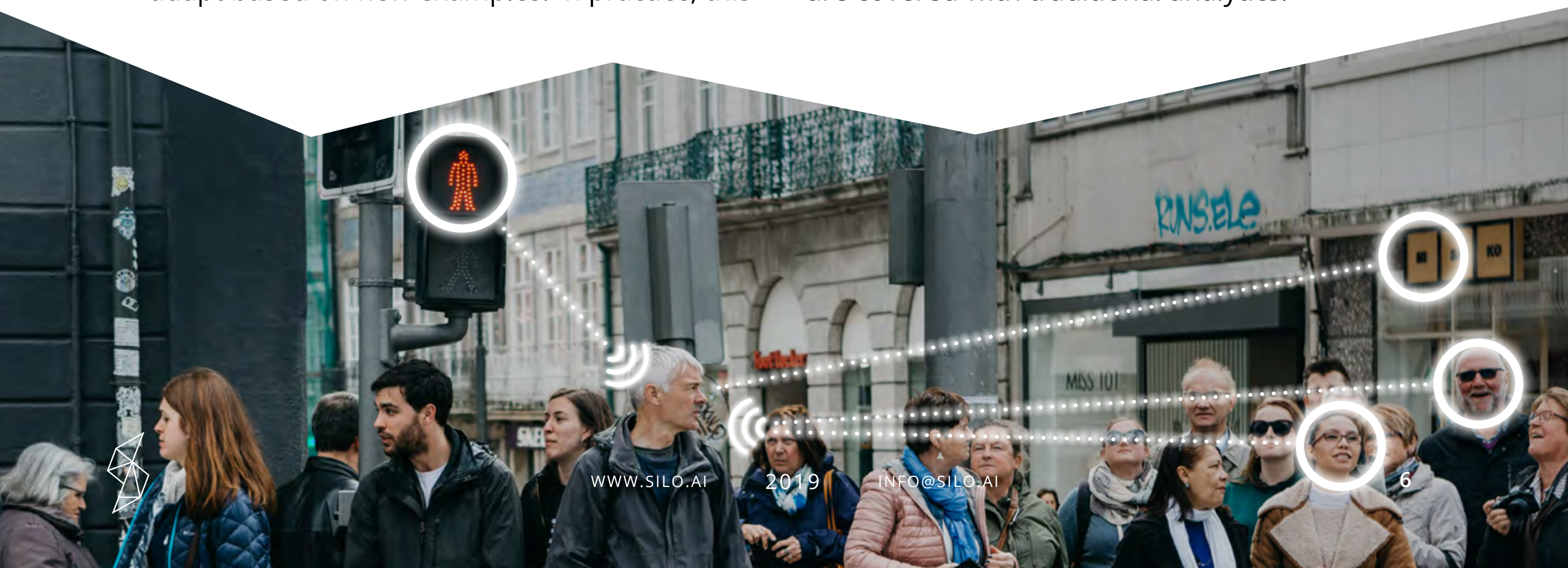
Natural language processing (NLP) on the other hand, could allow the utilisation of free text maintenance reports to steer how the climate control system works. It picks up on particular ways of reporting, despite different words (or even languages) used for reporting similar phenomena. This way, computer vision and natural language processing allows us to process data that previously could only be understood by humans.



The fourth activity, **learning**, is the other part that sets AI apart from traditional analytics. While this aspect of AI, commonly referred to as machine learning (ML) can be used for both predictive and prescriptive analytics (more on that below), one aspect in particular sets it apart: the ability for the algorithm to adapt without specifically being reprogrammed.

In a context of a rule-based system, such as our climate-control system, ML enables automatic rule re-definition. Machine learning implies learning from examples and, as such, when conditions change, the ML model can adapt based on new examples. In practice, this could mean that a ML model can decide on its own that more people in a room require more cooling.

To summarise, artificial intelligence implies a new way of sensing data that previously could not be automatically processed, specifically relating to **visual data** in the form of images or video and **textual data** without any predefined logical structure, other than that of spoken language. In addition, AI has an ability to adapt and learn without being specifically re-programmed. Sensing and learning are thus additions to the ability to reason and react that are covered with traditional analytics.



2 APPLYING **MACHINE LEARNING AND TRADITIONAL ANALYTICS**

Machine learning (ML) is at the core of any modern AI application. While ML can be used for both predictive and prescriptive analytics, there is one key aspect of it that sets it apart from traditional methods; it learns from examples.

