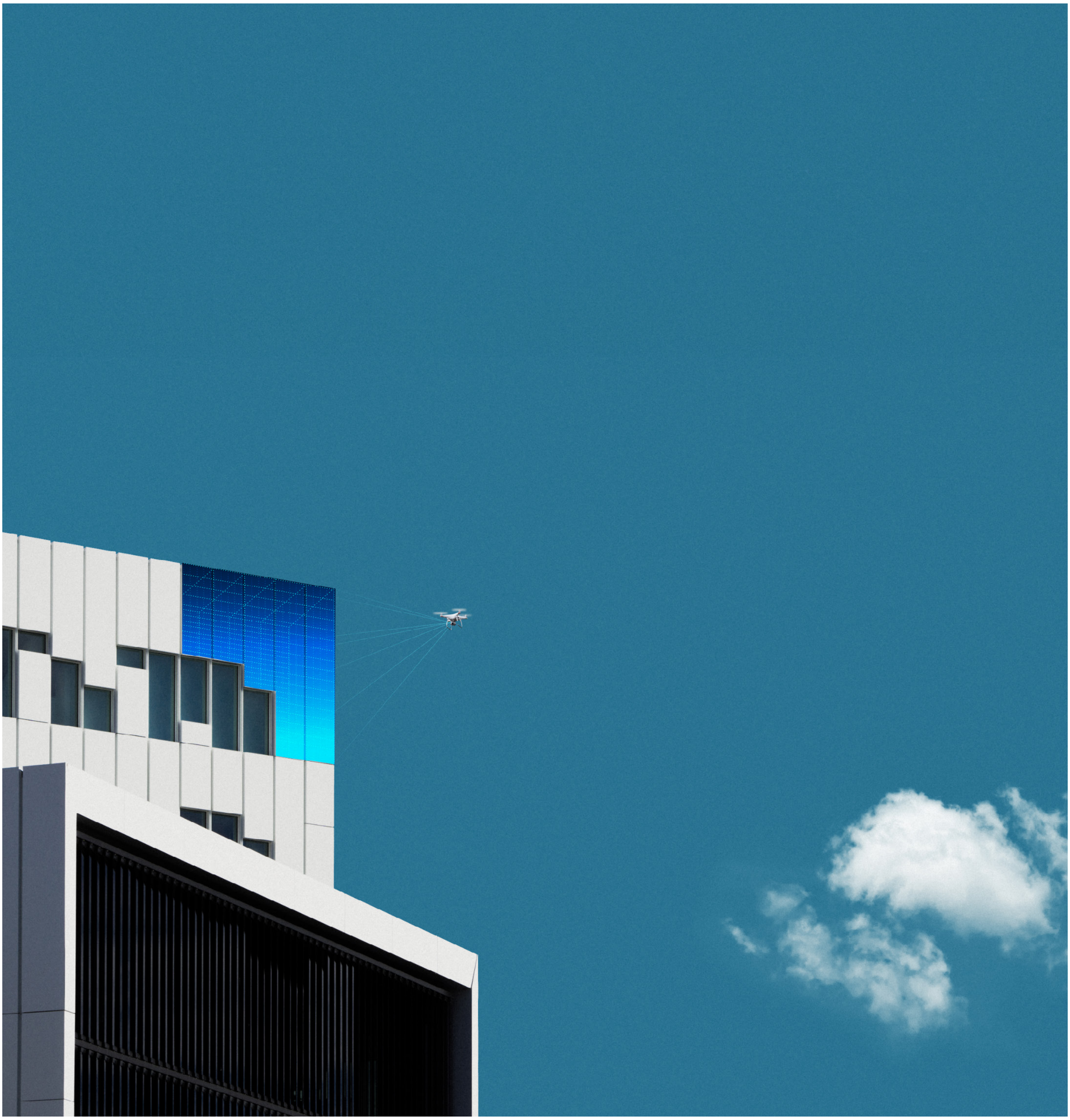
AI for Civil Engineers:

What You Need to Know to Build the Future



This whitepaper introduces civil engineers to the key concepts of AI, equipping them with the tools to identify AI opportunities in the built environment, assess potential solutions, and make informed, actionable decisions to drive transformative change within their organisations.

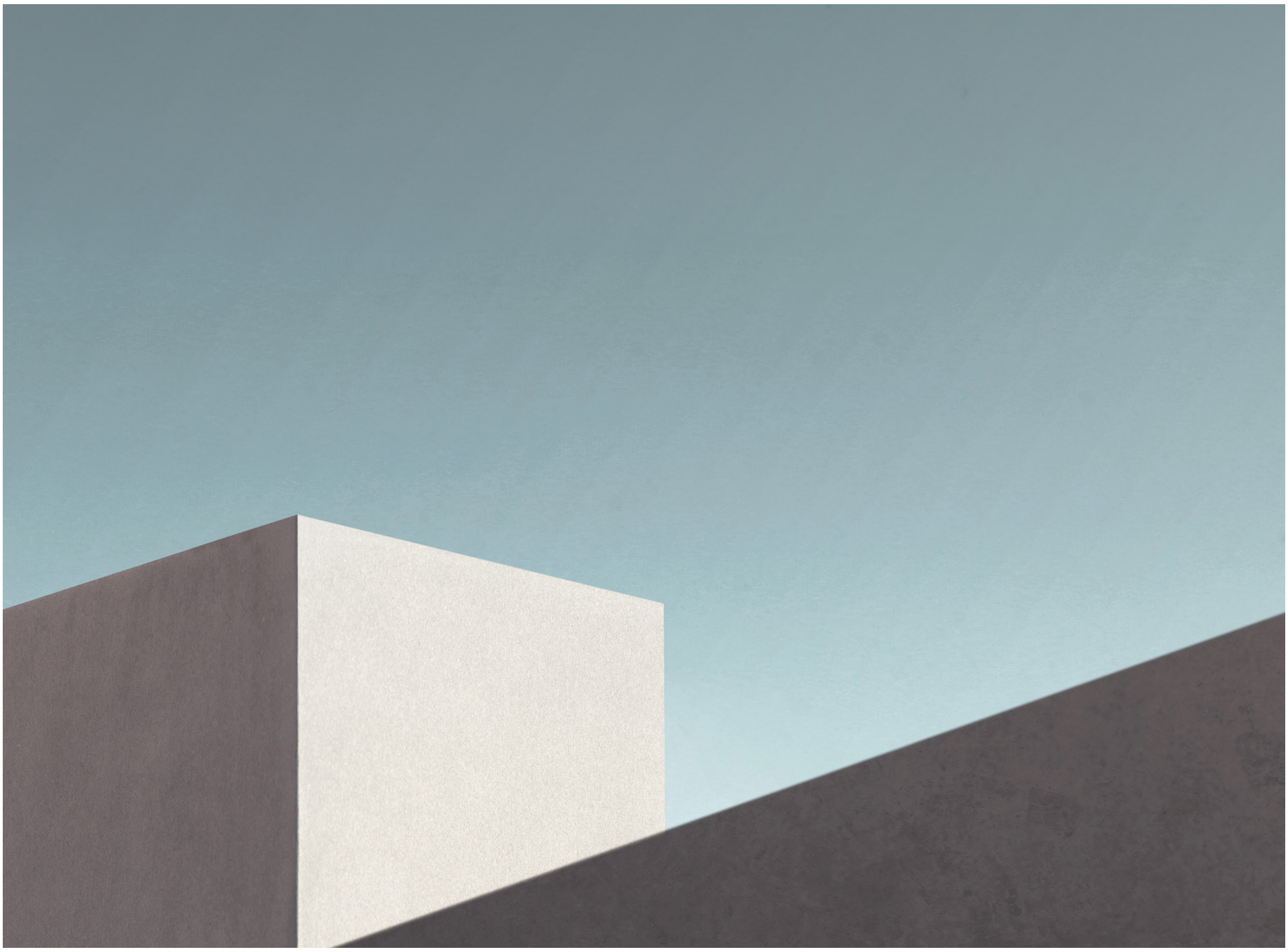


Table of Contents

Introduction

The field of Artificial Intelligence is complex, and considering how it can be applied to civil engineering can feel daunting. Large Language Models and Generative AI may be at the forefront of conversations around AI today, but as we'll explore in this whitepaper, the true potential for AI in civil infrastructure lies in other types of models and algorithms.

Fast forward to today and the development of Large Language Models and Generative AI has taken centre stage thanks to their easy and widespread access. The business value behind these models, however, is still being explored. At the same time, conversations around ethical and responsible AI, and the challenge of ensuring an AI system is aligned with human values have become serious fields of research.

A brief history

Artificial Intelligence (AI) has evolved dramatically since Alan Turing's "universal Turing machine", the first theoretical model capable of solving any problem with an algorithmic solution, was created in the 1940s. AI progressed at incredible speed over the next two decades, moving from conceptual and theoretical foundations to real algorithms and programs which could, for example, converse in limited English and play checkers.

The term "Artificial Intelligence" was coined during this period, but soon afterwards, AI entered its first "winter" due to disappointment in the lack of real-world applications. Government funding and academic research dropped significantly in the field until the 1990s. The 90s saw a revitalisation of AI thanks to the development of new powerful algorithms (such as deep learning algorithms), significant hardware improvements (e.g. cloud computing), and access to larger amounts of data. The creation of ImageNet (a massive database for object recognition) in 2012 and the defeat of a world champion Go player, by AlphaGo (a computer program developed by Google subsidiary DeepMind) in 2016 represented significant developments in the field of AI.

The Bombe device used by codebreakers during World War II.

AI Uncovered: The Essential Guide to Its Key Types and Uses

There is a wide range of AI models and algorithms, from decision trees to clustering models to deep learning. Understanding the differences between each of them and figuring out what they are best suited for can be a challenge.

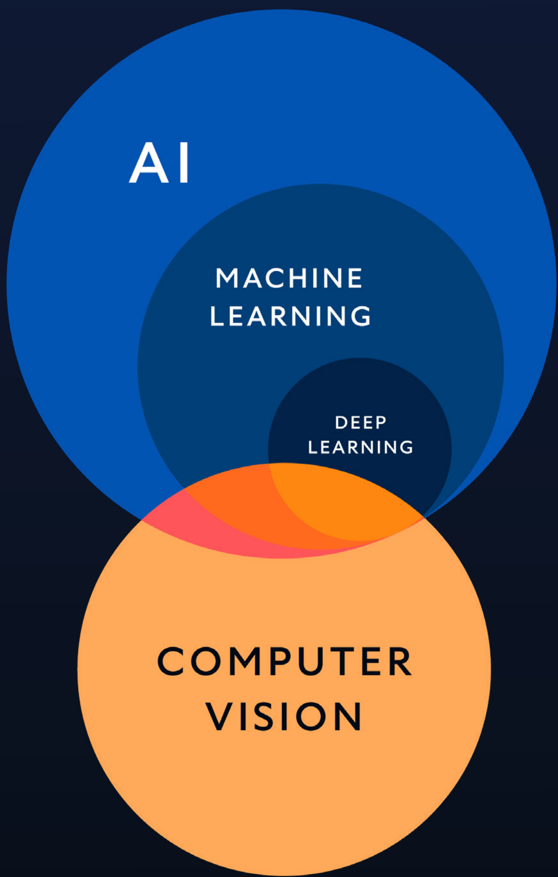
This first section of the whitepaper gives an overview of the main types of models a civil engineer should be aware of.

Understanding these models, algorithms, and techniques, and beginning to think of how they could be applied in the civil engineering context, will give you the knowledge and vocabulary to start enacting real AI-powered digital transformation on your projects by:

- Understanding the capabilities and limitations of different models and algorithms.
- Framing how different models and algorithms could solve different use cases.
- Facilitating informed discussions with data scientists, machine learning scientists, and software engineers when building and deploying AI applications.

Inside the world of AI

AI is a field of computer science in which machines perform tasks that had previously been possible exclusively by human intelligence. However, the term AI is not well-defined and can cover anything from simple rule-based systems to complex deep learning models. The diagram below provides a high-level overview of some different types of AI, and this section will outline how they can be used.



Rule-based systems

Rule-based systems run on predefined rules to make decisions and usually follow "if-then" statements. They are a basic type of model that uses logical reasoning and decision-making based on predefined rules and conditions. These predefined rules can be provided by a subject expert to create what is known as an Expert System. They are optimal for repetitive, rule-based tasks where information, or data, is collected and evaluated against expert rules to achieve some desired outcome.

