AIR QUALITY EXPERT GROUP

Statement on integration and adoption of AI in air quality

Report for the Department for Environment, Food and Rural Affairs;
the Scottish Government; the Welsh Government; and the Department of Agriculture, Environment and Rural Affairs in Northern Ireland

This is a report from the Air Quality Expert Group to the Department for Environment, Food and Rural Affairs; Scottish Government; Welsh Government; and Department of Agriculture, Environment and Rural Affairs in Northern Ireland, on scientific and technical opportunities for the use of Artificial Intelligence and Machine Learning in air quality science and evidence. The information contained within this report represents a review of the understanding and evidence available at the time of writing.

Front cover image credit: AI generated air pollution image by Alan Frijns.

United Kingdom air quality information received from the automatic monitoring sites and forecasts may be accessed via the following media:

Freephone Air Pollution Information Service 0800 556677

Internet http://uk-air.defra.gov.uk

PB1307 doi: **10.5281/zenodo.14038922**

Terms of Reference

The Air Quality Expert Group (AQEG) is an expert committee of the Department for Environment, Food and Rural Affairs (Defra) and considers current knowledge on air pollution and provides advice on such things as the levels, sources and characteristics of air pollutants in the UK. AQEG reports to Defra's Chief Scientific Adviser, Defra Ministers, Scottish Ministers, the Welsh Government and the Department of Agriculture, Environment and Rural Affairs in Northern Ireland (the Government and devolved administrations). Members of the Group are drawn from those with a proven track record in the fields of air pollution research and practice.

AQEG's functions are to:

1. Provide advice to, and work collaboratively with, officials and key office holders in Defra and the devolved administrations, other delivery partners and public bodies, and EU and international technical expert groups;
2. Report to Defra's Chief Scientific Adviser (CSA): Chairs of expert committees will meet annually with the CSA, and will provide an annual summary of the work of the Committee to the Science Advisory Council (SAC) for Defra's Annual Report. In exception, matters can be escalated to Ministers;
3. Support the CSA as appropriate during emergencies;
4. Contribute to developing the air quality evidence base by analysing, interpreting and synthesising evidence;
5. Provide judgements on the quality and relevance of the evidence base;
6. Suggest priority areas for future work, and advise on Defra's implementation of the air quality evidence plan (or equivalent);
7. Give advice on current and future levels, trends, sources and characteristics of air pollutants in the UK;
8. Provide independent advice and operate in line with the Government's Principles for Scientific Advice and the Code of Practice for Scientific Advisory Committees (CoPSAC).

Expert Committee Members are independent appointments made through open competition, in line with the Office of the Commissioner for Public Appointments (OCPA) guidelines on best practice for making public appointments. Members are expected to act in accord with the principles of public life.

Further information on AQEG can be found on the Group's website at:
https://www.gov.uk/government/policy-advisory-groups/air-quality-expert-group
and https://uk-air.defra.gov.uk/research/aqeg/

Membership

Report lead author

Professor David Topping
University of Manchester

Chair

Professor Alastair Lewis
National Centre for Atmospheric Science, University of York

Members

Dr James Allan
National Centre for Atmospheric Science, University of Manchester

Dr Jo Barnes
UWE Bristol

Dr Sean Beevers
Imperial College, London

Professor David Carslaw
Ricardo Energy and Environment and University of York

Dr Chris Dore
Aether Ltd

Professor Matthew Fisher
Imperial College London

Dr Gary Fuller
Imperial College, London

Professor Roy Harrison OBE
University of Birmingham

Professor Mathew Heal
University of Edinburgh

Dr Ben Marner
Air Quality Consultants Ltd

Dr Maria Val Martin
University of Sheffield

Dr Eiko Nemitz
UK Centre for Ecology & Hydrology

Ad hoc members

Dr Eloise Marais
University College London

Dr Sarah Moller,
National Centre for Atmospheric Science, University of York and Defra

Professor Anil Namdeo
Northumbria University

Ex officio members

Central Management and Control Unit of the automatic urban and rural networks: **Dr Richard Maggs**, Bureau Veritas

National Atmospheric Emissions Inventory: **Dr Tim Murrells**, Ricardo Energy and Environment
Non-automatic hydrocarbon monitoring networks and metals monitoring network: **Dr Nick Martin**, National Physical Laboratory

Quality Assurance and Quality Control of the automatic urban network and the non-automatic monitoring networks: **Dr Paul Willis**, Ricardo Energy and Environment

Assessors and observers

John Newington
Department for Environment, Food and Rural Affairs

Roger Herbert
Welsh Government

Barry McCauley
Department of Agriculture, Environment and Rural Affairs in Northern Ireland

Andrew Taylor
Scottish Government

Alison Gowers
UK Health Security Agency

Rob Kinnersley
Environment Agency

Secretariat

Shaun Brace
Department for Environment, Food and Rural Affairs

Michelle Brailey-Balster
Department for Environment, Food and Rural Affairs

Jenny Hudson-Bell
National Centre for Atmospheric Science / University of York

Statement on integration and adoption of AI in air quality

Summary

The field of Artificial Intelligence (AI) has evolved significantly over recent years, with increasing availability of underlying methodologies and adoption across academia, the public and private sectors. This coincides with increased public discourse around AI and the potential opportunities and potential dangers posed, partly driven by the public release of facilities such as ChatGPT and the creation of national and international forums around regulation and safety.

In parallel, increasing diversity in air quality sampling and simulation technologies, combined with the need for integration with complex health and socio-economic outcomes, positions air quality science and policy as a potential user of AI tools. With this in mind AQEG considers the utility of Artificial Intelligence (AI) and associated innovation landscape whilst considering risks and sustainable mechanisms for use.

The report differentiates between Artificial Intelligence (AI) as a goal of autonomous intelligence, Machine Learning (ML) as a subset of AI that concerns development of algorithms to extract meaning from data and Data Science as the broader umbrella that encapsulates AI, ML and includes consideration of regulation, standards and ethics. Within the remit of this report and at the time of writing, the majority of air quality exemplars use machine learning (ML). Evidenced benefits of ML for air quality include development of new methods to infer contributions from different emissions to measured instrument response, to development of causal inference of policy change. In addition, service-driven industries are rapidly adopting ML tools within their data pipelining, including air quality instruments. Data driven models offer a route to mitigate traditional challenges of computational resource barriers in regional to global air quality models. AQEG foresee use of time-series forecasting from single sites or networks through to development of personal exposure estimates built around mobility data. It is widely accepted that the pathway to increased model resolution and integration of increasing amounts of data can only be met through inclusion of machine learning approaches. National research centres and laboratories are already reframing existing models with this in mind.

Despite the dominant current use of isolated ML in air quality there is significant potential offered by AI. This includes the use of Digital Twins in automated systems management and the application of Foundational Models. A Digital Twin might use near real-time data to change the state of the system being studied in advance of a desired outcome. This feedback mechanism separates a Digital Twin from a Digital Shadow, or Digital Model (as considered in a traditional modelling sense); digital twins may require support communications infrastructure, cyber security and so on. Foundational Models have rapidly entered the public domain following the public release of the ChatGPT variants by OpenAI, with parallel responses by other vendors including Google, Meta and IBM. These large models, trained on vast quantities of unlabelled data, replace task specific models and are able to adapt to more generic use cases. Geospatial Foundational Models are now used to understand the impacts of extreme weather events such as predicting the extent of flooding

and forest fires. It is likely only a matter of time before they are applied to global air pollution. This may change the balance between services offered and maintained between the public, private and academic sectors.

Quantifying the success of AI adoption requires first for a clear strategy to be defined. This should identify where existing operations might benefit from potential use of AI technologies and create a value proposition with a range of stakeholders. This co-design with stakeholders could benefit government departments for a number of reasons including building trust, maintaining sustainable partnerships and positioning government at the forefront of discussions around standards and regulations. A strategy around adoption and use should also clarify a governance structure which goes beyond awareness of the IT or digital tools to definitions of roles and responsibilities on staff, operations and relationships with AI technologies. The creation of an advisory board, for example, for cross sector partnerships around adoption of AI could be an effective vehicle to maintain an appreciation of the breadth of activity. Membership could include representations from cross government departments, academia and industry.

Whilst regulation of AI is beyond the responsibility of a single organisation, there is strength in forming such partnerships across government departments, industry and academia. This is particularly important where the state-of-the-art can change in a short space of time along with calls for wider consensus around both regulation and standards. Partnerships with academia could usefully vary from individual secondments, joint PhDs through to co-funded programmes of development across e.g. UKRI. The benefit of such arrangements includes increased external problem visibility and the opportunity for knowledge transfer around successful demonstrations of technologies and work practices. AQEG recognise the importance of engaging with the AI industry which is likely to act as a significant source of technical solutions at the environment-human health interface.

Defra and Devolved Administrations should support staff to develop the necessary skills to be aware of, use and understand AI-driven technologies for air quality. Improved training could include joint programmes with HE institutions through to tailored training options provided by industry. Nurturing AI innovation for air quality science and policy, through partnerships with external organisations, would support a longer-term goal of attracting and retaining staff with AI skill sets. As with fluid movements around regulation and standards, retaining staff in the public sector is a much bigger challenge that would benefit from a collective vision across public sector organisations and academia.

With all that in mind, the graduate workforce is likely to embrace data science as a core tool in the future. This will inevitably reduce the burden on organisations investing in targeted training. However, the need to provide an innovative and nurturing ‘AI aware’ environment will remain, with proposed activities and initiatives given in this report designed to facilitate this. By considering the issues raised in this report, we would support Defra and the DAs in their commitments to exploring the exciting opportunities offered by AI in managing and improving air quality whilst retaining and building public trust in the policy decisions that may emerge.

Contents

1. Introduction

The field of Artificial Intelligence (AI) has evolved significantly over recent years with increasing availability of data science tools and methodologies and adoption across academia, the public and private sectors. There is use of a variety of AI approaches being applied across all scientific domains. With increasing diversity in sampling and simulation technologies, air quality data and modelling output exhibit widely varying features. This varies from high resolution multivariate signals at a single location (such as a multichannel mass spectra) to satellite products or model outputs with lower temporal resolution but higher geospatial coverage. Combined with common challenges around standards, sustaining heterogeneous networks and the need to determine impacts of interventions and integration with health and socio-economic outcomes, this positions the combination of air quality science and policy as a natural target for the potential offered by AI. With the increased public discourse around AI and the potential opportunities and potential dangers posed, this statement recognises the challenges of managing adoption, integration and sustainable use of AI within Defra to support air quality improvements.