THE USE OF MACHINE LEARNING IMPLIES THAT

- **LITTLE OR NO PRIOR KNOWLEDGE OF CORRELATIONS IN THE DATA IS REQUIRED. You only need a sufficient amount of examples to learn from.**

- **NO RE-PROGRAMMING IS NEEDED IF CONDITIONS CHANGE.**

- **DATA HAS TO BE 'LABELLED' TO BE USEFUL FOR MACHINE LEARNING.**

- **NEARLY ANY TYPE OF DATA (IMAGE, VIDEO OR TEXT) CAN BE USED AS EXAMPLES.**

In addition to the aspects covered above, often referred to as **supervised learning**, more and more focus is put on so called **reinforcement learning**. The main difference between reinforcement learning and supervised learning is covered in the last chapter. We will also explain how reinforcement learning can add an additional layer of intelligence to solving the problem, given that certain prerequisites are met.

DEFINITIONS

Supervised learning

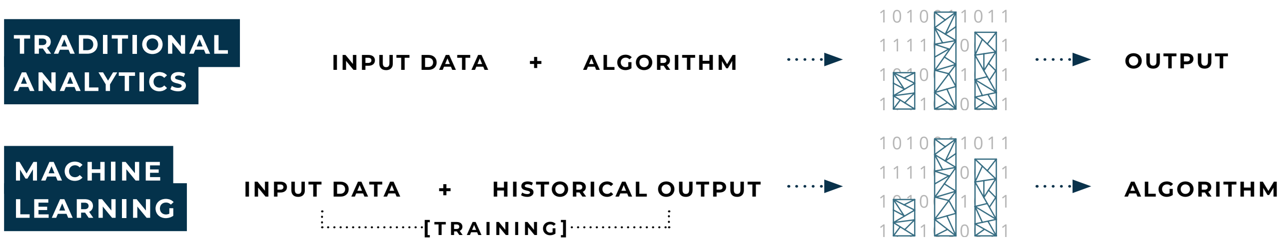
A system that learns from examples

Reinforcement learning

A system that learns based on taking a given action

Understanding data

Machine learning, traditional analytics, and the need for prior knowledge regarding correlations in data can be illustrated with the following simple diagram:



As can be seen above, the key difference between ML and traditional analytics is in whether we define how data is analysed and processed, or whether the computer suggests correlations in data.

Another way of looking at the difference is through how a conclusion is reached. In ML, we look at general observations to make a prediction or recommend a course of action, also referred to as **inductive reasoning**.

Traditional analytics, on the other hand, frequently uses **deductive reasoning** to reach a conclusion. In practice, the starting point is general knowledge about the problem, further analysis helps us reach a conclusion.

Traditional analytics requires some prior understanding of the problem as a basis for our analysis. In practice, it means that you have to frame the problem. One implication of this is that if conditions change, they also require that we “reframe” the problem. We adjust our hypothesis, add or change constraints in an optimisation model, or similar.

ML behaves slightly differently, as ML models will detect that conditions have changed, and automatically adapt accordingly. The ML model “sees” this in the examples we feed to it for training. This is the ‘learning’ aspect described in chapter 1.

Where to apply machine learning

When and where does it make sense to apply ML as opposed to traditional analytics? On a general level, ML is very useful when there are very complex relationships in the data, or when the input data consists of images, videos, or textual data in a non-standardised format (referring back to 'sensing' in chapter 1).

MACHINE LEARNING IN SUPPLY CHAIN MANAGEMENT, AN EXAMPLE

There are also examples where traditional analytics and ML can be used to complement each other. For example, in supply chain management, there is one specific area that illustrates this point well – demand forecasting. Typically, statistical forecasting is used to create a plan for a "baseline demand". In other words, we predict what future demand will look like based on historical demand. While ML could be used to do this as well, framing the problem is achievable and feasible with statistics.

However, if we introduce more complex relationships to the data, where we do not necessarily have a clear picture of their impact on demand, ML starts to shine. This could include signals from social media (people discussing your product in a positive or negative manner), competitor pricing, or competitor campaigns, as well as your own campaign activities. Campaign activities typically differ a lot, including adverts in social media, TV, or newspapers.

Price campaigns is another example where it can be difficult to estimate impact on demand. In this case, and in many other cases, the challenge is typically to find relevant data to teach and train the ML model.

Where not to apply machine learning

When is machine learning very difficult to apply? A “prime suspect” would be a case where there are too few examples for the ML model to learn from. However, the question of what is a sufficient amount of examples is not a trivial one. Unfortunately, the most accurate answer to the question is that “it depends”.