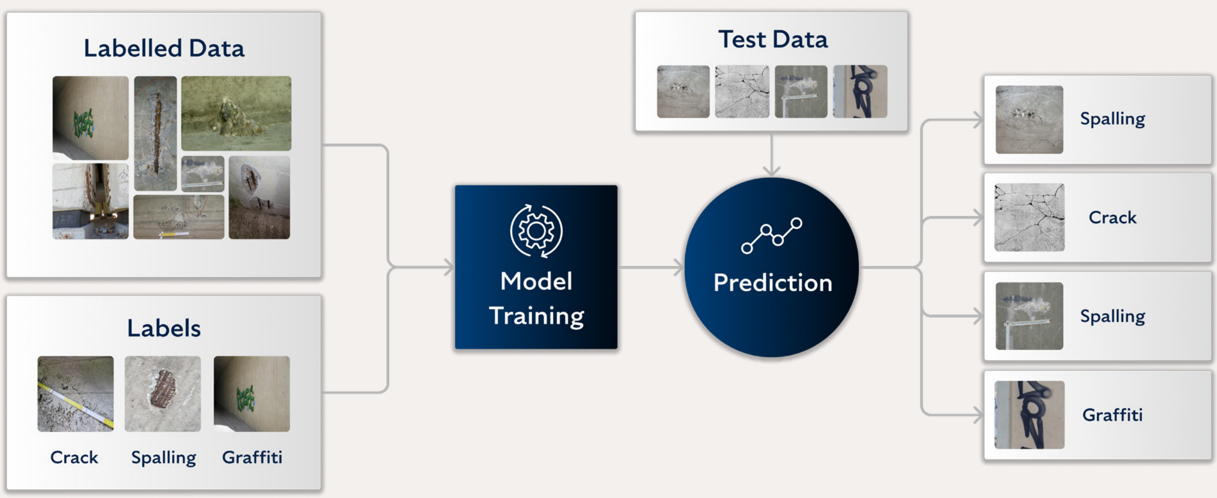
For example, a council might want to optimise road closure times for infrastructure repair. A subject expert would then provide specific rules (e.g. repairs cannot happen between 8 am and 8 pm, and if the temperature is below 2°C; if sensors installed on a bridge detect unusual frequency responses, then send an inspector within 24 hours) and the model would output repairs recommendations optimised to meet the rules.

Machine learning

Machine learning refers to systems which are able to learn and improve from relevant datasets without being explicitly programmed. Machine learning is a subset of artificial intelligence that involves training algorithms to recognise patterns and make decisions based on data. Instead of being explicitly programmed, machine learning models learn from examples and improve their performance over time. The output of a machine learning model is a prediction of what it understands to be the correct output. It's an educated guess - the better the training and data, the better the educated guess is. The advantage is that the computer can spot patterns and relationships in data that humans can't always see and can deal with vast data sets to detect even small signals in noisy data. Machine learning consists of three main learning paradigms: supervised learning, unsupervised learning, and reinforcement learning. These concepts will be explored throughout the following sections.

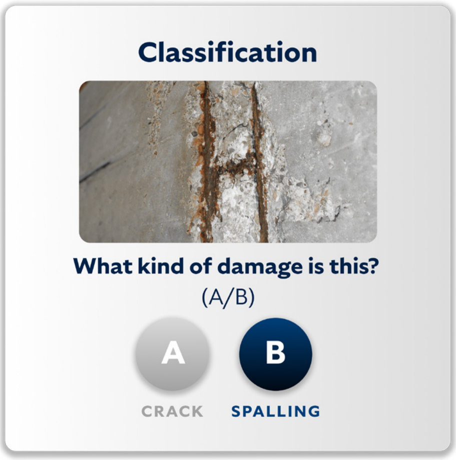
Supervised learning

Supervised learning is a type of machine learning where a model is trained using labelled data, meaning each input comes with a corresponding output. The model learns to map inputs to outputs by finding patterns in the training data, allowing it to make predictions for new, unseen data. This approach is commonly used in tasks such as classification and regression, where accurate examples are available for model training. As a result, supervised learning is beneficial when accurate predictions are required.

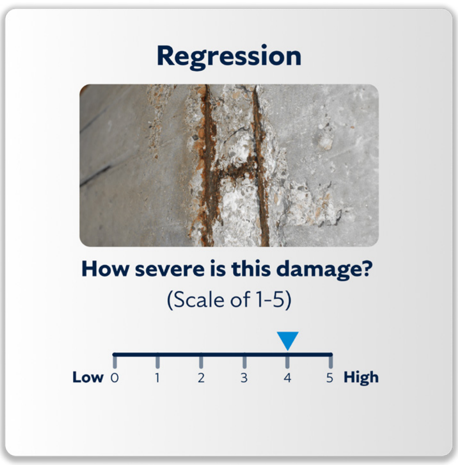


Common supervised learning models include:

Classification: Classification algorithms predict discrete outputs, often referred to as a class label. Models created using classification algorithms exploit the relationship between the data describing an entity, or instance, and the class label associated with that entity. The relationship can then be used to predict the correct label of a given input data with an unknown label.

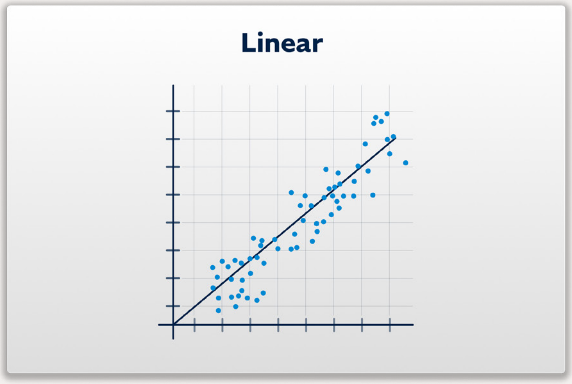


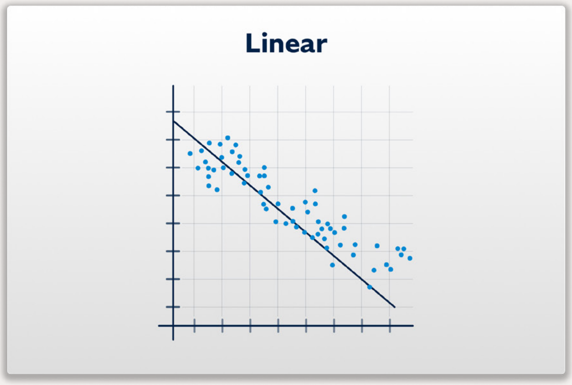
Regression: Regression algorithms predict continuous outputs. Models created using regression algorithms learn a function that describes the relationship between one or more independent variables in the input data and a response, dependent, or target variable.

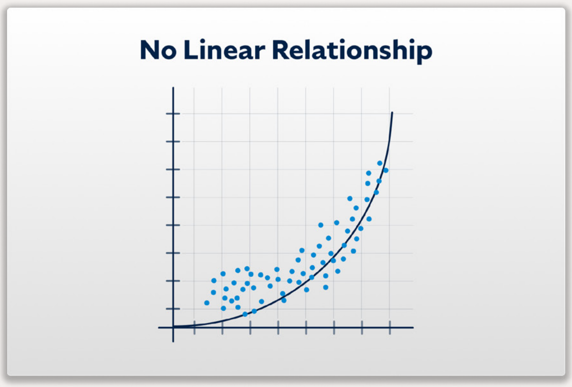


Commonly used types of supervised machine learning algorithms include:

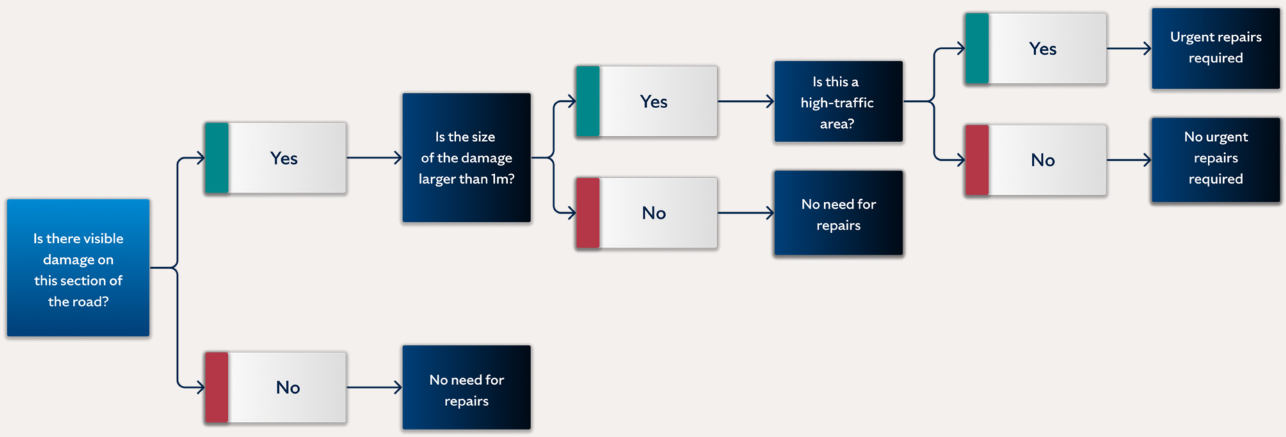
Linear models: Linear models are the simplest type of machine learning model. They search for the best-fitting line through a set of data points.



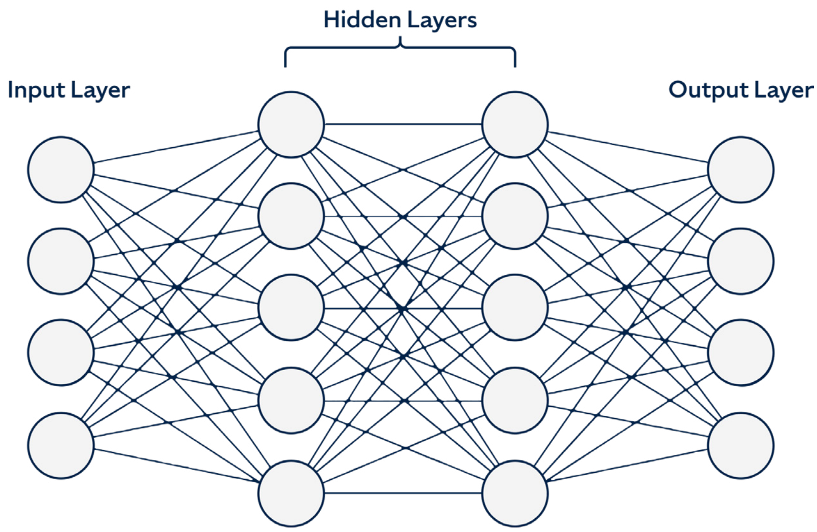




Decision trees: Decision trees are a non-parametric supervised learning algorithm which is utilised for both classification and regression tasks. They have a hierarchical structure, which consists of a root node, branches, internal nodes and leaf nodes. Decision trees are similar to expert systems in that they can be interpreted as if-then rules that result in a final outcome, however, unlike rule-based systems, decision trees learn the rules from labelled data rather than depending on an expert to define the rules in advance. Below is an example of a simple decision tree.



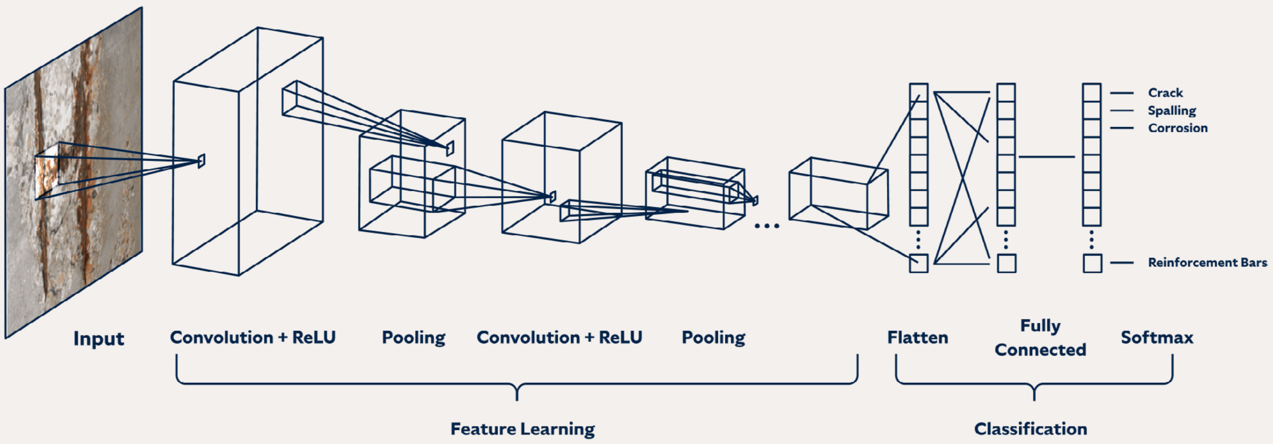
Neural networks: A neural network is a computational model inspired by the structure and functioning of the human brain, consisting of layers of interconnected nodes (neurons). It is designed to recognise patterns, learn from data, and make predictions or classifications, making it a key building block in many machine learning and deep learning applications. Neural networks are highly effective for tasks like image recognition, speech processing, and natural language understanding, where complex relationships in the data need to be captured.