Definition of some key terms:

- **Artificial Intelligence** (AI) as a goal of autonomous intelligence, which could include data driven systems or simply rule-based systems;
- **Machine Learning** (ML) as a subset of AI that concerns development of algorithms to extract meaning and build potential decisions around data;
- **Data Science** as the broader umbrella that encapsulates AI, ML, and consideration of regulation, standards through to ethics (figure 1a).

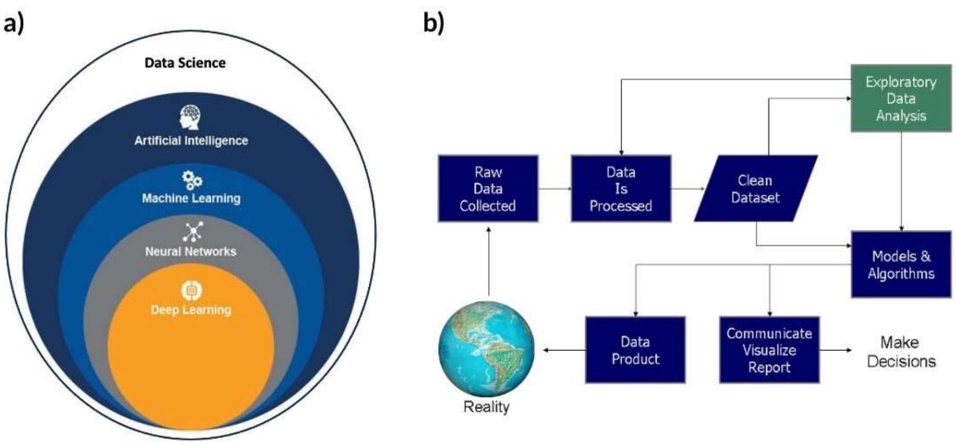

Figure 1: a) Data Science encompasses application and regulation of AI, built on Machine Learning algorithms which vary from traditional through to Deep Learning based architectures. b) The Data Science workflow, from data collection, EDA through to communications [O'Neil and Schutt, 2013].

This distinction is important when the current use-cases in air quality science are placed in the wider context of public discourse around AI. With the emergence of digital facilities such as ChatGPT from OpenAI (https://openai.com/), there is now widespread availability of AI tools not just within industry, but open to the general public at large. For example ChatGPT is a 'large model', specifically a Large Language Model (LLM) that has been expertly developed and trained around huge quantities of data. This LLM can respond to questions, appearing to offer 'advice' based on the contextual knowledge the model has been trained to. Whilst these models have been in development for some time (e.g. the model BERT was developed in 2018 (Devlin et al 2018)), ChatGPT demonstrates the impact of commercialisation for widespread use and will no doubt set the trend for emerging developments in this space. Since the release of ChatGPT, a number of industry giants such as Google, IBM and Microsoft have likewise positioned themselves to release a number of products in a short space of time. The consumer market is very fluid at the time of writing, with many digital consumer products offering AI enabled services as a positive development.

Large Language Models (LLMs) fall into a category of models known as Foundation Models, a set of tools we will revisit shortly when discussing potential future uses of AI with regards to air quality research and development (section 2). These tools represent the forefront of AI development, requiring certain levels of resourcing and skill sets to train and deploy. Despite these barriers, one would expect migration of such tools for use across the broader scientific spectrum following the typical technology 'hype cycle'. Published every year, the Gartner hype cycle (figure 2) positions the state of a particular technology from its point of inception [left hand corner] through to widespread adoption [right hand side], passing through a peak of expectation before the true potential is evaluated. Figure 2 presents the 2023 Gartner hype cycle which places Foundational Models towards the peak of expectations and projects a time to reach widespread adoption between 2 to 5 years. This includes use by the research community.

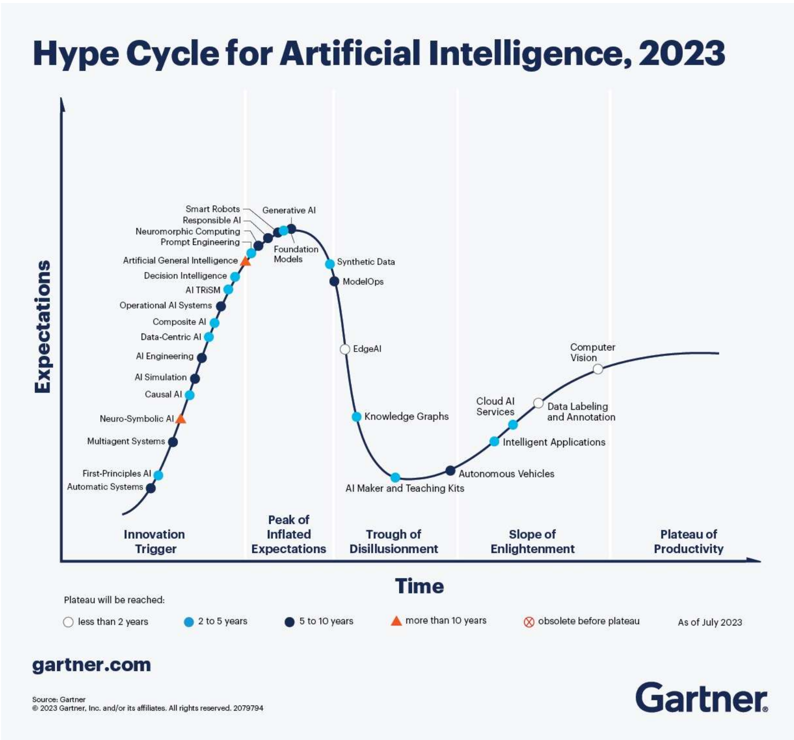

Figure 2: The 2023 Gartner Hype Cycle for Artificial Intelligence

Even without explicit use of Foundation Models by the wider research community these tools are already acting as aids in scientific workflows. For example, Github and Microsoft have released their own 'CoPilot', a tool to aid and generate code based on natural language requests which is significantly reducing the barrier for software development (e.g. https://github.com/features/copilot). This will inevitably accelerate the creation of new software and data analytics tools, in both academia and industry.

The key distinction within the remit of this report is that the majority of air quality research, at the time of writing, showcases the use of machine learning (ML), which is a subset of AI. Machine Learning represents the act of training algorithms to predict or detect metrics of interest and 'learn' how features within a dataset (e.g. concentrations of different pollutants or channels in a mass spectra) relate to each other. For example, one might build a ML model for predicting concentrations of NOx as a function of time and traffic levels (regression), or assign an emission source to a measured mass spectra (classification). Machine Learning is an essential component of autonomous intelligence but does not by itself ensure this is achieved; a machine learning algorithm is developed/trained and then interpreted by an expert within their domain who may then create further information to build a decision or action around. A ML algorithm may be used *automatically*, for example to correct an instrument response into a concentration of pollutant onboard the instrument, but autonomous intelligence can act and dictate without our intervention. In later sections of this report there are instances of AI that are of direct relevance to air quality management.

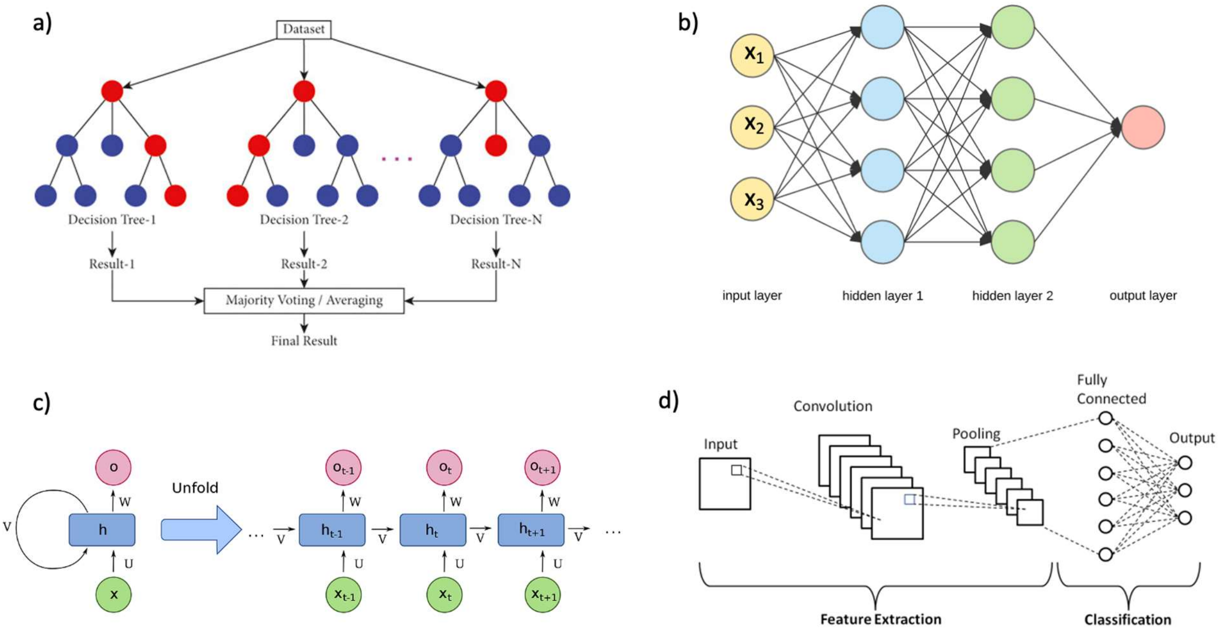

Figure 3: Example architectures of different machine learning frameworks - a) Random Forest [Khan et al 2021], where individual decision trees are built and then combined to produce an 'optimal' outcome, b) Deep Neural Network [modified from Ognjanovski, 2019] where an input array (left) is connected to a single output variable (label or number) via a series of connected layers, c) Long Term Short Term Deep Neural Network [e.g.https://www.techtarget.com/searchenterpriseai/definition/recurrent-neural-networks] where information in the sequence is exploited, d) Convolution Neural Network [e.g. modified from Fan and Truong 2022], where an 2D pixel matrix (left) is connected to an array of object labels (right) through a series of convolutional layers .