HOW MUCH TRAINING DATA IS NEEDED DEPENDS ON:

- **COMPLEXITY OF THE PROBLEM**
- **PREDICTIVE ABILITY OF THE EXAMPLES**
- **NUMBER OF VARIABLES**

In practice, the best way to find out if you have enough training data, is to test the ability of ML with the data you have. In addition, you can address possible shortcomings with traditional methods. Compared to ML, statistics can generate results with fairly small amounts of data. At the same time, results from statistical analysis can provide us with knowledge of the phenomenon, that can be used to enrich the ML models.

Another possible shortcoming with ML is that models can “find” patterns where there are no patterns to be found. One such field is in finance, where the predictive ability of examples can be low. As we all know, and as many investment funds are very keen to point out, previous market developments are not an indicator of future performance.

As you can see from above, there are differences in how you use data to reach a conclusion with traditional analytics and with ML. Both approaches have strengths and weaknesses, and they can also be used to complement each other.

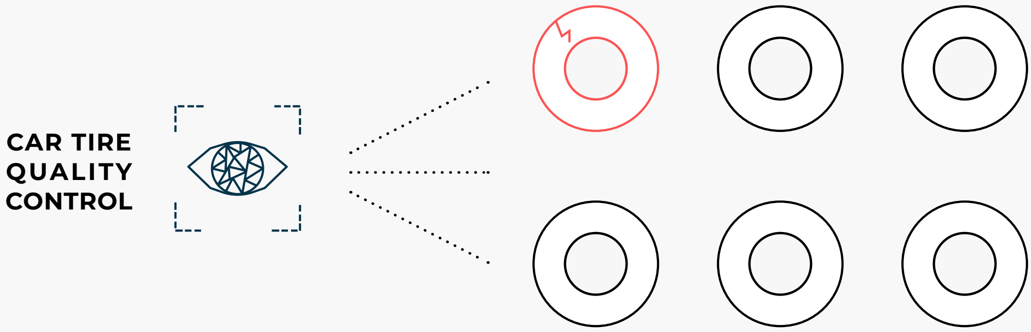
Labelled data

As we have discussed, machine learning requires examples for learning. The amount of examples needs to be sufficient, but there is an added complexity to consider. Data needs to be 'labelled' to be useful as an example for the ML model.

To understand what this means in practice, consider the phenomenon you are trying to predict. Sometimes it is a straightforward task to identify this from data; let's say you are trying to understand quality defects, and you have a clear historical record of when these defects have occurred.

However, sometimes data might be abundant, but there is no systematic way of distinguishing the phenomenon you are predicting from the dataset. In practice, there is no clear example to learn from.

Let's say we want to train an ML model to identify a particular object in a picture. For us to be able to train the model to do this, we need examples of pictures where the object in question can be found. These pictures need to be labelled with metadata indicating where in the picture the object is found. For certain objects, there are openly available datasets for training, but let's say we are looking at automatic detection of particular quality defects in your manufacturing; this is a much more specific and unique problem. In all likelihood, either you don't have a readily labelled dataset available, or then it has been created through manual effort. In this case, an operator responsible for quality assurance has tagged a picture or video with a particular defect type.