Deep neural networks: Deep learning refers to a particular type of modelling within machine learning whereby multi-layered neural networks are used to solve tasks. These deep neural networks can automatically learn representations from raw data, making them highly effective for tasks such as image recognition, natural language processing, and speech analysis. They often end up needing to have millions of parameters, and as such, they are not easily interpretable. Deep learning models have driven significant advancements in AI due to their ability to handle high-dimensional data and achieve state-of-the-art results in various applications.

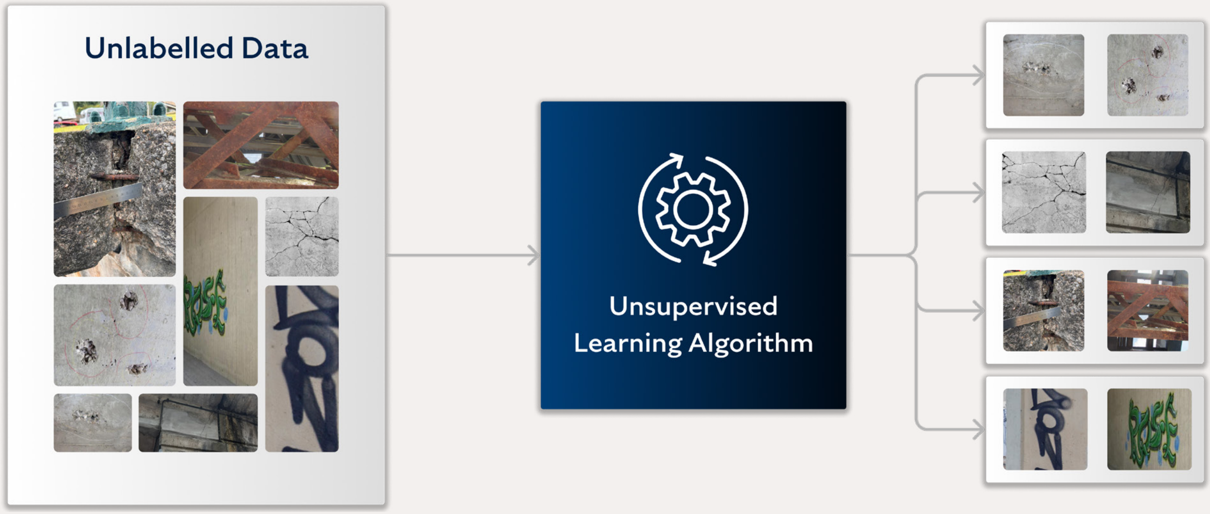
Time series forecasting: Time series forecasting is the process of using historical data to predict future values, where time is a key feature of the data. The model will output a time-based prediction for a given time stamp. For example, based on the history of the average daily price of steel and additional market indicators over the past year, this type of model could be used to estimate the steel price for a specific day in the future.

Convolutional neural networks: Convolutional neural networks are an example of deep neural networks. Convolutional neural networks are a type of artificial neural network for deep learning that are often used for image and video-based learning, where spatial relationships between pixels in images can be used to identify patterns, known as feature learning, making them highly effective at tasks such as object detection and classification. For example, a convolutional neural network could be built to identify different defect types - such as cracks, corrosion, spalling, and reinforcement bars - identified on various assets and prioritise maintenance of these assets based on the defect types and their anticipated deterioration pathways.



Unsupervised learning

Unsupervised learning models are trained on unlabelled data, so the model itself finds patterns in the data.



An example of unsupervised learning is clustering models.

Clustering models: Clustering models are a machine learning technique that identifies groups of similar records. Clustering techniques aim to find similarities within data to identify different groups, or "clusters", that exhibit different characteristics without requiring classification labels.

The characteristics that separate data belonging to different clusters can be further examined to determine how they differ and can be used as a proxy for classification labels. For example, a clustering model could be used to cluster a set of unlabelled defect images into groups of images with similar defects. Using an approach like this could separate images with cracks from images with spalling.

Despite using unlabelled data, the algorithm will be able to learn the difference between different desired labels, but it will take longer to do so and will require much more data than, for example, supervised learning, which uses labelled data.

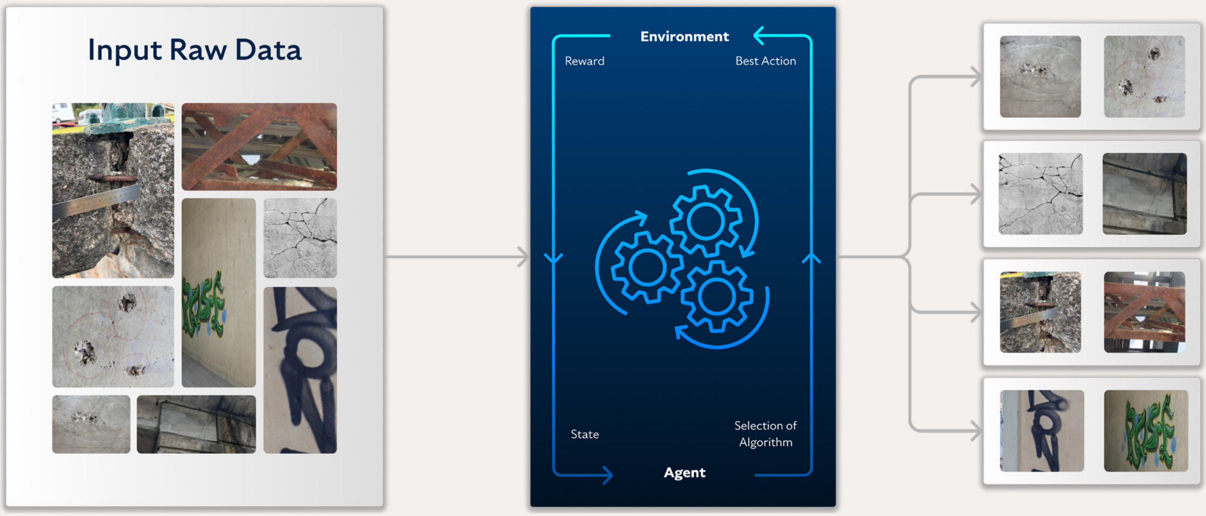
Generative AI and Large Language Models

Generative AI and Large Language Models are a subset of deep learning models. Generative AI refers to a methodology which learns relationships across disparate data domains and the concepts underneath them from incredibly large, broad sets of data. Generative AI can create a wide variety of outputs, such as images, text, audio, and datasets. Large Language Models are very similar, however they specifically generate natural language. More generally, generative models learn the patterns and structure of data and then create new data with similar characteristics. Generative AI and Large Language Models have recently exploded in popularity thanks to models like OpenAI's Dall-E and GPT-4, Google's Gemini, and Mistral AI's Mixtral.

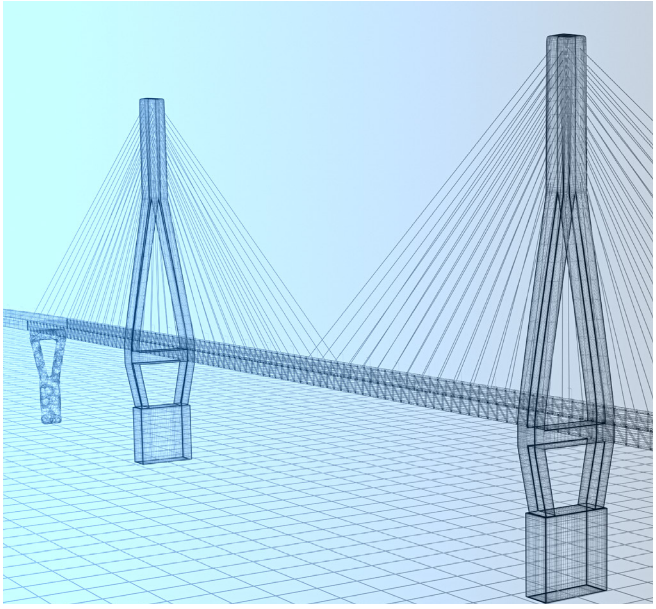
While impressive, the true value proposition and range of use cases of these models are still very unclear, so they should not be considered catch-all solutions. These models usually remain too generic to generate real business value, and their factual inaccuracies, or "hallucinations", do not inspire confidence in a high-stakes environment like civil infrastructure.

Reinforcement learning
Reinforcement learning models use techniques that train software to make decisions to achieve the most optimal results. They mimic the trial-and-error learning process that humans use to achieve their goals. They can be used to determine a sequence of actions to achieve a goal or desired outcome. The model will consider different actions and receive rewards for actions based on how close it is to achieving the goal or output. The model uses this reward system to determine how optimal each action is with respect to the desired objective. As such, the model can iteratively improve its behaviour. The reward is sometimes a human providing feedback on the model output (saying whether the output is correct or not).

For example, a reinforcement learning model has been deployed to forecast optimal bridge maintenance schedules. The model's reward system is looking to minimise the risk of bridge failure due to insufficient maintenance, while lowering the cost of maintenance. The users of that model will then provide feedback back to the model based on the bridge's true condition following inspections suggested by the model itself.



Preparing data for machine learning



In the previous section, concepts around “labelled” and “unlabelled” data were discussed in the context of machine learning models and algorithms. This section details these concepts to further explore the different characteristics to look out for in data when building machine learning models. Fully understanding the characteristics and limitations of the available or required data is important because data often dictates or limits what use cases can or cannot be solved, which algorithms can or cannot be used, and what alternatives exist.

Throughout this section, data refers to the data used to train a model. Part of that data is reserved to test and validate the performance of that model once it has been built. Finally, new data is fed into the AI model once it is deployed into production to receive an output.

Structured and unstructured data
Data comes in a variety of formats, ranging from images to tabular data to natural language. Structured data refers to data in a standardised format, like tabular data. On the other hand, unstructured data does not follow a standardised data format and includes images, videos, audio, and natural language, for example. Unsurprisingly, unstructured data usually requires more pre-processing than structured data to make it understandable to the machine learning model.

Labelled and unlabelled data
The concepts of “labelled” and “unlabelled” data dictate what can be done with it and what type of algorithm can be used. Reusing the terminology learned in the previous section, supervised learning requires labelled data, while unsupervised learning requires unlabelled data.

A label is a piece of information that indicates what a particular data point represents. Labelled data consists of input data paired with the correct output or classification. For example, in a dataset of images of buildings, each image might have a label indicating whether it is in need of repair or not. This labelling provides examples for the machine learning algorithm indicating what it is supposed to learn from the data. It’s important that the labels are relevant to the task the algorithm needs to perform. If the labels are unclear or not useful, the algorithm won’t be able to learn effectively, and the data might as well be unlabelled, as it won’t provide any guidance for making predictions.

If the data is labelled, the machine learning algorithm will find the relationship between the characteristics of the labelled data so that those relationships can be used to determine the label for unlabelled data. However, if none of the data is labelled, the algorithm must find similarities among the features to identify different groupings in the data.

Example 1: labelling unstructured image data

Figure 1 below presents two images: one of a crack and one of spalling. The top image is labelled and the bottom one is unlabelled. Both of these sets of images contain data about some concrete, however, the label of 'crack' or 'spalling' provides additional information about what is being observed in the image. Additionally, since this data is made of images, it is unstructured.