There are various families of ML algorithms. These include, for example, tree based methods through to deep learning (Figure 3). Choice of method depends on the question being addressed and the amount of data available. This can also be dictated by the need to understand what relationships have been learned by the ML algorithm and why. For example, 'explainable AI' refers to the development of methods that aid this interpretation, driven by the growing need to ensure provenance in decisions made [Xu et al. 2019]. Deep learning became prominent across scientific research in the early 2000s as Graphic Processing Units (GPUs) were exploited for faster training to much larger quantities of data. For example, a deep neural network is a neural network with more than one layer (Figure 3b); one can imagine each node (circle) on the left hand side of Figure 3b represents an ambient measurement and the single node on the right hand side the concentration of $PM_{2.5}$. In this case the end goal may be to predict $PM_{2.5}$ in the next hour as a function of weather and concentrations of other pollutants at the current time. Each node, and connections between the nodes, have parameters that are optimised through a training process. The number of layers, the number of nodes per layer and such parameters are known as hyperparameters. If one was to replace the right hand node with the number of hospital admissions within the local area, or classification of 'hazardous conditions' for example, a expert user might start to connect data from multiple domains. Another family of deep learning methods, shown in figure 3c, accounts for the sequence of information provided.

Thus the first example could be extended to predict $PM_{2.5}$ in the next hour as a function of measurements taken over the last 6 hours. Computer vision techniques including Convolutional Neural Networks (CNN) (Figure 3d) have revolutionised several fields. CNNs are deep neural networks that learn how localised features within an image translate to key properties of that image. These can be used to downscale remote sensing products, for example, treating 2D fields of total column variables and meteorological fields to generate predictions at ground level. This report provides a brief synopsis of peer reviewed demonstrators of relevance in section 2.

Overall, evidenced benefits of ML for air quality research include development of new peer reviewed research software tools designed to work across a number of scales; from new methods to infer contributions from different emissions to measured instrument response Lin et al. 2022], to development of causal inference of policy change [Song et al. 2022]. In addition, service driven industries are rapidly adopting ML tools within their data pipelining, including air quality instruments [Grant-Jacob and Mills, 2022].

It must also be recognised that, more broadly, Data Science encompasses all facets of AI, ML through to best practices (Figure 1a). Scientific research has always relied on data, but there is now more focused attention paid to key components in the lifecycle of a data product from collection through to end user delivery. This includes training data scientists and engineers in Exploratory Data Analysis (EDA) techniques with the aim of identifying outliers, for example, and more broadly constructing an interpretation on the state of the data collected. Whilst most of the public discourse tends to be on ML algorithms and AI demonstrators, such as ChatGPT, there is an important AI ecosystem which is being mapped into academic training programmes and service industries. For example, Data-Centre AI focuses on enhancing and enriching training data to drive better AI outcomes; Prompt Engineering concerns providing text or image inputs into generative AI models to optimise the response; Edge AI refers to the use of embedded AI on IoT endpoints such as mobile phones or autonomous vehicles etc. Perhaps more importantly, Data Science also brings attention to regulation and standards. With these developments in mind, one must recognise the importance of the wider ecosystem of digital tools, staffing and services to facilitate sustainable use of AI which are discussed in turn.

In the following report AQEG briefly covers the documented uses of ML and potential uses of AI as relevant to air quality research and management (Section 2). This is not intended to provide an exhaustive list of all activity in this space, but rather to provide referenced demonstrators. Emphasis is then placed on the required adoption of standards and governance to ensure trust in any adopted methodologies (Section 3). Whilst recognising that most, if not all, developments evident in the literature focus on applications of ML, there is no doubt that AI will infiltrate current ways of working as per the aforementioned hype cycle. Later AQEG consider a possible implementation strategy that requires further scoping and co-design with relevant stakeholders with regards to internal operations (Section 4). Explicit adoption of technologies held under the banner of AI will require understanding dependencies on data streams, data preservation and access to appropriate computing services which can be constructed through subsequent architectural designs.

2. Existing and potential use cases

Growing evidence in the scientific literature already demonstrates the potential uses of machine learning (ML) in air quality research, from real time data integration through to regional and global modelling. Whilst there is no clear adoption of these tools in operational air quality forecast services, or even standardised approaches for data analytics used across the community, it is likely that developments will move towards adoption as standard. With this in mind, the following section discusses specific areas of evidenced use that could be of interest to Defra and the DAs, and highlights future developments that are likely to emerge across the landscape of AI innovation, including Foundational Models and Digital Twins. Subsequent requirements of the supporting ecosystem, from standards to training, are described in section 3. Potential use cases are for measurements and modelling are included, with some discussion of the potential role for integrating unstructured data from social media, public interactions through to policy reviews. This is not designed to act as an exhaustive list of all activities, but provides relevant examples where the size and characteristics of generated data vary significantly, from individual sensors that provide high frequency time series datasets to hybrid global models that produce huge volumes of 3D geospatial data.

2.1 Measurements.

Generating insights from measured and laboratory-controlled data requires a combination of statistical and, often, ML methodologies built on top of robust EDA protocols. Modern sampling methods used in air quality monitoring can produce huge quantities of multivariate data including high resolution mass spectral through to optical signatures, including scattering and holographic images.

Classification:

Several ML methods are used to infer contributing sources from measured signatures that fall under the remit of unsupervised, supervised and semi-supervised machine learning. Briefly, unsupervised learning algorithms are designed to extract patterns and meaning from unlabelled data. For example, ambient time series of pollutant concentrations can be fed into unsupervised clustering algorithms to determine distinct contributions such as wood burning, diurnal patterns and so on. Supervised learning algorithms are trained to detect known patterns, or labels, once trained to that labelled data. For example, we might use a neural network to classify a scattering image into known types based on controlled laboratory studies. Those algorithms can then be applied to ambient data. Supervised algorithms, particularly neural networks, often require large quantities of data to build acceptable levels of accuracy. As generating controlled data can be a challenge, semi-supervised methods can be used to learn relevant underlying features from much larger quantities of ambient data before being tuned to the small amount of labelled data available [e.g. Gilik et al. 2022]. Combined with controlled laboratory studies, these methods have been used to quantify changes to measured signals and thus contributions from different emission sources. In Example 1 the use of supervised deep learning algorithms is applied to real time holographic images (as a classification problem). In Example 2 the use of supervised tree based methods is used to build a model to predict pollutant concentrations as a function of meteorology and time (as a regression problem), with the aim of applying weather normalisation.

Example 1: Detecting and classifying bio-aerosol

The Swissens Poleno uses a combination of classical image analysis and a deep neural network algorithm to identify a range of pollen taxa. Deterministic criteria based on the shape of the particles are used as a first step before images are processed through a trained Convolutional Neural Net (CNN). The figure below provides a schematic of both stages, showing that the architecture of the CCN used was based on an industry standard known as VGG16 which has 16 hidden layers. The pretrained industry standard VGG16 model was trained to millions of images and is used for several object detection tasks [Simonyan and Zisserman, 2015].

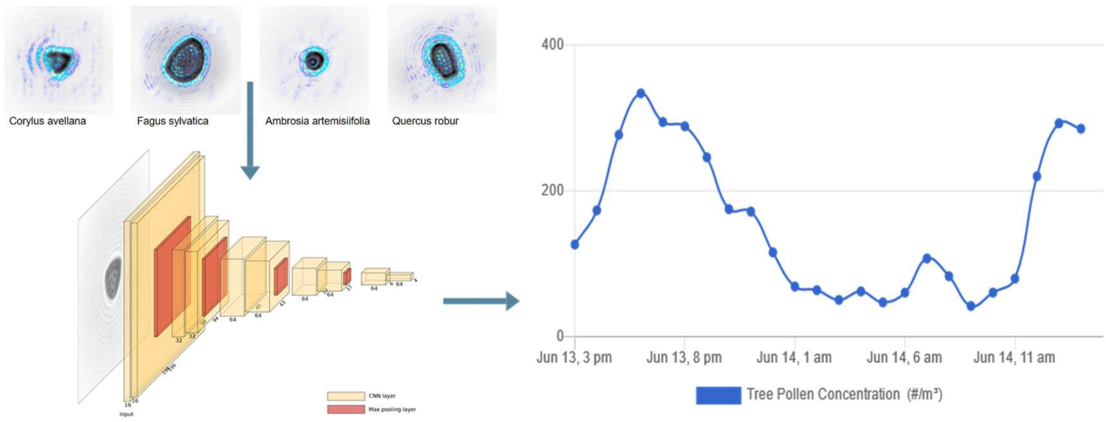

Figure 4: Classifying pollen particles through deep learning leading to generation of real time concentrations [Sauvageat et al, 2020].

Weather normalisation and changepoint detection:

Attributing changes in measured pollutant concentrations to natural conditions or technical interventions often requires weather normalisation combined with change-point detection. Weather normalisation concerns the process of fitting a model to predict a measured pollutant concentration as a function of meteorological variables and time. This is an example of ML regression where one temporal feature includes a trend term (e.g. unix time). By sampling the fitted model around typical conditions for a given time of day/year, but keeping the trend term, one can create a 'clean' dataset from which changes in the underlying trend can be evaluated. Attributing changes in the trends to policy interventions, for example, can be done through expert interpretation or through changepoint detection methods. Even with a changepoint detection method, expert interpretation is often needed to attribute detected points to expected changes.

Example 2: Policy intervention detection.

Figure 5 shows the measured (left) and weather normalised (right) NO_x and NO_2 at London Marylebone Road between 1997 and 2016. The weather normalised values were calculated by 50 random forest models (for each pollutant). The vertical lines show the changepoints identified by structural change analysis [Grange and Carslaw, 2019] which were found to match known interventions. The authors also note, in many cases there may not be sufficient information or metadata to help explain the changes observed.