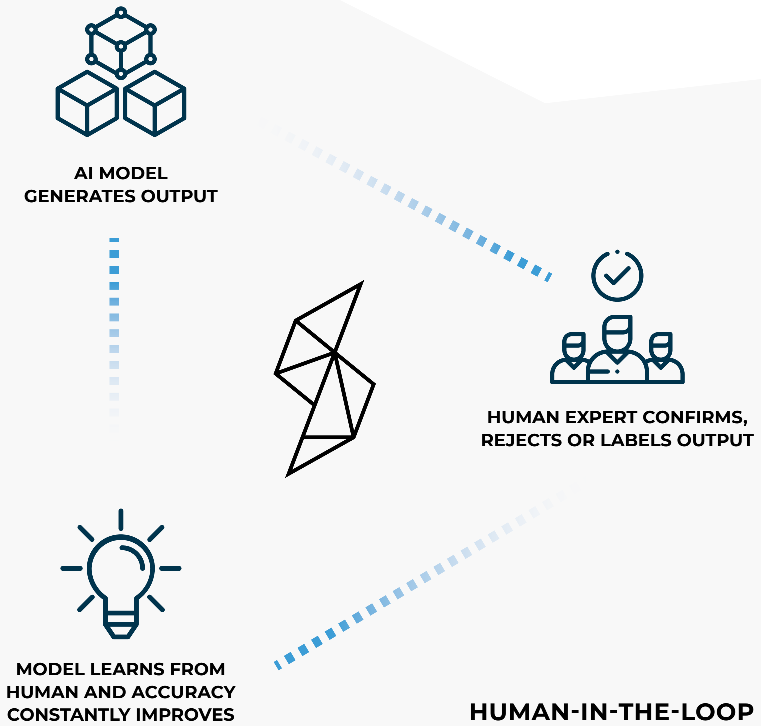


A HUMAN-IN-THE-LOOP APPROACH TO AI

The above example concerns computer vision (CV), but it is important to recognise that labelling is not restricted to visual data, nor is it restricted to labelling of training datasets. ML models will not produce perfect results; in order to steer the learning in the right direction, it is still important to involve a human operator to correct for inaccuracies. Adding human intelligence typically improves model performance, and it can also open up for capturing tacit knowledge for use by the ML model. This is described further in the Human-in-the-loop image below.

CV is an example of where labelled datasets can be difficult to obtain. The same applies for free text, natural language processing (NLP). As indicated in the definition of AI, both NLP and CV process data in ways that were difficult or impossible to do with traditional analytics. This might explain why metadata and labels are scarce in these datasets.

While modern AI provides us with opportunities to use new types of data, it is not necessarily a silver bullet for new insights, unless the text, image or video is labelled with the relevant metadata. It is also good to keep in mind that you still need a sufficient amount of examples to learn from.



Reinforcement learning

Machine learning can be divided into supervised learning, reinforcement learning, and unsupervised learning. So far, we've focused on supervised learning, which accounts for the vast majority of current ML implementations. However, there is a growing intrest in applying reinforcement learning to solve real-world problems.

The key difference between supervised learning and reinforcement learning is that reinforcement learning models learn through iteration when trying to maximize a pre-defined reward. Unlike supervised learning, there is no need for examples, as the reinforcement learning model will test alternative solutions to the problem on its own. In essence, it is creating its own training dataset through the evaluation of different approaches. This obviously sounds very good; no training data is needed.

Why then isn't reinforcement learning used to a larger extent? In its quest to learn, the reinforcement learning model will test all options available. Many of these could have unwanted consequences if tested in a real-world context. If a reinforcement learning model would, for example, steer a car, it would crash many times in the process of learning to drive. As such, the only way to safely train a reinforcement learning model is in a simulation environment, or another similarly contained environment. Examples of such contained environments are games, which is why reinforcement learning has indeed been used to beat human players in games such as Go. No harm is done if and when the reinforcement learning model loses a game. The model starts over, learns, and improves. Rules of the games are relatively easy to model, but more complex systems require a simulation environment with sufficient detail to allow for proper training of a reinforcement learning model.

This is also why reinforcement learning applications are fairly rare compared to supervised learning. If, however, you do have a simulation environment or a so called "digital twin", reinforcement learning is something worth looking into.

3 SUMMARY AND **CONCLUSION**

In this white paper, we summarised AI as a system capable of sensing, reasoning, reacting, and learning. Sensing and learning are of particular interest as they set AI apart from traditional analytics. Sensing implies that the AI system can process visual and textual data in new ways. The learning aspect refers to machine learning, which perhaps is at the "core" of any modern AI application. At the same time, machine learning extends to the reasoning and reacting part, providing us with new ways of understanding patterns in data to drive better decision making and automation.

To get started on a journey towards making use of your data for competitive advantage, it is important to answer two questions:

1) WHAT ARE THE BEST SUITED METHODS FOR DATA-DRIVEN DECISION SUPPORT AND AUTOMATION IN YOUR BUSINESS?

2) WHAT DATA DO YOU HAVE READILY AVAILABLE, AND WHAT PRIOR WORK HAS BEEN DONE TO UTILISE THIS DATA?

Putting this in a slightly different context; start with understanding the problem. What is it that you want to achieve? Can AI help with solving the problem, and should it be complemented with traditional analytics? What data do you have available to help in solving the problem, and has this data already been analysed in some way?

> **"The first step in getting started is often to ask the right questions. Hopefully this guide has given you insight to help in finding answers to the above questions, and to get started with implementing AI in your business.**
>
> **– Henrik Nyman**

If you believe that the questions above deserve a more in-depth discussion, get in touch with **Henrik Nyman (henrik.nyman@silo.ai), Head of Operations at Silo.AI.**

Get in touch to discuss how AI can help your business.

Henrik Nyman
HEAD OF OPERATIONS
henrik.nyman@silo.ai

Silo.AI
www.silo.ai
+358 40 359 1299
info@silo.ai

Silo.AI is the largest private Artificial Intelligence lab in the Nordics that builds AI as a service.

SILO.AI
THE LARGEST PRIVATE AI LAB IN THE NORDICS