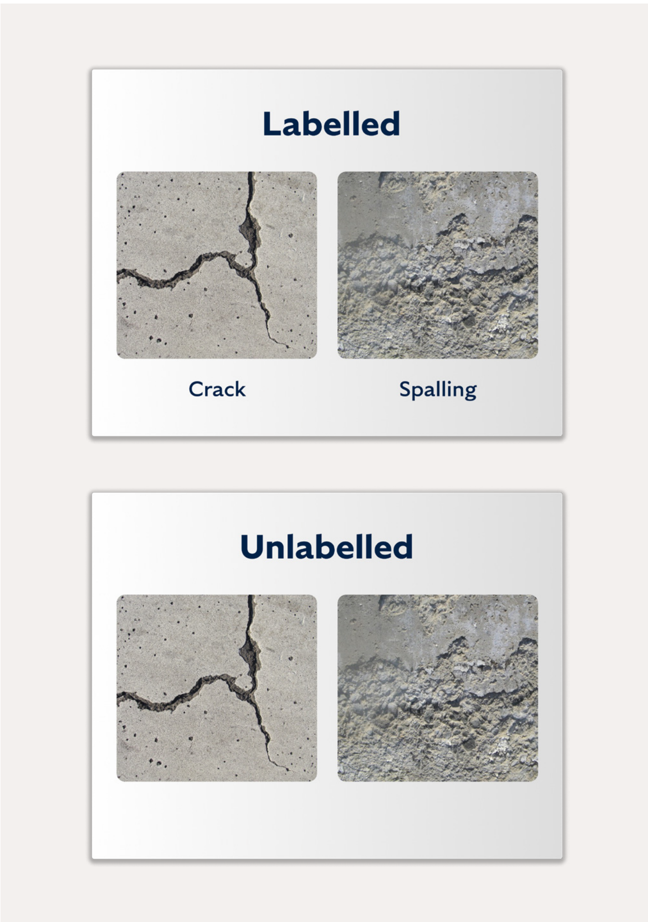


Figure 1 - labelled and unlabelled image data

Going a step further, there are multiple ways an image can be labelled. Figure 2 below illustrates the different types of labelling used with images, going from higher-level image classification, where the entire image is labelled as belonging to one class, to object segmentation, when pixel delineation provides the most accurate type of labelling. Object segmentation is more time-consuming (and often more expensive) than image-level classification, so it is important to understand the labelling requirements for each use case to determine what level of detail is required.

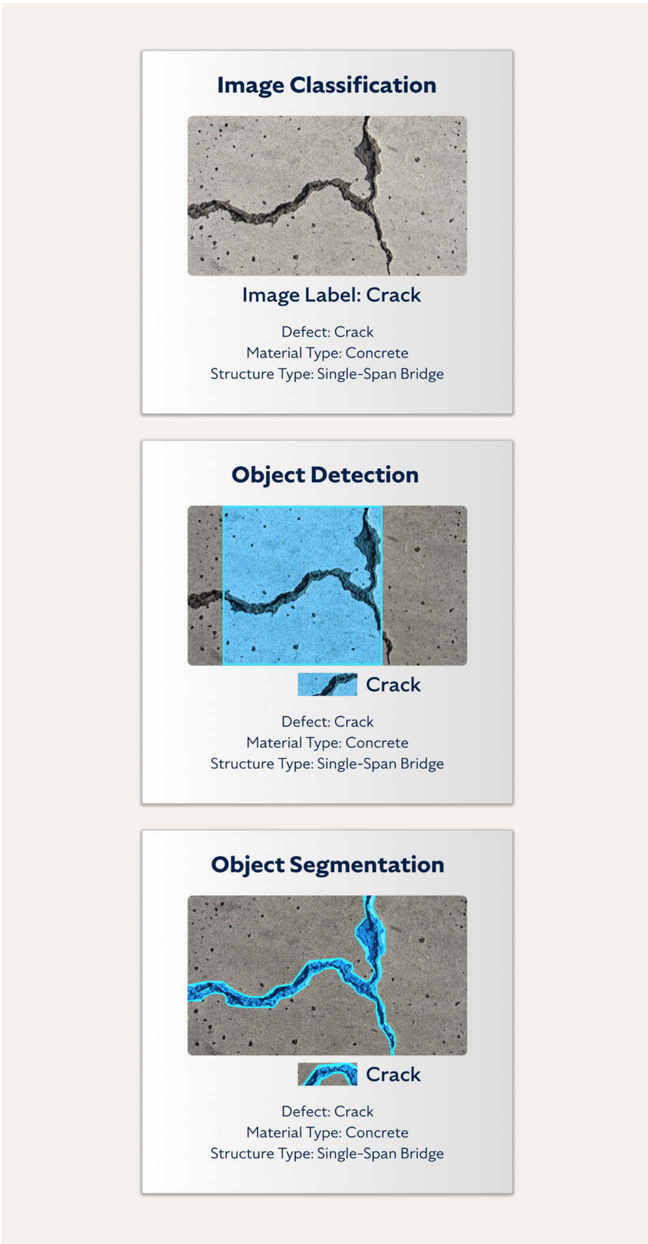


Figure 2 - Image classification, object detection, and object segmentation

Example 2: labelling structured tabular data

Figure 3 below presents an example of labelled and unlabelled tabular data. Entries with a condition class are labelled, and those without a class are unlabelled. Each condition class corresponds to a label. Additionally, since this data is a standardised tabular dataset, it is structured.

Retaining wall ID	Wall height (m)	Type	Material	Wall rotation	Condition (class)
1	6.0	Gravity	Masonry	Yes	Excellent
2	5.7	Embedded	Concrete	No	Poor
3	3.4	Cantilever	Concrete	No	Fair
4	4.5	Embedded	Concrete	Yes	-
5	6.2	Cantilever	Concrete	No	-

Figure 3 - labelled and unlabelled tabular data

Data features

The last important concept to understand about data is the concept of features. Features are characteristics or properties of a dataset. For example, in Figure 2, each image contains multiple labelled features, like the defect type, the material type, and the structure type. In Figure 3, each column header (wall height, type, material, etc.) corresponds to a feature.

Gathering quality data

Obtaining well-labelled data can be challenging and expensive. For example, there may not be enough data to train a model on, even if it is labelled. The data may be outdated, incomplete, or not representative of what the trained model will see in production.

When quality data is not easily available, or data collection is infeasible, some solutions exist, such as purchasing third-party data, paying a company to label data for your organisation, and generating synthetic data. However, these solutions are not always guaranteed to work, and are usually expensive.

Understanding these data concepts and limitations around data is the first step in determining whether AI can be applied to a specific use case and which types of models can be considered based on the data available.

Application areas

Natural language processing

Natural language processing is a field of AI that focuses on the interaction between computers and human language. It enables machines to understand, interpret, and generate human language, allowing for applications like chatbots, language translation, and sentiment analysis. Natural language processing combines techniques from linguistics and machine learning to bridge the gap between human communication and computer understanding. For example, natural language processing can be used to extract text from a Computer-Aided Design drawing, which can then be fed into a rule-based system to automatically generate a client report.

Computer vision

Computer vision is a field of AI that enables machines to interpret and understand visual data from the world, such as images and videos. It involves techniques like image recognition, object detection, and scene understanding, allowing systems to analyse visual inputs and take actions or make decisions based on them. Their applications focus on image processing and are therefore relevant for image-related tasks, as illustrated in Figures 1 and 2 on page 13. For example, the use of cameras or drones with computer vision algorithms to measure the volume of earth that has been moved during a construction project.

Unleashing the power: AI meets cutting-edge technology

In this section of the white paper, we delve into the potential of AI and its ability to amplify its capabilities when combined with other advanced technologies. We will explore a variety of innovative combinations—such as AI integrated with the Internet of Things (IoT), robotics and autonomous systems, edge computing and more—that create transformative solutions across infrastructure.

Internet of Things and Big Data

The field of Internet of Things is a network of physical objects that can connect and exchange data with other devices and systems over the Internet, for example, communication protocols (like Bluetooth) and communication networks (like the Cloud). It enables communication between the devices themselves, as well as between the devices and the Cloud on which data is stored, accessed, and analysed. These devices generate a lot of data, making them a perfect use case for AI and big data, a field that looks at how to manage, store, and process incredibly large amounts of data to efficiently extract insights.

For example, strain-gauge sensors could be placed along a bridge to monitor strain over time. These sensors collect a lot of data, so big data techniques could be used to process it, and AI could be leveraged to make predictions about the bridge's condition.

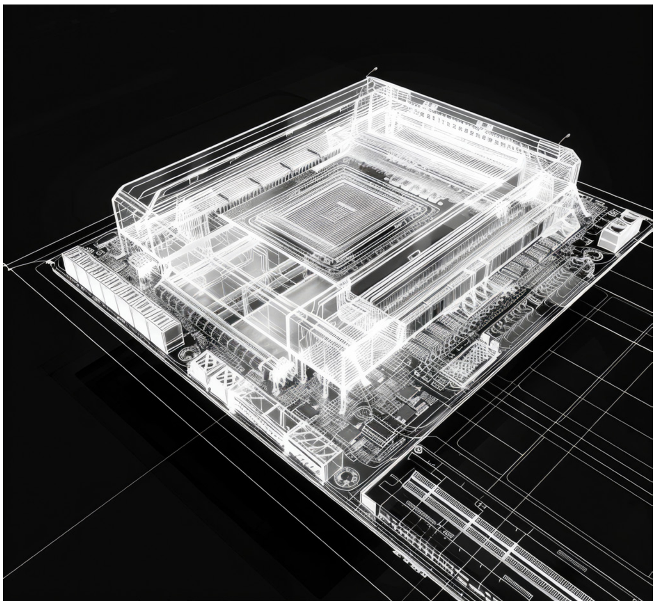
Robotics and autonomous systems
A robot is a programmable machine capable of carrying out tasks autonomously or semi-autonomously. On the other hand, automation refers to a broader concept that involves using technology to perform tasks automatically without direct human intervention. Drones, robots, and autonomous systems are increasingly used in civil infrastructure. However, because this technology is expensive and requires extensive customisation, it is often used solely for large and unique assets that are difficult, if not impossible, for a human to access. AI is then used in tandem with this technology to plan and optimise routes for autonomous systems.

Computer vision can be leveraged here too to analyse what these systems see (e.g. detecting the presence of any defects in its vision). For example, a drone could follow an automatically generated and optimised path around a very hard-to-access structure to capture photos of every surface of the asset. Alternatively, a remotely operated vehicle could be guided around confined underwater spaces, like tidal culverts.

Edge computing
Edge computing is a computing model that moves data storage and computation closer to the devices that generate the data. It enables the storage and processing of data at the location where it was collected, rather than having to send it somewhere else (like the Cloud or transferring it via physical means). By building on the examples of robotics and autonomous systems, these systems could identify defects in real-time through edge deployment, enabling faster identification and quantification of defects.

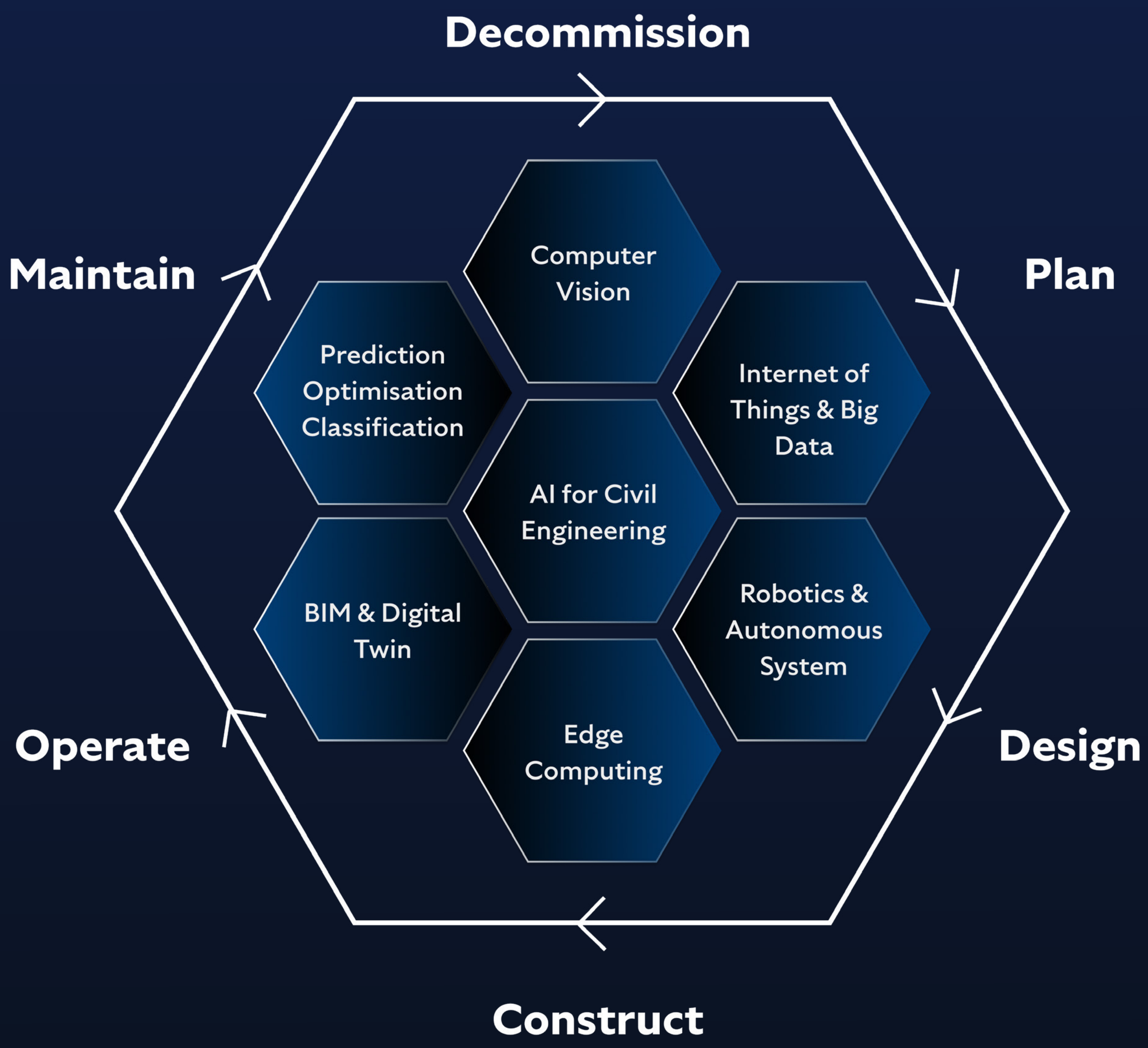
Building Information Modelling
Building Information Modelling (BIM) is a collaborative process for creating and managing digital models of buildings and infrastructure projects. It provides detailed information on assets throughout their entire lifecycle, from planning to decommissioning. This requires bringing together a lot of information from different sources in different formats. As such, AI used in conjunction with BIM can optimise the various processes and workflows required during and between each phase of the asset lifecycle, for example, with resource allocation and project planning. More examples of AI through the asset lifecycle will be covered in the next section.