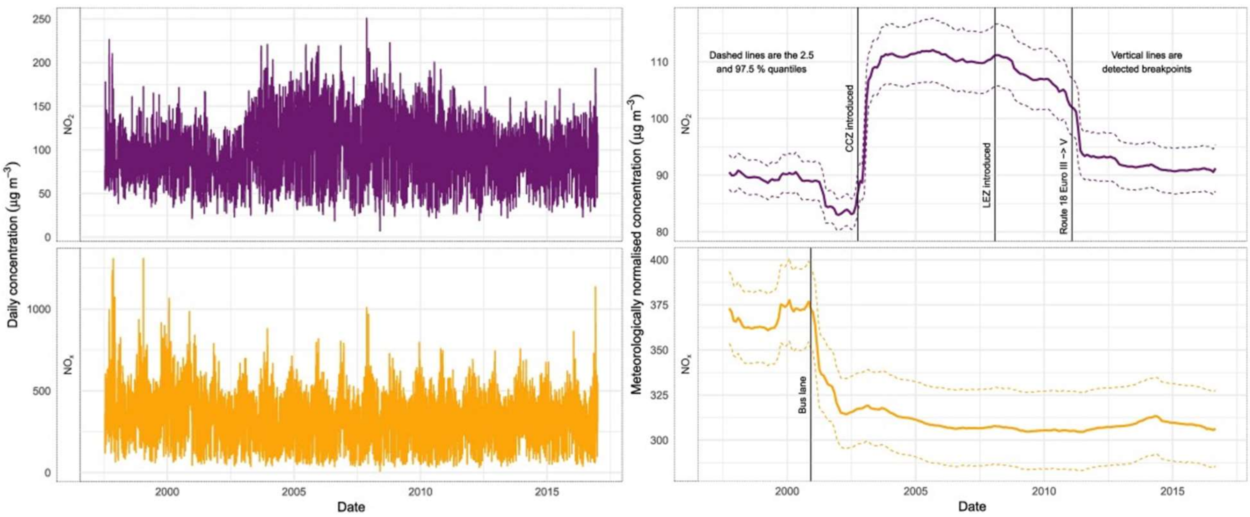

Figure 5: Raw times series for both NO_2 and NO_x (left) compared with the weather normalised concentrations (right) [Grange and Carslaw, 2019].

2.2 Models.

Traditional numerical models of air quality are built on underlying parameterisations of pollutant emission, along with physical and chemical processes that simulate dispersion and evolution. A translation of that science through to code is then coupled with a strategy for solving coupled and decoupled equations to deliver predictions of pollutant concentrations over space and time. The computational challenges of solving many types of equations and tracking concentrations across a 3D space, for example, can lead to a range of model scales from street level through to global. These scales can likewise be coupled in a variety of ways to improve broader understanding of regional to local drivers. These models have supported decades of regulatory assessments, supported by a huge collection of scientific studies investigating the impacts of newly identified physical processes, emission sources and so on. The challenges of developing and running these models include access to high-performance computing (HPC) systems and wider understanding of the software engineering required to deploy, maintain and interpret these tools. Data driven models offer a route to mitigate some of these challenges by utilising developments in open-source programming environments and accessible computing hardware. For example, this includes time-series forecasting from single sites or networks. Data driven models also offer a route for including data that is pertinent to the air quality outcome but would otherwise be difficult to include within traditional frameworks. For example, this includes development of personal exposure estimates built around mobility data. Even with access to HPC resources, model frameworks are having to change to meet requirements of high resolution. Figure 6 displays a schematic from a seminal review of next generation earth system models [Reichstein et al. 2019]. The key message here is that the path to increased resolution, fidelity and integration of increasing amounts of data can only be met through inclusion of machine learning approaches in regional to global models. Moving towards development of AI for air quality modelling and management we see potential in the use of Digital Twins in automated systems management and the rise of Foundational Models.

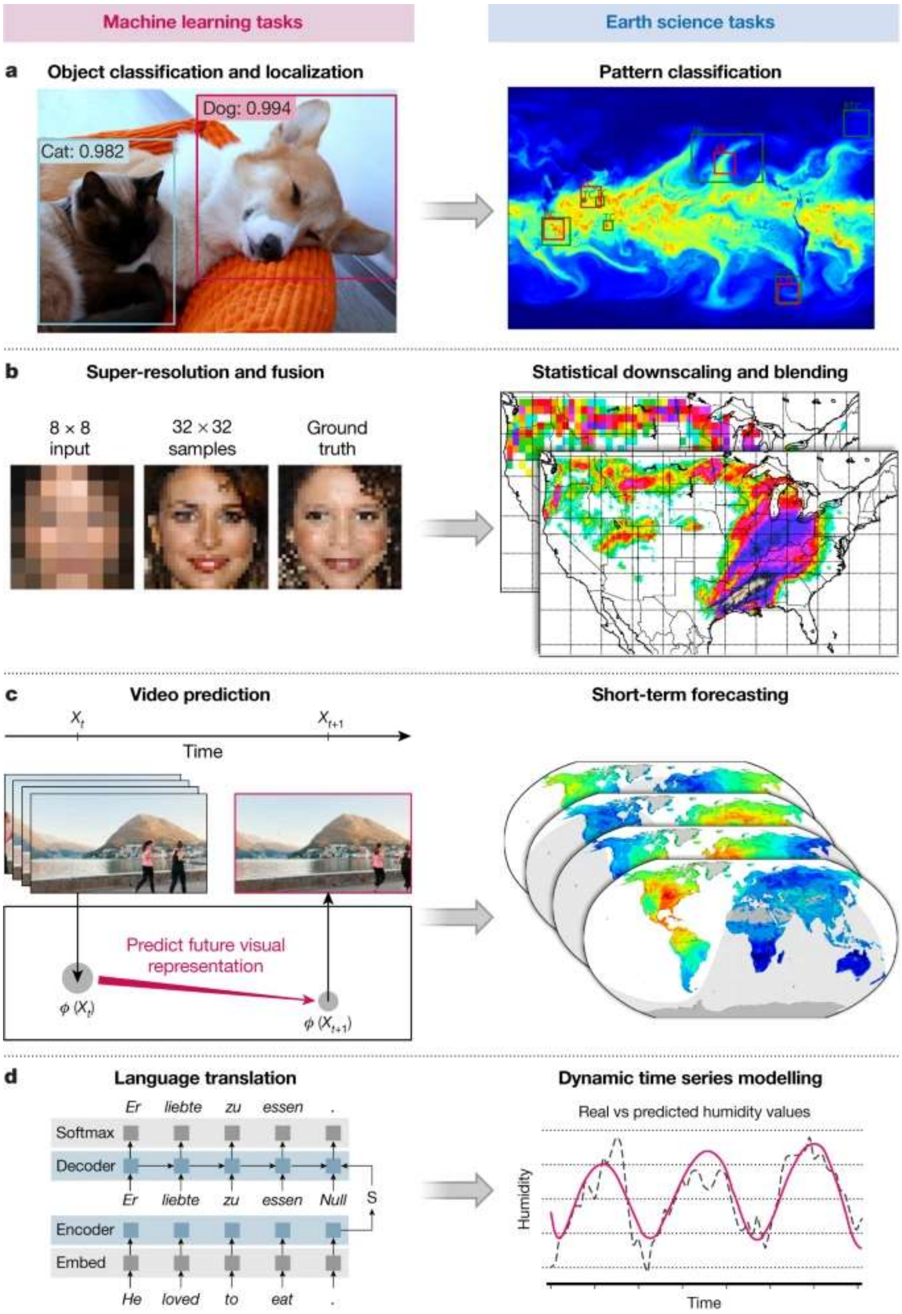

Figure 6: Four examples of typical deep learning applications (left panels) and the geoscientific problems they can be applied to (right panels) [Taken from Reichstein et al. 2019].

Time-series forecasting

Time-series forecasting methods explicitly account for information held within the sequence of measured data. These are typically built and optimised to measured or synthetically generated data, rather than on an architecture and model parameters built on underlying physics or chemistry. The most common use case in air quality is the development of time-series forecasting tools. Combined with best practices around data cleaning and preprocessing, there is evidenced potential for predicting future change from measured data

[Liu et al. 2021], which includes differentiating the impact from changes in weather versus transport interventions [Sulaimon et al. 2022].

Personal exposure estimates

Agent-based models have emerged across a number of scientific domains. These models create agents that have pre-determined characteristics and thus ‘interact’ with other agents and a model environment. These models are clearly relevant to predicting personal exposure under different microenvironments, provided data on mobility and behaviour response can be generated and then used to fit the relevant model. In response to the COVID19 pandemic, for example, the SPENSER model was used to predict where people might spend their time by LSOA [Spooner et al. 2021]. Once mapped to pollutant fields and concentrations in different environments, estimates of total exposure can be created.

Example 3: Predicting personal exposure by microenvironment.

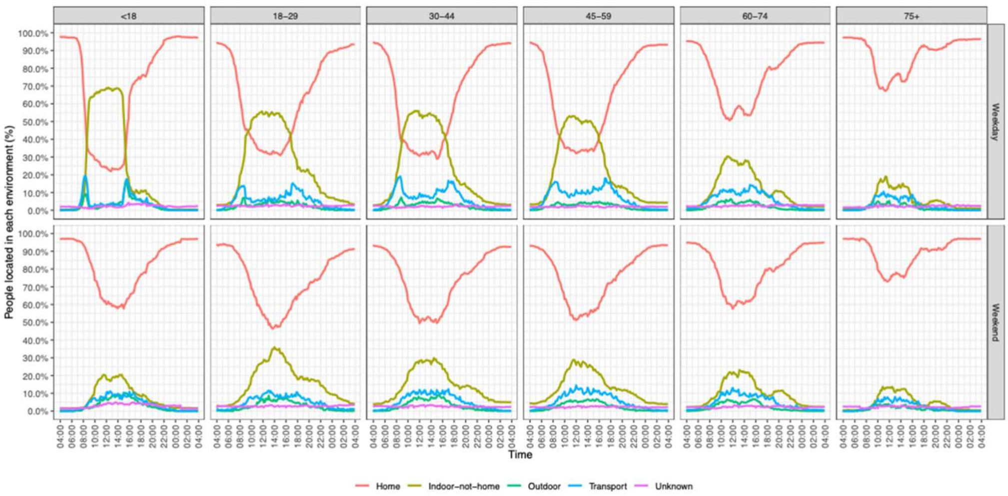

Figure 7: Proportion of time spent by simulated individuals from Greater Manchester within different micro-environments (Home, Indoor-not-home, Outdoor, Transport and Unknown locations), by age and gender [Thomas et al. 2022].

The Data Integration Model for Exposures (DIMEX) integrates data on daily travel patterns and activities with measurements and models of air pollution using agent-based modelling to simulate the daily exposures of different population groups [Thomas et al. 2022]. Figure 7 shows predictions of time spent in different microenvironments. When combined with diurnal patterns of pollutant concentrations in those environments, these models are able to generate predictions of total exposure.

Downscaling

Integrating remote sensing data with model and ground-based data offers the potential to provide estimates of air quality in poor sampled areas. There are a number of challenges in performing such an integration, not least related to the variable characteristics each data contains. This includes sparse ground based measurements with high time resolution to remote sensing products with a daily total column observation. Nonetheless, computer vision techniques including CNNs are often used to develop regression models that are trained to predict concentrations as a function of land-use, remote sensing product and meteorological fields. There is also use of random forest models with the same aim in imputing data in areas with poor sensor coverage. Downscaling can also be applied directly to model output with the aim of increasing the resolution for a fraction of the cost of running very high resolution models.

Example 4: Downscaled predictions of atmospheric composition

Geiss et al. (2022) used CNN based super resolution to downscale atmospheric chemistry simulations using physically consistent deep learning. Single-image super resolution (SISR) artificially enhances the resolution of images after they are captured. Following the training of a CNN for downscaling, the idea is that a user can generate coarse simulations at relatively low computational cost. Model ensembles are an ideal use case for this application, where a large number of high-resolution simulations is too resource-intensive to complete. In this study the authors found SISR methods that incorporated high-resolution climatological data performed significantly better than those that did not when compared with high resolution models.

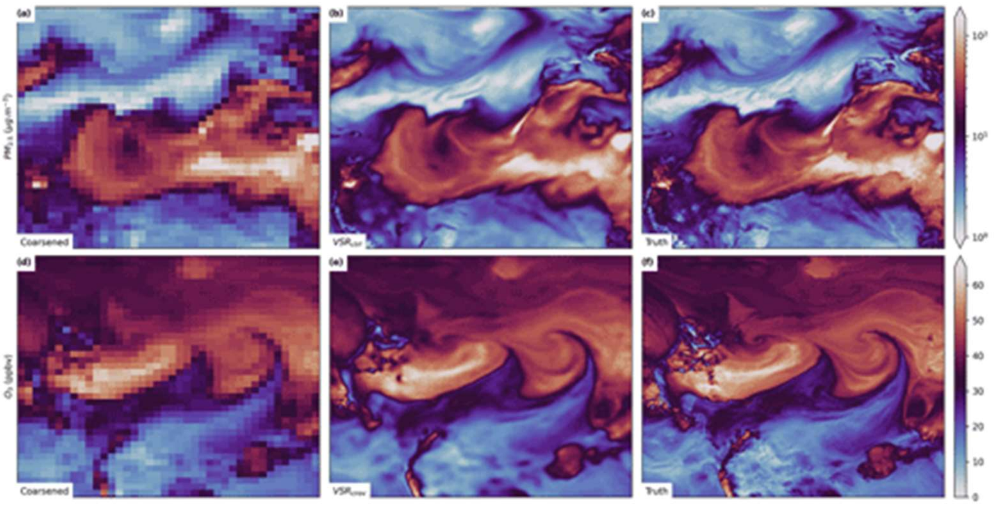

Figure 8: Predictions of $PM_{2.5}$ (top row) and O_3 concentrations (bottom row). The left column shows the coarsened data, the middle column shows the super-resolved output from the best-performing CNN, and the right column shows the 'ground truth' from the high resolution model [Geiss et al. 2022].

2.3 Emerging technologies

Physics-informed machine learning

Retaining the numerical basis of existing global models whilst increasing the physical/chemical complexity they represent requires improvement in computational efficiency of the solution process. It is common to integrate a complex process into existing geospatial models through the process of parameterisation. In emerging numerical model developments traditional numerical methods are replaced with hybrid process level-ML frameworks to mitigate this challenge. With concerns around explainability of AI systems, there is movement towards applying physics informed machine learning [Karniadakis et al 2021], where the underlying physics/chemistry of the problem at hand constraints and improves the behaviour of the data driven component (e.g., Neural Ordinary Differential Equations). This offers significant potential to inform the development of next-generation regional models, for example, where heterogeneous data streams and environmental/urban infrastructure could be used to improve forecasts of emerging events. These hybrid “physics informed” approaches could improve trust in new predictive systems and allow us to adapt to emerging data streams.

Digital Twins

Digital Twins are emerging as the next generation of digital simulation tools across multiple domains. Whilst existing numerical models allow a user to determine the potential concentration changes of e.g. an emission reduction, or data driven time-series forecasting tools to predict potential impacts of a traffic intervention, a Digital Twin would use near to real-time data streams to change the state of the system being studied in advance of a desired outcome. The key distinction between a Digital Twin and Digital Shadow (see Figure 9) is this feedback mechanism. Digital Twins have been used predominantly in applications of engineering up to this point. Despite evolving narratives around potential use of Digital Twins, there are some clear potential use cases. This could include, for example, optimal management of traffic systems to reduce concentrations of pollutants, or optimal control of building ventilation systems to reduce personal exposure from indoor-outdoor sources. A Digital Twin is thus built around feedback loops to build automated decision support systems, requiring consideration of real time communications and data security (see section 3).

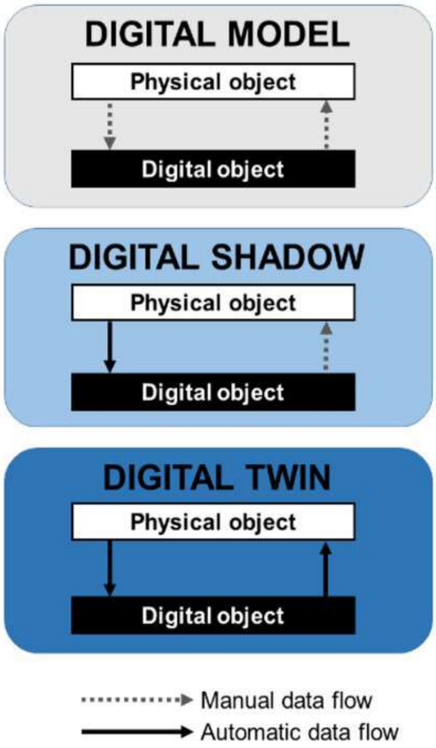

Figure 9: Data flow in digital models, digital shadows, and digital twins [Nikula et al 2020].

Foundation models

Foundation Models have rapidly entered the public domain following the public release of the ChatGPT variants by OpenAI, with parallel responses by other vendors including Google, Meta and IBM [https://research.ibm.com/topics/foundation-models]. These large models, trained on vast quantities of unlabelled data (Figure 10), replacing task specific models and are able to adapt to more generic use cases. More specifically, in the discussion of ML models up to this point, normally an input is provided with a specific output as the task (e.g. predicting $PM_{2.5}$ classification as a function of optical image). The data used to train ML models are restricted to certain types, for example a collection of ambient or lab generated optical images. Foundational models on the other hand can train on vast amounts of unlabelled data with varying types. The goal is that, through self-supervised learning, these frameworks begin to learn patterns and relationships within that data. From this learning they are then able to respond to multiple queries. This includes several tasks such as responding to user requests for summaries of large quantities of published evidence or augmenting the software development processes through automatic code generation. Geospatial foundational models are now being developed to ingest remote sensing data to underpin efforts to understand impacts of extreme weather events, such as predicting the extent of flooding and forest fires [https://research.ibm.com/blog/geospatial-models-nasa-ai].

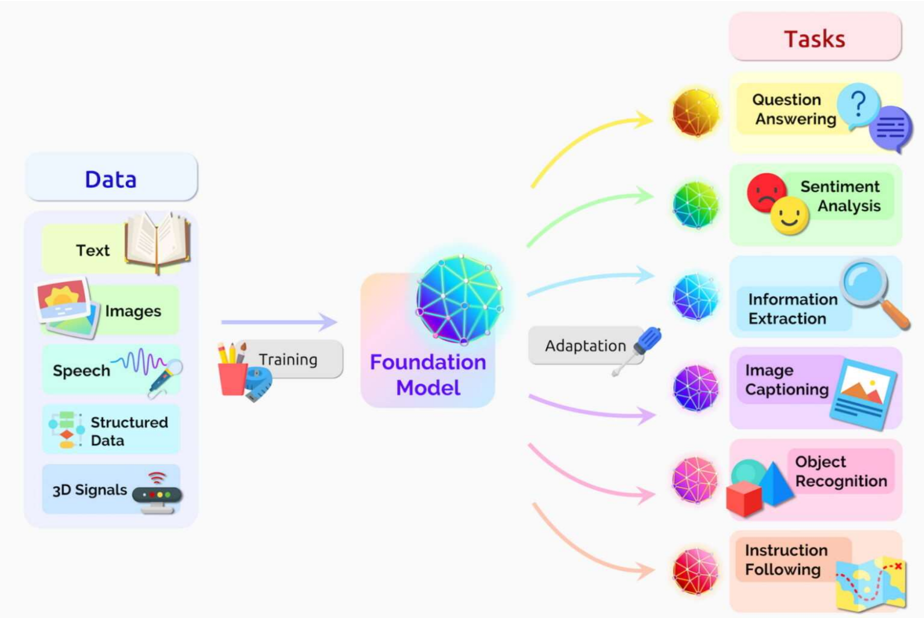

Figure 10: A foundation model can centralize the information from all the data from various modalities. This one model can then be adapted to a wide range of downstream tasks. [Bommasani et al. 2021].