Digital twin
Digital twins are a virtual representation of an object or system designed to reflect a physical object as accurately as possible. It spans the object or system's lifecycle, is updated from real-time data and uses simulation, machine learning, and reasoning to help make decisions. Digital twins of assets have recently started to be created for major CapEx projects. Since they digitally represent an asset, bringing together a variety of data like sensor data and images, they can be significantly enhanced with AI for more proactive maintenance of an asset.



AI Applications Across the Asset Lifecycle

This section of the whitepaper details specific use cases for AI across the asset lifecycle, looking at planning, design, construction, operations, maintenance, and decommissioning of assets.



The different phases of the asset lifecycle can easily blur into one another, but they've been presented here as separate stages to contextualise and present possible use cases which can leverage AI. This doesn't necessarily mean that a use case should be limited to that phase.

Planning

The planning stage requires bringing in a wide variety of information to understand what can be built, where it can be built, and what the constraints are (e.g. surveys, ground investigations, impact assessment, and economic feasibility assessment). This phase can create significant pain points due to the challenge of bringing together and making sense of a variety of data from multiple sources with varying degrees of uncertainty. Consequently, most AI use cases at this stage will focus on the optimisation of information and planning to help make the most informed decisions.

Since this phase is very iterative and blends in with the design feasibility stage, the following use cases are also relevant in the design phase.

Use cases

- Supply/demand and pricing forecasting for materials to make more informed decisions during the optioneering phase.
- Optimising analytics to find the best site using constraints such as environmental, demographic, and geological factors.
- Optimising resource planning and allocation based on specific headcount and budget constraints.

Design

The design phase ranges from concept designs to technical designs and brings together the work done by architects and structural, civil, mechanical, electrical, and geotechnical engineers. As such, multiple use cases can be explored for each of these professions. As models are often used at this phase, exploring AI applications to enhance BIM can be very valuable.

Use cases

- Architects and structural engineers can speed up their design process by leveraging AI-powered CAD for blueprint drawing, structural analysis, simulations, and identifying risks and blockers in proposed designs. Plans and documentation can be automatically drawn, written-up, and shared with other stakeholders like mechanical, electrical, and civil engineers, who need to receive that documentation as soon as possible.
- Designers can look into Generative AI (specifically image generation) to speed up the ideation process by transferring ideas to paper more quickly and conveying these ideas in reports.
- Finally, BIM can be enhanced with AI to predict risks and possible failure points as early in the process as possible before construction begins.

Construction

Applying AI to monitor construction phase safety can be extremely valuable, since on-site accidents are one of the largest safety risks in the construction process. As on-site construction isn't done in a controlled environment like a factory, there are more variables that need to be controlled in order to mitigate unforeseen events. As these variables become more complex, AI can assist us in understanding them in real-time.

Use cases

- Applying computer vision to video captured on-site (either from a static video feed or video captured from a moving vehicle) as an additional safety monitor to detect possible accidents.
- Optimising project management by proactively recommending actions to resolve resourcing, financial, and material delays.
- Predicting and forecasting material orders based on pricing and project plans.
- Leveraging optical character recognition to read receipts and manage budgets and accounting.
- Using robotics and machine control to automate dangerous or high-precision tasks.

Operations

The operations stage is crucial as it accounts for up to 80% of total asset costs, so leveraging AI to lower costs can be extremely valuable. Using AI in tandem with other technologies (such as IoT sensors and robotics) can also make operation processes more efficient.

Use cases

- Using AI to optimise maintenance operations based on resource and budget constraints, asset type, and asset information from the BIM model.
- Leveraging computer vision in tandem with robotics and satellite imagery to detect, classify, and quantify asset deterioration.
- Ensuring accurate and live structural health monitoring of assets by using AI on IoT sensors to monitor for anomalous structural changes.
- Using AI to operate systems associated with an asset's day-to-day use.

Maintenance

The maintenance phase can gain the most value from AI by ensuring the timely and proactive maintenance of assets. Furthermore, managers can save time and resources by eliminating unnecessary inspections. In the maintenance phase, AI use cases revolve around understanding and quantifying the asset's condition and deterioration, often alongside other technologies, to ensure proactive rather than reactive maintenance. This approach will lower capital expenditure costs and increase the asset's lifespan.

Use cases

- Using AI to assist on-site inspections:
    - Combine computer vision with drones to facilitate the inspection of hard-to-reach or dangerous-to-reach areas on the asset, as well as to detect and quantify defects observed by the drone.
    - Combine computer vision to objectively quantify the defect extent observed during an inspection.
- Optimising maintenance plans and programmes based on predictive deterioration models.
- Applying computer vision to satellite imagery of retaining walls to quantify their rotational movement over time.

Decommissioning

The decommissioning stage will be very similar to the construction stage in terms of AI use cases due to similarities in the problems that need to be tackled, specifically questions around how a decommissioning project can be optimally managed while balancing resource, time, and financial constraints, and how the decommissioning of the asset can be made as safe as possible. Additional tasks specific to the decommissioning phase can be enhanced with AI as well, like managing waste disposal, creating an environmental impact assessment, and understanding what can be retrofitted versus what must be demolished.

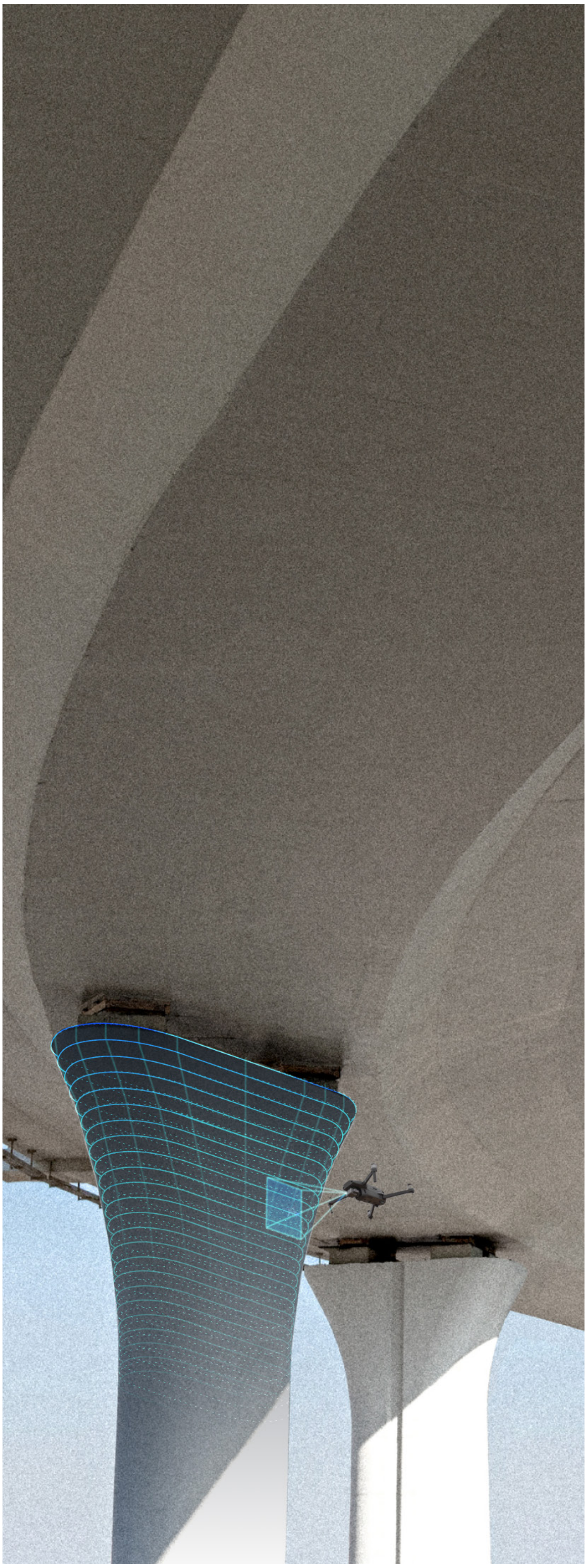
Use cases

- Using AI and computer vision to aid in overcoming uncertainties, e.g. identifying potential asbestos items in buildings with undocumented construction details.
- Leveraging computer vision to preemptively detect possible accidents, optimising project management based on given constraints.
- Using computer vision to determine and quantify asset conditions to assess the possibility of retrofitting.
- Using AI for compliance procedures such as waste disposal documentation.

Conclusion

As we have explored, there are an incredibly large number of AI use cases that can be applied through the asset lifecycle. Figuring out the best use cases for your organisation will vary based on the number of assets, their type, and their condition, as well as financial constraints, what stage of the lifecycle you focus on, data availability, digital maturity, and more.

Reflecting on what use cases would bring the most value to your organisation is, therefore, the first step. Start by implementing and testing smaller use cases as minimal viable products and develop from there. This will ensure value is delivered with AI transformation without requiring a huge budget.



Crucial Factors for Civil Engineers when Choosing AI Systems

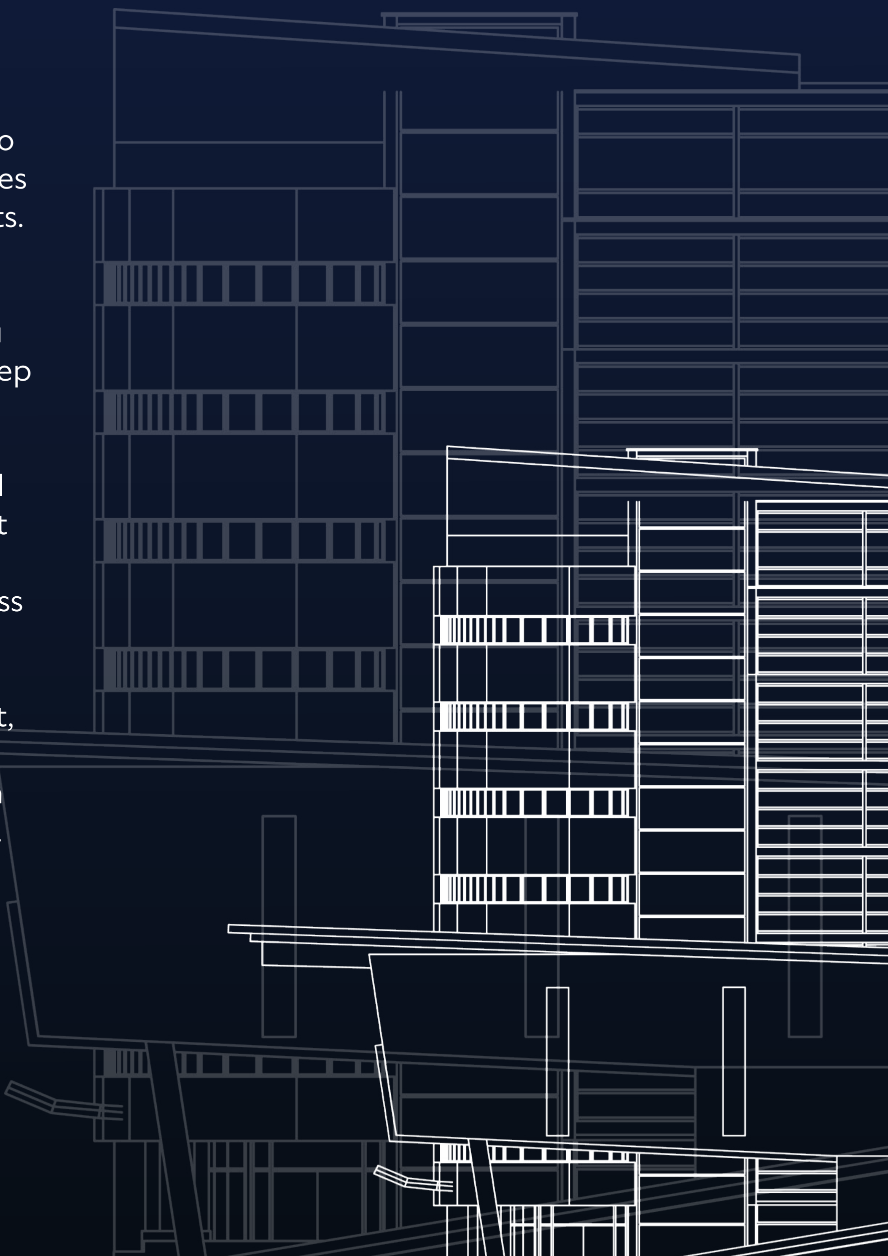
This section of the whitepaper looks at what to consider when evaluating a use case and selecting an appropriate AI system.