2.4 Behavioural change, reporting and policy development.

Aggregated air quality data can exhibit the three characteristics of 'big data': large volume, high velocity and large variety. More specifically, existing monitoring networks could enable data sharing and ingestion from high resolution multivariate instruments. This data could be coupled with data streams from networked IoT devices, remote sensing platforms and growing focus on air quality policies across social media platforms. The combination of each data stream is of course dependent on the question being answered, where individual projects may focus on each in isolation (see above). However, there is potential for a collective understanding and improved outcomes by considering the combination of unstructured data in detecting responses to and perceptions of interventions or naturally occurring events. Natural Language Processing (NLP) paired with social media analysis can detect behavioural changes and reactions to environmental stressors [Hodorog et al. 2022]. This can include monitoring local responses to a number of interventions. These tools have evaluated public reactions to natural hazards [Vongkusolkit and Huang, 2021], demonstrating their utility in adding supporting information to observations of evolving real time events. Studies on air quality using NLP have shown the capability to connect citizen observations with air-quality variations [Juanals et al. 2018]. This includes an assessment of sentiment responses in social media, which may provide valuable insights into how communities respond to interventions and adapt to evolving environmental conditions.

3. Implementation Strategy:

As developments and demonstrations of AI-machine learning within air-quality research and services continually evolve, such tools are likely to become a core component of the standard air-quality data lifecycle. As already presented in section 2, there are demonstrable benefits across academia and industry around adoption of machine learning.

An implementation strategy for any group or organisation lays the ground for integration and adoption of AI, whilst being agile to developments from a range of sectors and requirements around standards and regulations.

The initial step in any implementation strategy is to iteratively create a strategy around adoption and use. This strategy requires identifying where existing operations will benefit from potential technologies and creating a value proposition with a range of stakeholders. Some of those benefits are listed in section 2. An approach of co-design with Defra stakeholders is important for a number of reasons, including building trust, maintaining sustainable partnerships and positioning Defra at the forefront of setting standards. Co-design mechanisms could form a number of specific auditing and workshop exercises. For example, the US Food and Drug administration has issued an initial discussion paper to communicate with a range of stakeholders in its drive to develop an ‘agile regulatory ecosystem that can facilitate innovation whilst safeguarding public health’ [FDA, 2023a].

The following sections reflect on the fluid state of AI regulation (section 3.1) before commenting on the value of partnerships across government, academic and industry to develop several activities that would improve internal awareness and involvement in the AI landscape (section 3.2). This is then aligned with a discussion of training needs and sustainable use (section 3.3).

Beyond the areas briefly covered in sections 3.1 - 3.3, a strategy around adoption and use should also clarify a governance structure which goes beyond awareness of the IT or digital tools to definitions of roles and responsibilities on staff, operations and relationships with AI technologies. There are several international models of governance that would be useful to review, an example being the Singapore AI governance structure [https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI-Governance-Framework] that covers development of an advisory council, research programme and providing guidance to businesses in adoption of appropriate best practices.

3.1 Regulation:

Regulation of AI, and thus implementation of ML, is a fluid area of development both nationally and globally. The UK hosted the AI safety summit in November 2023 which aimed to bring together international governments, AI companies, academics and civil groups to consider the risks of AI and appropriate mitigation. There have been mixed reviews surrounding the outcomes of this event (https://www.theguardian.com/commentisfree/2023/oct/31/rishi-sunak-ai-safety-summit-tech-challenges), but the narrative around safety and this regulation is becoming more widely presented. The latest McKinsey Global Survey on the current state of AI [McKinskey, 2023] shows many organisations are not yet addressing potential risks from generative AI, with 21% of survey respondents noting their organisation had now established policies governing employees use of generative AI technologies in the workplace. This might cover; from generating reports, developing technology roadmaps and even automating data analytics. This will be partly linked to a lack of understanding on potential uses of such technologies and the speed at which solutions can be built around them. Previous reports published by the AI Council note that the UK will only feel the full benefits of AI if all parts of society have full confidence in the science and the technologies, and in the governance and regulation that enable them [AI Council Roadmap, UK]. Whilst conversations around regulation are often associated with assumed restrictions on innovation [e.g. https://www.forbes.com/sites/jackkelly/2023/06/05/artificial-intelligence-is-getting-regulated/?sh=5c96f6fb7a09], there are growing calls to clarify this does not have to be the case [Ada Lovelace Institute, 2021].

From the perspective of air quality there is benefit to be had by joining relevant conversations around regulation for internal and external use, even if the responsibility for setting regulatory procedures is positioned at both higher levels of government and different contributors to the air quality data lifecycle (e.g. instrument vendors, research modellers etc). Given the evolving public perceptions around the use of AI in particular, there are a series of mechanisms by which Defra might look to champion both regulation and innovation jointly with other government departments and with the research community. It is particularly important to support the development of innovative, safe, technological solutions to improving outcomes from air quality management given the social drivers surrounding the science of air quality. One might consider a future example of Digital Twin of a transport system, designed to improve air quality, inadvertently increasing emissions through low income neighbourhoods. An appropriate analogy around access and adoption of digital tools (not just AI) includes challenges around the performance of low-cost sensors, an ongoing area of research and development. This may require annual reviews of sampling and modelling technologies. Using the US FDA as an example, a recent report on medical devices notes that '*As of October 19, 2023, no device has been authorized that uses generative AI or artificial general intelligence (AGI) or is powered by large language models.*' [FDA, 2023b]. As far as one can understand, there is no clear review of relevant air quality sampling technologies and services at the time of writing.

Even if a strategy is focused solely on maintaining awareness of AI technologies, external changes in the underlying data sources (e.g. monitoring networks and model outputs) and supporting data ecosystems (data curation through to standards) to match AI regulation have a global scale, far beyond the responsibility of a single organisation. For example, concerns around the provenance around 'pure machine learning' based models has led to development of physics informed machine learning models to ensure provenance in decisions made [Karniadakis et al. 2021, Kashinath et al. 2021]. In these frameworks, a machine learning model is constrained by the underlying physics of the system. Adoption of these tools requires the research community to redevelop software tools to meet these requirements. With this in mind, both academia and industry are adopting a range of technologies that will impact both data and modelling tools available for use in Defra and more widely in the consulting and academic sectors.

Whilst regulation is an agile area of development there is strength in forming partnerships across government departments, industry and academia to better understand and inform AI regulations around the current state of monitoring and modelling technologies.

Alongside future regulations that will set mandatory requirements around the use of AI, there are ongoing conversations around developing standards. Technological standards from an air quality management perspective are not new; a previous AQEG report (PM measurement technologies, 2024] summarised the development and adoption of e.g. INSPIRE standards through regulatory monitoring networks [INSPIRE], with a recognition of ongoing challenges around components such as meta-data standardisation. AQEG recognise that technological standards surrounding AI are equally important, covering areas including data interoperability, model evaluations and so on. However, non-technical standards are equally important if somewhat less obvious when planning implementation of technical solutions. This includes ethical standards and social-cultural standards [e.g. Auld et al. 2022]. These aim to ensure the benefits of AI respect cultural nuances and ensure benefits are realised regardless of socio-economic status and cultural background. Whilst there are no universally accepted standards that work across all domains, AQEG would recommend a clear commitment to co-designing plans that move towards clarity around the development and use of relevant standards internally and through Defra managed networks.

As regulation and standards continue to evolve, there are several components of the air quality data lifecycle that Defra could act to support/champion in readiness for these and further AI developments. These are briefly listed below under data and software accessibility, data and software protection and method evaluations. These are highlighted as essential to developing a robust platform on which to develop partnerships around regulation and innovation as mentioned previously.

- **Data and software accessibility**: There have been significant advances in open data access. This is, in part, forced by a requirement to make data open access after a limited embargo period if generated through public research funding. Whilst there is no guarantee the same can be found for privately managed networks, there has been a cultural shift to adopt open science practices that include generating open software and releasing open data, supported by a range of open licensing options. FAIR data are data which meet principles of findability, accessibility, interoperability, and reusability [Wilkinson et al. 2016]. FAIR principles for software have been proposed [Barker et al. 2022]. Both are examples of initiatives that encourage open collaboration but are also key components in ensuring trust in actions made through adoption of data or implementation of software. Making data and software open access is perhaps the first step in ensuring readiness for incoming standards and then regulations around AI. There has been a ‘reproducibility crisis’ within the field of AI, driven in part by lack of data and software sharing [Ball, 2023].

  Whilst providing open data and software is an important first step, it does not guarantee it is either discoverable or usable by other end users [Weerakkody et al. 2017]. Here we see the importance of maintained data platforms and clear meta-data that enables other users to understand the remit of the provided product. In other disciplines, particularly the biosciences, there has been development of workflow management software. These are tools that allow an end user to not only visualise the steps used to generate the data, or run a model, but to repeat that process through a series of interactive steps. Defra may wish to conduct an audit of available datasets already in operation, including an evaluation of alignment with FAIR principles, and mapping those on to existing processes. This would allow an assessment of further development and support needed to improve readiness for the use of such data in the AI and air quality landscape.

Machine learning tools are built on data that represents a given system or process. With regards to air quality, this could include a mass spectral dataset of multiple emission sources or 3D model output from a regional model study on reduction in ammonia emissions. Defra could also help champion the need for better availability, and maintenance, of labelled datasets. This refers to the process of labelling, or tagging, a subset of a dataset that represents a source, change or fingerprint. In the wider AI community these central datasets are a key feature in the innovation landscape, with annual competitions based and judged against these facilities. At the time of writing there are no clear repositories that provide labelled air quality datasets on which to build machine learning tools that have clear provenance. This is partly driven by the nature of research funding. Championing and supporting such a facility would position government as a key enabler in the AI and air quality landscape. This is referenced within section 3.2 and 4 on partnerships and potential measures of success.

- **Data and software security**: Reference has been made to several emerging technologies, including Digital Twins. These tools could, in principle, connect incoming data streams to automated decisions around, for example, traffic flow or building ventilation. Several studies have now inferred the influence of 'bad actors' in some countries where networked data is used to influence resource management. Where there is dependency on near real-time data to automate decision support systems, one can imagine a number of scenarios that require stringent cyber security to be in place. This is a complex area, but would likewise benefit from an audit of existing monitoring systems in place and the range of cyber security standards. Detecting the influence of bad actors on decision support systems can also be influenced by choice of machine learning algorithm on which such systems are built. Some algorithms are easier to interrogate than others and thus discover the influence of bad/suspect data. For example, this includes the ability to relate changes in a forecast property (concentration) to previously sampled data which can then be verified separately through expert interpretation.