Business process and benefits

As covered in the “Types of AI systems and capabilities” section, AI comes in all shapes and sizes, from simple rule-based systems to automated repetitive administrative tasks and complex computer vision algorithms to identify defects in asset images to deep learning algorithms to optimise asset management at a portfolio level. AI can also use a variety of data modalities, from images and natural language to structured datasets.

Whatever problem needs to be solved, it is crucial to understand the use case that you are looking to augment with AI. The first step includes mapping out the corresponding business process as it stands today and determining what it would look like after AI implementation. Also, you need to map out the desired outcomes of implementing AI and tie them to specific steps in the business process. This ensures that value delivery (addressing the problem that needs to be solved) remains at the centre of the project, rather than over-indexing on a technical solution (how the problem is solved), which may not be suited for the problem at hand.

This thought process is important to ensure that AI is the right solution to tackle a particular problem. Where organisations go wrong is attempting to use AI as a blanket solution for all digital transformation projects without first deeply understanding their current business processes.



Data availability

Data availability will inform your understanding of what type of AI to use and whether AI can be used at all. Here are a few guiding questions to better understand what data you have at your disposal.

1. **Is there any existing relevant data in the organisation?**
2. **How is the data stored?** Is it accessible? Is it (consistently) structured?
3. **Is there sufficient data for the model to learn what it needs to learn?**
4. **Is the data good quality?** (e.g. if the data is image-based, what is the resolution of the images? When is the data from: 2 months ago, 2 years ago, 20 years ago? Is it still relevant today?)
5. **Does the data reflect what the model will see when deployed in the real world?**
6. **Is the data labelled or unlabelled?**
7. **If the data is labelled, how well has it been labelled?** Does the type of labelling match the requirements? How accurate or reliable are the labels?
8. **If the data is unlabelled, can it be labelled?** Or can similarly relevant labelled data be found elsewhere?
9. **How much variability is there in the data?** (For example, if all images of cracks are on clean, smooth concrete surfaces, this would be considered low data variability, while a wide range of highly textured materials such as brickwork and decorative render would be considered high data variability).
10. **Does your organisation have the right to train a model on this data?** Ensure that users have explicitly given approval for their data to be used for this purpose. Note that retroactive approval, given after the data has been collected, cannot be given.
11. **Where will your model be used?** Ensure the data being used and the model being created meet the legal framework of the relevant countries. For example, if a model is built in the US and deployed in Europe, it must adhere to GDPR rules.
12. **Who will the IP belong to?** This question concerns the data, the model, and the deployed solution containing the model.
13. **Based on the questions above, how much available, usable data is there today?** If there isn't enough, what would a path forward look like?

It is imperative to seek advice from a data scientist or a machine learning scientist to answer these questions accurately.

For example, let's say you want to build a model to identify cracks on images of concrete assets specifically.

1. **Is there any existing relevant data in the organisation?** The organisation has images of cracks from past inspections, specifically for concrete assets. If possible, it would be useful to quantify how many.

2. **How is the data stored?** The data is stored in PDF inspection reports in the organisation's asset management system. Extracting those images might be tricky.

3. **Is there sufficient data for the model to learn what it needs to learn?** To train a good starting computer vision model, about 1000 labelled images are required.

4. **Is the data good quality?** The data ranges from 20 years ago to today. It's worth looking into the resolution and quality of the older images.

5. **Does the data reflect what the model will see when deployed in the real world?** The data is made of images of cracks on a variety of concrete buildings, which matches the goal of the model.

6. **Is the data labelled or unlabelled?** The data is currently unlabelled. However, by matching up text in the report with images in the report, it may be possible to label the images.

7. **If the data is labelled, how well has it been labelled?** Image classification. This will pose limitations to eventually quantifying crack sizes.

8. **If the data is unlabelled, can it be labelled?** N/A

9. **How much variability is there in the data?** The images seem to vary well in terms of lighting, crack sizes and shapes, crack orientations, and the material on which the cracks have been photographed.

10. **Does your organisation have the right to train a model on this data?** The images were captured by your organisation. However, the assets belong to the asset owners.

11. **Where will your model be used?** The images were captured in the UK, and the model will be used only in the UK, so it is recommended looking up the UK's AI and data framework and policies to ensure the model can be used and deployed as desired.

12. **Who will the IP belong to?** Your organisation will build and deploy the model and develop the solution. The ownership of the data needs to be clarified.

13. **Based on the questions above, how much available, usable data is there today?** If there aren't enough images, other avenues can be explored, like looking up 3rd-party datasets online, getting the images labelled by humans, or partnering with a company which has such data. (In most cases, the solution is likely to require the creation and deployment of collaborative human-AI data pipelines, which start their life by delivering immediate process efficiency savings and pave the way for increasingly higher degrees of automation and better model performance. Until the model reaches a steady state where the human experts are only tasked with minimum supervision of AI in the mainstream, and most of human attention is spent on messy rare edge cases.)

This example illustrates how to get an initial feasibility assessment based on available data. It also illustrates that the path to integrating an AI solution into an organisation can be challenging but rewarding.

Responsible AI

This section of the whitepaper will break down some frequent terms used in responsible AI.

Responsible AI refers to the development and deployment of artificial intelligence systems in a manner that is ethical, transparent, and aligned with societal values. It emphasises fairness, accountability, privacy, and the minimisation of biases to ensure that AI technologies benefit individuals and communities without causing harm. Responsible AI practices also include ensuring that AI decision-making processes are understandable and accessible, fostering trust and inclusivity.

AI governance is also frequently mentioned in the context of responsible AI. AI governance refers to a set of principles or processes that ensure your AI is developed and deployed in a way that is responsible, ethical, and aligned with business strategies and current regulations.

Fairness and bias

Fairness is about ensuring that an AI system makes decisions without bias and without unfairly discriminating against any groups or individuals. Bias refers to the occurrence of biased results due to human biases that skew the original training data or AI algorithm, leading to distorted outputs and potentially harmful outcomes. Data bias is a big factor when considering the fairness of an AI system. If there is bias in the dataset used to train the AI, then the bias will be reflected in the AI system. The challenge is that bias may not be apparent in the data but may still be present through proxy variables.

For example, take a model that is being built to assign a maintenance prioritisation score to assets. It is determined that the model should not consider any potentially biased geographic indicators as a feature to avoid discriminating against less affluent areas that may not have received objective priority scores in the past. Data scientists may have removed all features explicitly indicating a geographic location, but another feature (like the monetary amount spent on asset maintenance) may be unknowingly acting as a proxy for geographic location, thus propagating the bias from the data. It is essential, therefore, to ensure tests and frameworks are in place during the model validation stage to check whether this bias is still occurring or not.

Trustworthiness

Trustworthiness refers to the reliability of an AI system. It refers to a combination of different values, such as accuracy, safety, and robustness. Accuracy is about how well an AI system does what it is meant to do. Safety ensures that the AI does not have any harmful or unintended consequences. Robustness ensures the AI system’s ability to handle anomalies, such as missing data, consistently.

Explainability

Explainability refers to methods and techniques that make the decision-making processes of AI systems transparent and understandable to humans. Unlike “black box” models, explainable AI provides insights into how an AI model reaches its conclusions, allowing users to interpret, trust, and verify the outputs. This is particularly important in high-stakes applications like infrastructure, where understanding the rationale behind AI decisions is critical. Generally, more easily explainable models are less complex. Therefore, there is often a slight trade-off between model accuracy and model explainability.

For example, an organisation might be using a model to optimise asset inspection scheduling. In this case, it is crucial to have transparency on the model's predictions because the organisation will most likely have to justify their inspection prioritisation due to the cost of running such inspections. On the other hand, if the organisation is using a model to alert of urgent maintenance required on a bridge based on sensor data readings from IoT devices installed on it, transparency will probably not be as crucial. These alerts usually indicate safety-critical situations, and the risk of losing human lives due to a bridge collapse far outweighs the need for transparency to justify money spent on sending someone to inspect the bridge.

Impact of responsible AI

Following responsible AI principles is critical when deploying AI systems. Not doing so can lead to inaccurate and/or outdated predictions at best. At worst, it can negatively and unfairly impact individuals and societies at scale and put people in dangerous situations. As a result, this can lead to financial losses and reputational risk, severely impacting organisations.

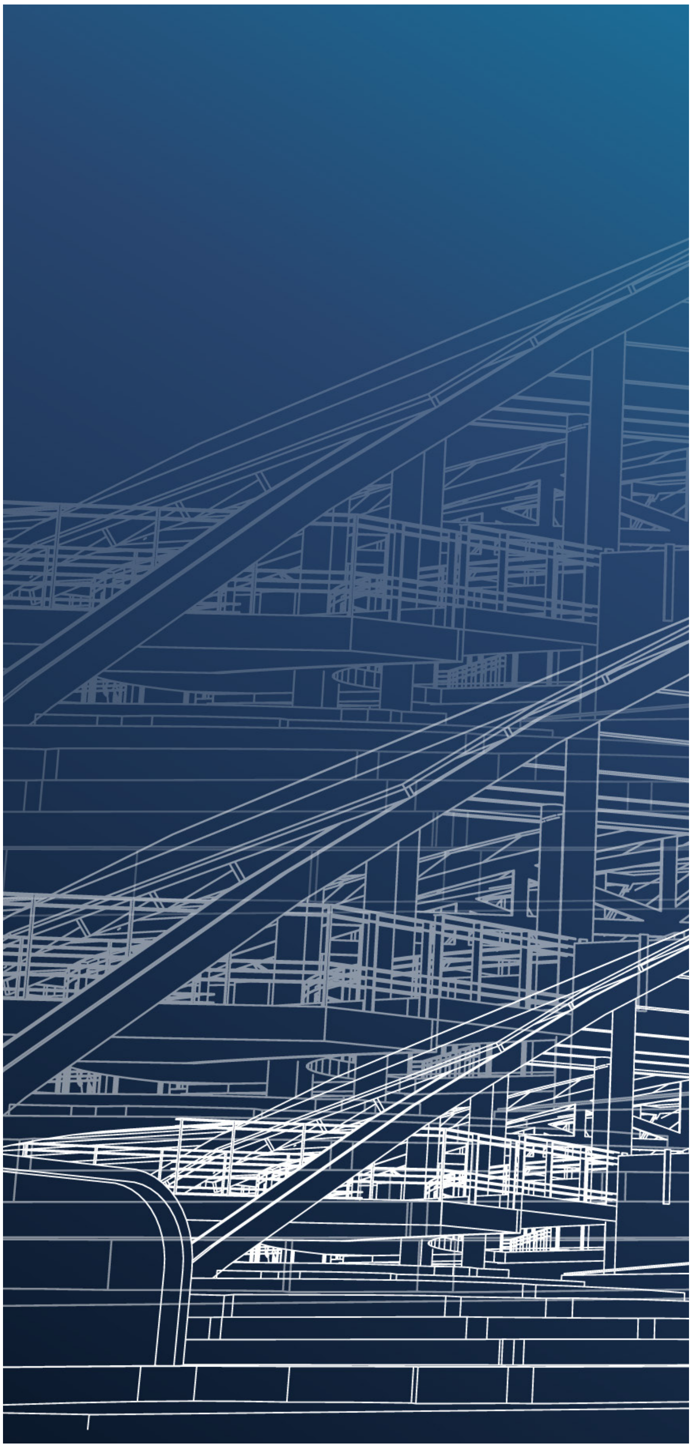
Responsible AI principles concern the entire model lifecycle, from the moment a use case is discussed to the moment the AI system is decommissioned. It is recommended that organisations set out and reflect on their own responsible AI principles, and ensure they are used as a north star throughout the AI applications.

When starting out on an AI application, it is recommended to run a responsible AI workshop with all members of the team to build an AI risk register, and implement a machine learning operations strategy and solution to manage these risks. Gathering viewpoints from different perspectives is crucial.

This topic will be further explored in the last section of this whitepaper (see page 27).

Additional resources on responsible AI can be found here:

- Approaching Ethical AI Design
- What Makes a Machine Learning Model 'Trustworthy'?
- Explaining AI Explainability



Essential Insights for Civil Engineers Deploying AI Systems

This whitepaper has explored what various AI systems can look like and how they can be applied throughout the asset lifecycle. This section will detail what deploying AI in an organisation looks like.

Deploying AI in an organisation will reflect the AI lifecycle, which is a continuous process from the moment the use case is explored to the moment a model is retired. Each stage requires bringing together different people with different expertise to ensure a successfully built and deployed AI system.

Identifying the problem space and defining the use case

This first stage of the AI lifecycle is crucial as it informs all the other steps. This stage will usually be led by a digital transformation manager, product manager, project manager, and/or solutions engineer. This person will be the expert on the use case and be responsible for progressing the project throughout the AI lifecycle.

As detailed in the “Business process and benefits” section, this stage provides the space to explore and understand the current business problem and the pain points to be solved.

The outputs of this stage are a solid understanding of what this digital transformation project entails and ensuring whether or not AI is the best solution for this problem.

Responsible AI workshop

Once the problem and solution spaces have been clearly articulated, it is time to conduct a responsible AI workshop. These workshops are most successful when the participants represent a range of experiences and perspectives.

The output of this workshop is an AI risk register, a document that tracks all possible risks throughout the project, who is in charge of them, and what will be done to mitigate them. The project owner and those responsible for risks in the risk register should update this document throughout the project.

A risk register is more than just paperwork required for the AI-building and deploying process. The outputs of AI models, especially in civil infrastructure, can impact lives and populations. Anyone who contributes to the building and deployment of an AI system, whether a project manager or machine learning scientist, has a moral responsibility to ensure that the implications of the system have been mapped and mitigation strategies are in place.

Data collection and preparation

At this stage, data scientists and/or machine learning scientists (depending on the structure of the organisation) will start to have a central role. This stage is about working with the project owner to gather data, whether that is using company data, purchasing 3rd party data, hiring a company to label unlabelled data, etc.

Once the data has been collected and/or labelled, it needs to be cleaned and prepared so it is suitable for model building. This includes a variety of data processing steps, such as handling missing data values and creating features to augment the data. This stage can also be done parallel to the data collection stage to verify that the collected data meets the technical and project requirements.

Definition
Feature engineering is a crucial step in building AI models to extract as much value as possible from the dataset. A feature is a measurable characteristic in a data set. For example, a model is being built to predict a bridge's condition score. Some features might include the bridge's material, street address, and historic weather patterns by postal code.

Some features may not have any predictive power, which is a key discovery in the data preparation step. For example, the bridge street address alone probably does not indicate anything about the bridge's condition.

However, features can be enhanced. For example, the bridge street address might not have predictive power, but converting the street address into a postal code might have predictive power when associated with historical weather patterns provided by postal codes because bridges in regions with harsh weather conditions may be in worse condition than bridges in regions with tempered weather.

Feasibility exploration

After the required data has been collected and prepared, but before jumping into model training, feasibility exploration work will be required to ensure the data is suited for the use case being solved with AI. This stage involves further data exploration and initial exploratory modelling work to validate the technical feasibility of this AI solution.

Getting the IT team involved at this stage is crucial as well. One of the top challenges that exist when deploying machine learning solutions is integrating them with existing systems. The best way to overcome this challenge is to involve IT as early as possible in the model-building process, so they can prepare.

Model training, evaluation, and validation

This stage is usually led by machine learning scientists, who have the expert knowledge to create AI models. However, it will require close collaboration with the project owner so they understand the project requirements and can integrate them into the model-building process. Machine learning scientists will need to choose the right model to use based on the type of problem that is being solved, the data available, and the required level of interpretability. Multiple models can be considered and compared against each other.

Model training consists of feeding a training dataset into the AI model and comparing its output with expected outputs to learn and optimise for different feature values. The machine learning scientist can then move on to the evaluation stage to validate that the model outputs correct information.

The evaluation stage quantifies the model's performance through a set of standardised metrics, which will depend on the model type. The machine learning scientists will switch back and forth between the training and evaluation stages until the model performs as desired.

Machine learning operations, model deployment, and model maintenance

Model deployment is the stage at which the AI system is brought to production. Depending on the size and structure of the organisation, model deployment can be managed by machine learning engineers, scientists, operations engineers, software engineers, and IT. The model can be deployed in multiple formats: as API endpoints, integrated into an already existing software, or in a new piece of software which requires its own custom interface. If it is the latter, a product designer will be required to design the interface, in collaboration with the project owner, to understand the customer, business, and technical requirements and constraints.

Another aspect of model deployment to think about is where the model will be deployed, whether in the Cloud or locally on an edge device. Deploying models at the edge is advantageous when inference times need to be very fast, or the internet connection is not reliable. For example, if a drone is flying around an asset in a rural area to capture images of defects, the speed at which the drone identifies defects is crucial, and an internet connection may not be guaranteed. In this case, it is advantageous for the model to be deployed as part of the drone's software. However, deploying at the edge can also bring other additional engineering challenges and may require specialist engineers. So the pros and cons of each deployment method must be weighed within the context of each use case.

Human-AI interface and experience

Understanding and defining how the model output will be consumed is a large piece of work when bringing AI systems to production. The term "Human - AI Interface" looks at how humans interact and work with the AI system and, more specifically, what the UX (user experience) and UI (user interface) will look like. Having a suitable human-AI interface for the AI system is crucial for two main reasons:

- Ensuring the AI system is used as intended, as defined through responsible AI principles.

- Ensuring that humans drive the decisions, rather than AI. Notice how the word "human" is placed before the word "AI", highlighting the crucial roles of humans in AI systems. This concept is referred to as "AI-in-the-loop", and is described below.

The first question to kick off the design of a Human-AI interface is, who will users be? Not everyone needs the same information displayed. Users with a technical background may want to know the internal model mechanisms, whereas non-technical users may want the model output explained in natural language. Equally, providing too much information on the screen can be just as confusing as not presenting enough. Finally, most model predictions don't have 100% accuracy. Highlighting and framing model accuracy is important so that users can observe how good the model predictions are and make informed decisions about how model predictions are used.

When designing human-AI interfaces, remember that AI should not replace humans but rather enhance their work and take over repetitive tasks.

Human-AI interaction is about how users engage with the model output. Enabling the user to correct and provide feedback to the model is crucial to enable it to continuously improve over time. This concept is often referred to as AI-in-the-loop, which emphasises that AI can augment humans as an assistant, but humans should remain in charge of decision-making.

Equally, it is important to explore how to empower humans to make confident decisions in AI systems. In AI, automation bias is a phenomenon in which humans tend to over-rely on the decision made by an AI system and not question it, even though they are more knowledgeable than the AI. This phenomenon is especially common when the user must repeatedly verify a model output in the same format and goes into "auto-pilot" mode. Designing a system that puts human decisions first is crucial.

Continuous Metalearning

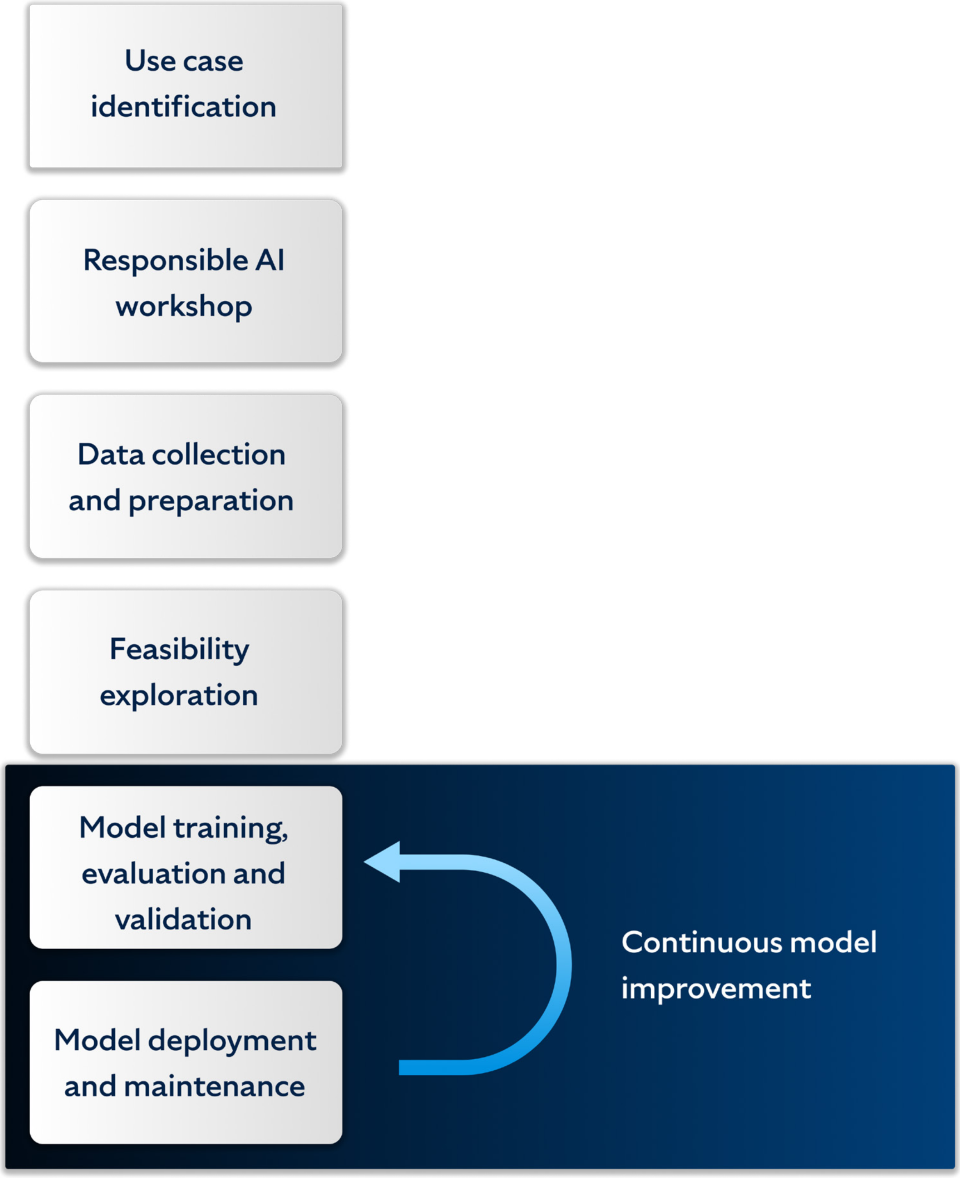
Continuous Metalearning is a technological capability that allows AI models to optimise their learning process and adapt to changes in the real world. It offers protection from the risks associated with misaligned AI and enables models to get better over time. Once the model has been deployed to production, it will be actively surveilled for a couple of months. When confidence in the model and AI system has been earned, maintenance of the model can be performed on an ad-hoc basis, and changes can be made when necessary (e.g. the format of the data has changed, a new requirement needs to be implemented). However, continuous monitoring of the AI system should be implemented at this stage. This process can be automated by surveilling specific model metrics and performing automated improvement of the model over time. The latter is also known as Continuous Metalearning, which ensures the model is retrained over time using human feedback so it remains relevant and does not degrade. Building responsible and transparent AI systems and including AI-in-the-loop to gather continuous feedback can also facilitate timely model monitoring.

Studies have shown that 91% of machine learning models degrade over time. This can be due to two things, model drift and data drift:

Model drift (also known as concept drift):
Concept drift occurs when the relationship between the input data and the target value changes in some way, potentially making the model inaccurate or unreliable. Concept drift can lead to a decline in the model's accuracy because it was trained on data that no longer reflects the current patterns or relationships. For example, historical retaining wall maintenance data is used to predict when future maintenance is required. Recently, construction companies have been using a new type of concrete to build retaining walls, which has a different lifespan and degradation pattern (which correspond to different parameters from older retaining walls). The model wasn't trained on these new parameters, so its maintenance recommendations will be inaccurate.