  Trusted Research Environments (TREs) are data storage and analytical environments that host sensitive data. A federated TRE landscape exists across the UK, where individual patient data, for example, are maintained by separate health authorities. There are wider national movements around standardisation of TREs (SATRE), with emerging open standards that define whether a system meets the criteria of a TRE. Health data cannot leave a TRE; rather ancillary data has to be ingested into a TRE for subsequent analysis. For example, this could include ingestion of environmental data within a TRE to determine the importance of residential postcode on disease outcomes. Whilst environmental data by itself does not require a TRE, we can imagine a future scenario where air quality data could be used to infer human behaviour. This is particularly evidenced in the indoor air quality domain where concentrations could be combined with other products to determine where an *identified* individual spends their time. This may require the adoption of TRE standards within the air quality space which, in turn, requires an understanding of air quality data characteristics that would align with a TRE environment.

It is appropriate to also consider the ethical implications around the use of AI. In all domains this covers issues around bias, fairness, transparency and privacy. Within the domain of air quality management this includes mitigating impacts of decision systems that create problems in low-income neighbourhoods, for example. We know that digital literacy is a heterogenous problem across the UK [ONS, 2019], creating a dependency on access and interpretation of tools developed. This may also align with sparse coverage of tools and technologies in poorly monitored areas. All of these factors should be considered when developing a strategy for AI adoption.

3.2 Collaboration, Partnerships and Stakeholder engagement:

The potential utility of AI across wider environmental and public sector bodies presents an opportunity to form lasting collaborative partnerships and share lessons learned. This is particularly important in a rapidly evolving environment where the state-of-the-art can change in a short space of time and there is a need for wider consensus around regulation and standards. This could extend to formation of partnerships with international allies.

Partnerships with academia could usefully vary from individual secondments, joint PhDs through to co-funded programmes of development across e.g. UKRI. The benefit of such arrangements, though varying in size and capacity, is placing Defra as the problem holder but also offers the opportunity for knowledge transfer around successful demonstrations of technologies and work practices. This also benefits from the growing momentum across academia to adopt and evaluate a range of machine learning solutions relevant to the air quality domain. This is partly fuelled by the improvement in accessibility of the software and programming environments that underpin such tools, coupled with an increasing number of graduate learning opportunities that support uptake and deployment. It should also be noted that, through these partnerships, focus should also be given to the underlying infrastructure (data and software accessibility) which is not in an optimal state. Readiness for regulation and standards would be improved by ensuring resources are given to such matters.

It is also important to engage with the AI industry which is likely to act as a significant source of technical solutions at the environment-human health interface. For example, the recent announcement of a Digital Twinning of the Earth demonstrates the potential future service offering from technology providers; the development of strong science-technology partnerships key to avoid unintended consequences of change and retain public trust in proposed solutions. Defra could act as a gateway, or identified end-user, to relevant data on which such services are built. There are similar partnerships between, for example, Foundation Data Lab Europe, NVIDIA and the European Space Agency [https://nvidianews.nvidia.com/news/nvidia-announces-digital-twin-platform-for-scientific-computing]. Likewise, the partnership between NASA and Space-X demonstrates the recognition of joining forces where problem holders can benefit from interaction with organisations with higher financial agility and workforce skill sets.

The creation of an advisory board for cross-sector partnerships around adoption of AI could be an effective vehicle to maintain an appreciation of the breadth of activity. Membership could include representations from cross government departments, academia and industry with clear terms of reference around mapping successful adoption and management of the broad AI ecosystem.

3.3 Skills, Training and Sustainability:

Here AQEG consider how Defra and the DAs can enable their staff to develop the necessary skills to be aware of, use and understand AI driven technologies in an air quality context. The broader 'AI driven technologies' expression is used here to reflect the fact that much of the current developments in the air quality space are centred around the use of machine learning, whilst there is appropriate attention to the use of generative AI [e.g. ChatGPT] in the media. Generative AI, including tools such as ChatGPT, is very much at the cutting edge. There is value, however, in development of training programmes around the foundations of AI, from different types and applications of machine learning through to best practices around evaluating and interpreting models, meta-data standards and open science. In other words, the development of new tools for forecasting air quality and detecting the impact of interventions, for example, will remain as emerging research methodologies for some time as the community migrates these into operational models.

Improved training could include the following initiatives:

- Joint training programmes with HE institutions. There has been a significant growth in data science masters programmes across the UK; many of which involve public and private sector partners to define problem lead activity. Doctoral Training programmes may offer part time positions for existing employees to attain a PhD in data science. A number of international Universities provide free course content alongside recorded lectures.

- Tailored training options. There is a significant number of MOOC courses available now, with a low-cost entry point. Technology partners also offer very popular zero cost training portals where employees can build a personal portfolio of AI training. This includes, for example, the NVIDIA Deep Learning Institute (https://www.nvidia.com/en-us/training/). Microsoft have created the OpenDS4all (Open Data Science for all) programme which is built around a set of foundational concepts designed to upskill organisations around data science and cloud technologies.

Improved appreciation of potential uses could benefit from interaction with the broader AI community through joint workshops, literature reviews and targeted research programmes. Nurturing AI innovation within the department, and through its partnerships with external organisations, would support a longer-term goal of attracting and retaining staff with AI skill sets. As with fluid movements around regulation and standards, retaining staff in the public sector is a much bigger challenge that would benefit from a collective vision across public sector organisations and academia. Indeed, recent reports highlight the need for improving digital skills across the public sector. This runs in parallel with the 'digital divide' across the UK public, which of course has implications on the widespread adoption and trust of AI driven services at local to national government level. Forming partnerships between Defra and academia offers a potential route to bringing air quality science at the forefront of outreach with the public.

With all that in mind, the graduate workforce is likely to embrace data science as a core scientific tool in the future. This will inevitably reduce the burden on organisations to invest in targeted in-house training. However, the need to provide an innovative and nurturing 'AI aware' environment will remain, with proposed activities and initiatives given in this report designed to facilitate this.

4. Measuring Success

Quantifying the success of AI adoption relies on a clear strategy to first be defined. This would also depend on any governance structures and thus reporting in place (e.g. to an advisory board). Peer review literature on strategies for adoption and thus success of AI technologies appear in their infancy, largely as organisations continue to explore ways to improve overall performance through improved predictions and decision-making (Dirican, 2015). The latest McKinsey Global Survey on the current state of AI confirms the rapid growth of generative AI (gen AI) tools, less than one year after the widespread emergence of tools such as ChatGPT [McKinsey, 2023]. That same report notes that the most commonly reported uses of generative AI are in marketing and sales, product and service development whilst noting that knowledge-based industries (including education) could experience significant effects. Given the remit Defra has and the opportunities/barriers discussed in this report, there are several strategic Key Performance Questions (KPQs) that may be of use in the short (12 months) to medium term (2 years) to lead to improved readiness and adoption of AI. Potential strategic KPQs are suggested below, which could be partly addressed with the information provided in this report.

Strategic goal:	**Key Performance Questions:**	**Measures of success:**
Define AI strategy	What are the advantages and disadvantages around adoption of AI according to Defra remit? What is an appropriate governance structure to enable successful adoption and/or awareness of AI use in air quality mitigation?	Mapping of existing and emerging AI technologies to air quality management under Defra remit. Development of a strategy around a clear governance structure. This includes identifying roles, responsibilities, advisory board and reporting mechanisms.

Place Defra at the heart of shaping AI regulation in environmental policy.	Which areas of activity are open to potential challenges around unregulated use, closed source algorithms, and so on? What are appropriate mechanisms to encourage/force best practices of technologies provided through funding portfolios/contracts?	Documented requirements and commentary around regulation and standards for air quality management. Formation of cross government department partnerships. Horizon scanning activity around AI regulations Audit/cataloguing of data repositories and data streams on which AI driven air quality services/methods could be built. Partnerships with academia/industry on championing improved access to data, databases and workflows on evaluating new tools.
Develop and maintain training and development opportunities for staff.	What roles, responsibilities and skill sets would be needed in response to a clearly defined AI strategy? How are all staff made aware of AI as a general tool and evolving landscape of innovation? How are staff enabled to co-develop innovative solutions?	Mapping of existing and emerging machine learning techniques used in air quality science to training needs. For example: Deep learning, explainable AI, Forest based methods etc. Build a training portfolio for staff members. Research partnerships with academia and industry. Joint positions between Defra staff and academic institutions.
Develop and maintain an active role in the AI innovation landscape.	How to retain and nurture talent around the use and development of AI in the public sector? How to associate key developments and benefit	Create AI roles and responsibilities within the organisation Develop use cases for the AI landscape: e.g. set questions to be addressed and data to address that with

	associated with the use of AI in air quality with Defra?	Build new tools and techniques in-house; co-designed with academic/industry

5. Glossary

Artificial Intelligence (AI)	Artificial Intelligence (AI) as a goal of autonomous intelligence, which could include data driven systems or simply rule based systems.
Machine Learning (ML)	A subset of AI that concerns development of algorithms to extract meaning and build potential decisions around data.
Large Language Model (LLM)	A Deep Learning algorithm that can perform a variety of natural language processing (NLP) tasks. LLMs use transformer architectures and are trained using huge quantities of data (large).
ChatGPT	Chat Generative Pre-trained Transformer (ChatGPT). A class of Large Language Model (LLM) which is also a Foundational Model in that it can adapt to a wide range of requests.
Transformer	A Machine Learning architecture developed by Google.
BERT	Bidirectional Encoder Representations from Transformers (BERT). A Natural Language Processing (NLP) model.
Foundational Model	A large AI model trained to vast quantities of unlabelled data which is then able to be used for a wide range of tasks. ChatGPT is an example of a Foundational Model.
CoPilot	A Generative AI code generation facility. This enables end users to, for example, submit a natural language request and receive a sample of code based on that request.
Regression	An algorithm to understand the relationship between dependent and independent variables. Regression models are used to predict numerical values based on different data points (e.g. NO_2 as a function of time and weather).
Classification	An algorithm to accurately assign test data into specific categories, such as separating NO_2 from O_3.
Tree based ML method	An algorithm that works by partitioning the combinations of inputs into a number of

	smaller (non-overlapping) regions that connect with an output.
Deep Learning	Neural networks, modelled after the human brain, made of many layers of artificial neurons.
Explainable AI	A set of processes and methods that allows human users to comprehend and trust the results and output created by machine learning algorithms
Graphic Processing Units (GPUs)	Specialised hardware, most often connected to a motherboard via PCI that accelerates computer graphics and image processing.
Hyperparameters	Configuration variables of a machine learning model (e.g. number of deep neural layers) that scientists use to manage model training.
Convolutional Neural Networks (CNNs)	A machine learning model designed to process images; used for image and object detection.
Exploratory Data Analysis (EDA)	Used by data scientists to analyse and investigate data sets and summarise their main characteristics, often employing data visualisation methods.
Data Centric AI	The discipline of systematically engineering the data used to build an AI system.
Prompt Engineering	The process to guide generative artificial intelligence (generative AI) solutions to generate desired outputs.
Edge AI	Referred to as 'edge' because the AI computation is done near the user at the edge of a network, close to where data is collected. This includes AI computation on embedded devices within instruments, for example, rather than done via cloud computing.
IoT endpoints	Devices that are monitoring a service, process, or machine, by collecting data and sending it back to be analysed.
Data Driven Models	a process by which a algorithm learns from an input dataset and develops or improves a structure through the data for a particular purpose.
Unstructured Data	Information that isn't stored in a specific format. It can contain images, audio, or

	documents.
Unsupervised learning	Uses machine learning algorithms to analyse and cluster unlabelled datasets.
Supervised learning	Defined by its use of labelled datasets to train algorithms that to classify data or predict outcomes accurately.
Semi-supervised learning	A combination of supervised and unsupervised learning, often when there is no enough data required to build a robust supervised learning algorithm. An unsupervised approach first learns latent
Self-supervised learning	A training approach that enables algorithms to train and thus learn from unlabelled data
Transfer learning	The approach of taking a mature algorithm, often train on huge quantities of data, and modifying layers for use in domains where labelled data is difficult to accumulate.

6. References

6.1 Market and industry reviews

McKinskey. latest annual McKinsey Global Survey on the current state of AI (2023) https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year

Gartner. Gartner Hype Cycle™ for Artificial Intelligence (AI) (2023) https://www.gartner.com/en/articles/what-s-new-in-artificial-intelligence-from-the-2023-gartner-hype-cycle

6.2 Academic literature

Auld, G., Casovan, A, Clarke, A. & Faveri, B.: Governing AI through ethical standards: learning from the experiences of other private governance initiatives, Journal of European Public Policy, 29:11, 1822-1844, 2022. Doi: 10.1080/13501763.2022.2099449

Ball P.: Is AI leading to a reproducibility crisis in science? Nature, 624 (7990), 22-25, 2023. doi: 10.1038/d41586-023-03817-6.

Barker, M., Chue Hong, N.P., Katz, D.S. et al.: Introducing the FAIR Principles for research software. Scientific Data 9, 622, 2022. doi.org/10.1038/s41597-022-01710-x

Rishi, B., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M/S.: On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258, 2021.

O'Neil, C. and Schutt, R.: Doing Data Science, O'Reilly: A Very Short History of Data Science. ISBN: 978-1-449-35865-5. Forbes-Gil Press, 2013.

Devlin, J., Chang, M-W., Lee, K. and Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. 2018. arXiv:1810.04805v2

Hodorog, A., Petri, I. and Rezgui, Y.: Machine learning and Natural Language Processing of social media data for event detection in smart cities. Sustainable Cities and Society, 85, 104026, 2022.

Fan, X. and Phuong, T.: 2022 https://medium.com/sfu-cspmp/an-introduction-to-convolutional-neural-network-cnn-207cdb53db97

Juanals, B. and Minel, J-L.: An Instrumented Methodology to Analyze and Categorize Information Flows on Twitter Using NLP and Deep Learning: A Use Case on Air Quality. In Foundations of Intelligent Systems: 24th International Symposium, ISMIS 2018, Limassol, Cyprus, October 29–31, 2018, Proceedings 24, pp. 315-322. Springer International Publishing, 2018.

Karniadakis, G.E., Kevrekidis, I.G., Lu, L.: Physics-informed machine learning. *Nat Rev Phys,* 3, 422–440, 2021. doi.org/10.1038/s42254-021-00314-5

Kashinath, K., Mustafa, M., et al.: Physics-informed machine learning: case studies for weather and climate modelling, *Phil. Trans. R. Soc.* 379*,* 2194, 20200093, 2021. doi.org/10.1098/rsta.2020.0093

Khan, M.Y., Qayoom, A., Nizami, M.S., Siddiqui, M.S., Wasi, S. and Khaliq-ur-Rahman Raazi, S.M.: Automated Prediction of Good Dictionary EXamples (GDEX): A Comprehensive Experiment with Distant Supervision, Machine Learning, and Word Embedding-Based Deep Learning Techniques, Complexity, vol. 2021, Article ID 2553199, pp 18, 2021. doi.org/10.1155/2021/2553199

Nikula, R-P., Paavola, M., Ruusunen, M. and Keski-Rahkonen, J.: Towards online adaptation of digital twins. Open Engineering, 10, 776-783, 2020.

Ognjanovski, G.: 2019. https://towardsdatascience.com/everything-you-need-to-know-about-neural-networks-and-backpropagation-machine-learning-made-easy-e5285bc2be3a?gi=b2aaeb9a3fff

Geiss, A., Silva, S.J. and Hardin, J.C: Downscaling atmospheric chemistry simulations with physically consistent deep learning. Geoscientific Model Development 15, no. 17, 6677-6694, 2022.

Gilik, A., Ogrenci, A.S. and Ozmen, A.: Air quality prediction using CNN+ LSTM-based hybrid deep learning architecture. Environmental science and pollution research, 29, 11920-11938, 2022.

Grange, S.K. and Carslaw, D.C.: Using meteorological normalisation to detect interventions in air quality time series. Science of the Total Environment 653, 578-588, 2019.

Grant-Jacob, J.A. and Mills, B.: Deep learning in airborne particulate matter sensing: A review." Journal of Physics Communications 6, no. 12, 122001, 2022.

Lin, J-Y., Liu, H.T. and Zhang, J.: Recent advances in the application of machine learning methods to improve identification of the microplastics in environment. Chemosphere, 136092, 2022.

Reichstein, M., Camps-Valls, G., Stevens, B. *et al.*: Deep learning and process understanding for data-driven Earth system science. Nature, 566, 195–204, 2019. doi.org/10.1038/s41586-019-0912-1

Sauvageat, E., Zeder, Y., Auderset, K., Calpini, B., Clot, B., Crouzy, B., Konzelmann, T., Lieberherr, G., Tummon, F. and Vasilatou, K.: Real-time pollen monitoring using digital holography. Atmos. Meas. Tech., 13, 1539–1550, 2020. doi.org/10.5194/amt-13-1539-2020, 2020.

Song, C., Liu, B., Cheng, K., Cole, M.A., Dai, Q., Elliott, R.J.R., and Shi, Z: Attribution of Air Quality Benefits to Clean Winter Heating Policies in China: Combining Machine Learning with Causal Inference. Environmental Science & Technology, 57, 46, 17707–17717, 2023.

Spooner, F., Abrams, J.F., Morrissey, K., Shaddick, G., Batty, M., Milton, R., Dennett, A. et al.: A dynamic microsimulation model for epidemics. Social Science & Medicine 291, 114461, 2021.

Sulaimon, I.A., Alaka, H., Olu-Ajayi, R., Ahmad, M., Ajayi, S. and Hye, A.: Effect of traffic data set on various machine-learning algorithms when forecasting air quality, Journal of Engineering, Design and Technology, 2022. doi.org/10.1108/JEDT-10-2021-0554

Simonyan, K. and Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

Vongkusolkit. J. and Huang, Q: Situational awareness extraction: a comprehensive review of social media data classification during natural hazards. Annals of GIS, 27:1, 5-28, 2021. DOI: 10.1080/19475683.2020.1817146

Weerakkody, V., Irani, Z., Kapoor, K. *et al*. Open data and its usability: an empirical view from the Citizen's perspective. Inf. Syst. Front., 19, 285–300, 2017. https://doi.org/10.1007/s10796-016-9679-1

Wilkinson, M.D., Dumontier, M., Brand, I.J.S. and Aalbersberg, J.: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. 3 (1): 160018, 2016. doi:10.1038/SDATA.2016.18.

Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D. and Zhu, J.: Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges. In: Tang, J., Kan, MY., Zhao, D., Li, S., Zan, H. (eds) Natural Language Processing and Chinese Computing. NLPCC 2019. Lecture Notes in Computer Science(), vol 11839. Springer, 2021. doi.org/10.1007/978-3-030-32236-6_51

6.3 UK and International organisation AI papers/relevant commentaries:

AI Council Roadmap, UK, 2021: https://assets.publishing.service.gov.uk/media/5ff3bc6e8fa8f53b76ccee23/AI_Council_AI_Roadmap.pdf

Ada Lovelace Institute, UK, 'Regulate to innovate: 'A route to regulation that reflects the ambition of the UK AI Strategy'. 2021. https://www.adalovelaceinstitute.org/report/regulate-innovate/#:~:text=A%20report%20which%20sets%20out,beneficial%20for%20people%20and%20society.

FDA, 2023a: Using Artificial Intelligence and Machine Learning in the Development of Drug and Biological Products - Discussion Paper and Request for Feedback.

https://www.fda.gov/media/167973/download?attachment

FDA, 2023b: Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

INSPIRE, Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE): https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32007L0002

UNESCO, Recommendation on the Ethics of Artificial Intelligence (2022): https://unesdoc.unesco.org/ark:/48223/pf0000381137

Gov-uk, Data ethics and AI guidance landscape. A *page that collates the various existing ethical principles for data and AI, developed by government and public sector sector bodies. It intends to provide clarity and guidance for public servants working with data and/or AI.* (2020), https://www.gov.uk/guidance/data-ethics-and-ai-guidance-landscape

Leslie, D.: Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. The Alan Turing Institute. 2019. https://doi.org/10.5281/zenodo.3240529

Office for National Statistics (ONS).: Exploring the UK's digital divide, 2019. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/articles/exploringtheuksdigitaldivide/2019-03-04

6.4 Open source code repositories and tools of interest:

SATRE Specification Repository: Technical Document for the Standard Architecture for Trusted Research Environments Project https://github.com/sa-tre/satre-specification