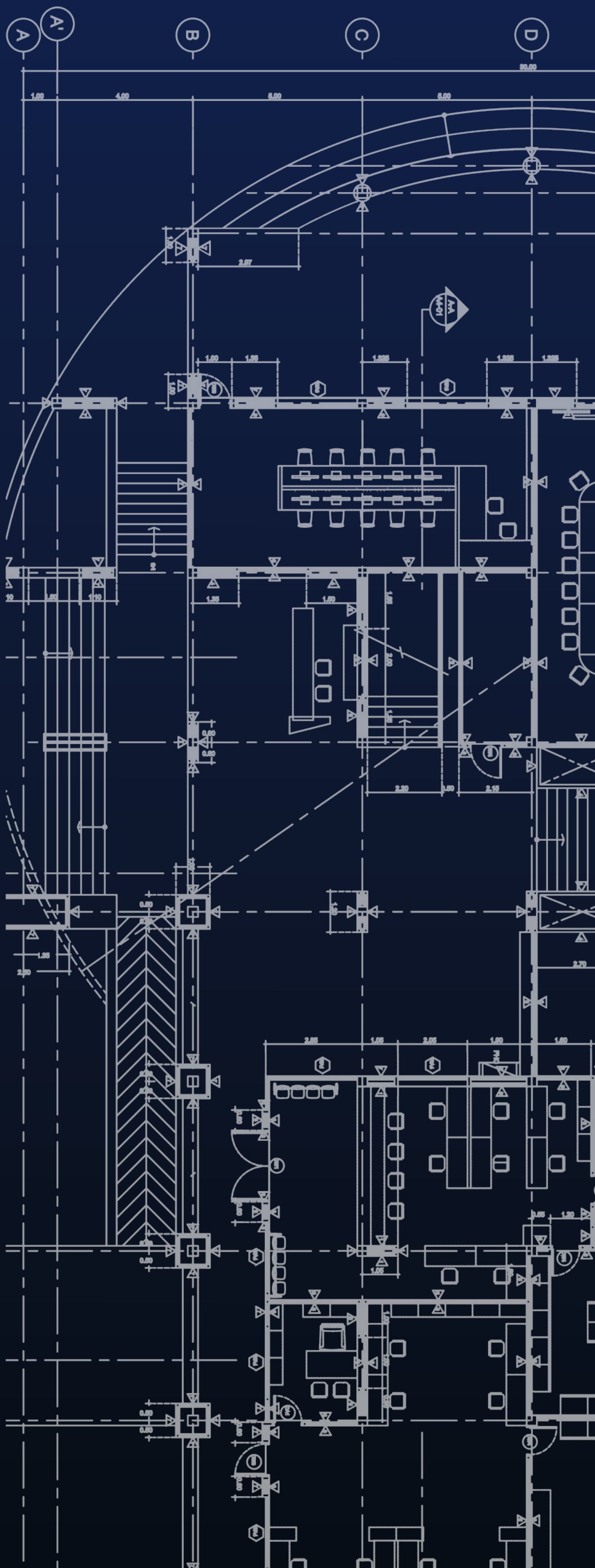
Data drift (also known as covariate drift):
Data drift occurs when the distribution between the inputs of a machine learning model changes, which can cause a machine learning model to perform poorly. This shift happens when the characteristics of the data used for training the model differ from the data it encounters in production. Monitoring and addressing data drift is crucial to ensure that models continue to make accurate predictions in real-world applications. For example, sensors have been placed on cables of a suspension bridge to monitor for unusual physical activity. As part of a significant maintenance campaign, the cables have been recently replaced so the bridge can support additional weight. However, this change wasn't communicated to the model, so the model frequently sends alerts to detect tensions above the accepted threshold.

A way to address this issue is Continuous Metalearning. This technique enables the auto-retraining of a model to adapt to changes in the environment or respond to human feedback provided on the model output. More can be read about this topic here.



Overview of bringing AI systems into production

What next?



Depending on the organisation's technical maturity, openness to innovation, and in-house knowledge, there is a range of options to explore to start leveraging AI in civil infrastructure, from self-serve off-the-shelf solutions and leveraging internal data and machine science teams, to partnering with specialised AI companies. This section of the whitepaper explores what these different options require and offer.

Buying or building in-house

In-house development requires a solid understanding of the use cases that need to be solved in the civil engineering space. And depending on the levels of technical know-how within the organisation, building in-house or buying an external solution are suitable options. It is worth exploring both of these options and seeing how they match up with business constraints such as budget, time, and resource investment, as well as the organisation's long-term goals.

The table on the next page lists different in-house development options, ranging from the option with the least upskilling required to the option with the biggest time and resource commitment.

Solution	Description	Pros	Cons
Off-the-shelf AI solutions	For organisations that have a solid idea of the use cases they want to address and a basic technical understanding of AI concepts. These solutions offer pre-built and pre-defined AI models and pipelines which require very little customisation.	Little upfront investment is often required. A good way to validate ideas. Requires a basic understanding of AI concepts.	Difficult to gain a competitive edge because competitors can use the same solution. Locked into the solution. More expensive in the long term than building in-house. Don't always incorporate domain expertise in their offerings.
Off-the-shelf AI solutions and domain experts	More specialised than general off-the-shelf AI solutions, these solutions differentiate themselves by having AI directly integrated into their civil infrastructure-oriented solutions.	Solutions built for civil infrastructure use cases can help identify relevant AI use cases.	Difficult to gain a competitive edge because competitors can use the same solution. Locked into the solution. More expensive in the long term than building in-house.
AutoML	AutoML tools require a more advanced understanding of bringing AI to deployment, as they require the user to process their own data, build their data features, and build, deploy, and maintain their own model.	Offers more customisation options than off-the-shelf solutions. A good way to validate idea feasibility based on data available.	Can become locked in and dependent on the solution, difficult to gain a competitive edge because competitors can use the same solution, and don't always incorporate domain expertise in their offerings.
Built internally	Implement everything from data exploration to model deployment and maintenance using languages (e.g. Python) and open-source libraries (e.g. PyTorch, TensorFlow, and Pandas).	Cheaper and more autonomy in the long term. Full customisation is possible. Advanced to more cutting-edge AI.	Requires significant upskilling and long-term technical investment to maintain the AI system and team. Vulnerable to team turnover/change and can be limited by in-house knowledge.

Table 1 - in-house development options

The partner approach

Another option is partial external development through the partner approach. These companies will have varying degrees of AI and civil infrastructure knowledge, ranging from consultants who can provide initial advice on AI systems or civil infrastructure digital transformation use cases to AI specialist companies with a focus on civil infrastructure, who will guide an organisation in identifying and defining a use case all the way through model deployment and monitoring. Similarly, the output of these partnerships can range from an initial discovery paper to a fully deployed model with API endpoints and continuous model monitoring.

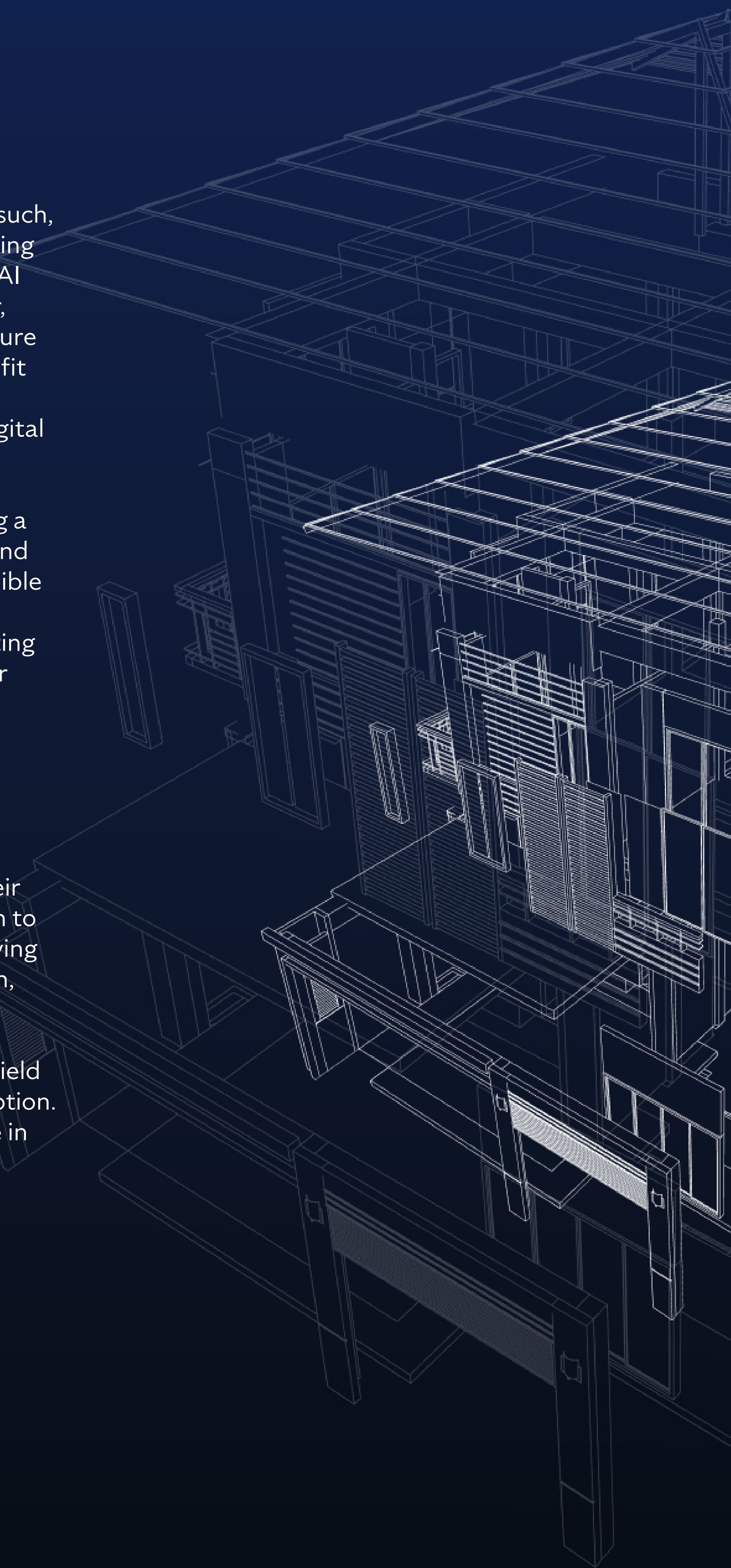
Solution	Description	Pros	Cons
Consultants	Digital transformation experts who can help identify use cases to enact digital transformation in an organisation.	Can help identify potential use cases. Can be civil infrastructure domain experts.	Rarely experts in deploying AI, usually focused on high-level digital transformation support, and more expensive than in-house solutions.
AI and machine learning specialists	Experts in applying cutting-edge AI and machine learning. Since working with these companies can be a commitment, they usually offer free, initial exploratory workshops.	Experts in machine learning, digital transformation, and responsible AI. Can build, deploy, and manage custom solutions responsibly to ensure they remain performant. Guidance from use case identification to model monitoring and long-term partnerships to drive profound digital transformation.	Can lack domain knowledge. More expensive than in-house solutions in the short term. Usually a long-term commitment and engagement.

Table 2 - external development options

Conclusion

In recent years, there has been a surge in AI due to the advancements in Large Language Models and Generative AI. As such, organisations are focusing on implementing Large Language Models and Generative AI to enact digital transformation. However, most organisations in the civil infrastructure and built environment space would benefit from considering a wider range of AI capabilities that are best suited to the digital transformation their organisation needs. Rather than picking a specific type of AI technology and then retroactively finding a use case, start by identifying a use case and problem space and then comparing possible technical solutions. Sometimes, digital transformation doesn't entail implementing complex deep learning models but rather automating a business process through simpler rule-based decision-making.

This white paper has broken down the concepts and vocabulary surrounding AI so that any civil engineer now has the tools to lead digital transformation in their organisation, from use case identification to defining responsible AI principles, deploying and maintaining AI systems in production, and collaborating with people in their organisation to make this happen. Civil infrastructure is an important, complex field and is still in its infancy regarding AI adoption. Now is the moment to make a difference in your organisation with AI.



About Mind Foundry

Mind Foundry is an Oxford University company operating at the intersection of innovation, research, and usability. We empower organisations with AI built and deployed to tackle high-stakes, real-world problems at both individual and population scales.

We aim to address the most pressing issues in infrastructure, from asset to portfolio. We deploy responsible AI to deliver direct cost savings in operational expenditure and unlock long-term benefits. Mind Foundry partners with owners and consultants across infrastructure sectors to create targeted impact through deployed AI systems.

To help you kick-start or progress your AI journey, we're offering a workshop where you and your team will cover the fundamentals of machine learning and AI, how to identify opportunities, and the different options available to get started.

We will come to your company (or you can come to us in Oxford) and deliver a session tailored to the attendees on a no-obligations basis and absolutely FREE of charge.

Sign up here.