The Alan Turing Institute

# Tackling climate change with data science and AI

Successes and recommendations from The Alan Turing Institute's AI for science and government programme

The Alan Turing Institute

**About The Alan Turing Institute**

The Alan Turing Institute is the national institute for data science and artificial intelligence (AI). Established in 2015, we are named in honour of Alan Turing, whose pioneering work in theoretical and applied mathematics, engineering and computing laid the foundations for the modern-day fields of data science and AI. Headquartered at the British Library in London, we partner with organisations across government, industry, academia and the third sector to undertake world-class research that benefits society.

**About the ASG programme**

Established in 2018, the AI for science and government (ASG) programme is a five-year, £38.8 million research programme at The Alan Turing Institute that aims to deploy data science and AI to address significant societal challenges. Funded through UK Research and Innovation's (UKRI's) Strategic Priorities Fund, and delivered in partnership with the Engineering and Physical Sciences Research Council (EPSRC), the programme brings together researchers from diverse disciplines to tackle problems including climate change, health emergencies, economic instability, and online harms.

# Contents


| | |
|---|---|
| **Executive summary** | **3** |
| **1. Introduction** | **4** |
| **2. Using data science and AI to tackle climate change** | **6** |
| 2.1 Monitoring the environment | 7 |
| 2.2 Forecasting environmental change | 9 |
| 2.3 Simulating the human cost of climate change | 11 |
| 2.4 Adapting to climate change | 13 |
| **3. Recommendations** | **16** |
| **4. Conclusions** | **21** |
| **5. Acknowledgements** | **22** |



**Authors**

The AI for science and government programme's 'Environment and sustainability' working group (Alden Conner, Scott Hosking, James Lloyd, Achintya Rao, Gavin Shaddick, Malvika Sharan)



DOI: 10.5281/zenodo.7712969

# Executive summary

In this white paper, we share how The Alan Turing Institute's AI for science and government (ASG) programme has been using collaborative and multidisciplinary data science and AI to help tackle climate change.

We describe four key research areas in which data science and AI can address the climate crisis:

1. Monitoring the environment.
2. Forecasting environmental change.
3. Simulating the human cost of climate change.
4. Adapting to climate change.

Within each of these areas, we spotlight an ASG project that has demonstrated particular success, encompassing software for automatic plankton classification, a framework for forecasting Arctic sea ice change, a model for simulating the health impacts of extreme heat, and a digital twin that is being used to optimise the world's first underground farm.

Based on our experiences with these projects, and within the broader ASG programme, we provide four recommendations for how researchers and funders can better use data science and AI to tackle climate change:

1. Apply cutting-edge data science and AI to environmental decision-making.
2. Foster a community of AI specialists, environmental researchers and stakeholders.
3. Build robust digital pipelines.
4. Develop digital twins to support decarbonisation.

The climate emergency needs urgent solutions. This paper demonstrates how data science and AI can play a key role by advancing our scientific understanding of this planet-threatening problem and providing new pathways to mitigation and adaptation.

> **"The future health, wealth and safety of our species is intimately linked to the fate of our planet."**

# 1. Introduction

Climate change is arguably the biggest threat facing humanity. From rising sea levels to changing weather patterns, the effects are global and catastrophic. And the evidence is unequivocal: human activity is the main cause of rising temperatures. We are emitting unprecedented amounts of greenhouse gases that are disrupting our planet.

We need urgent solutions to climate change from science, engineering and politics. The Alan Turing Institute is the national institute for data science and artificial intelligence (AI), and these two disciplines have a critical role to play in understanding, mitigating and adapting to the climate crisis. Data science techniques can help us to make sense of the vast amounts of information being collected about our planet, while AI tools and methods can provide decision-making capabilities, enabling researchers to more accurately and efficiently study climate change, predict its impacts and develop solutions.

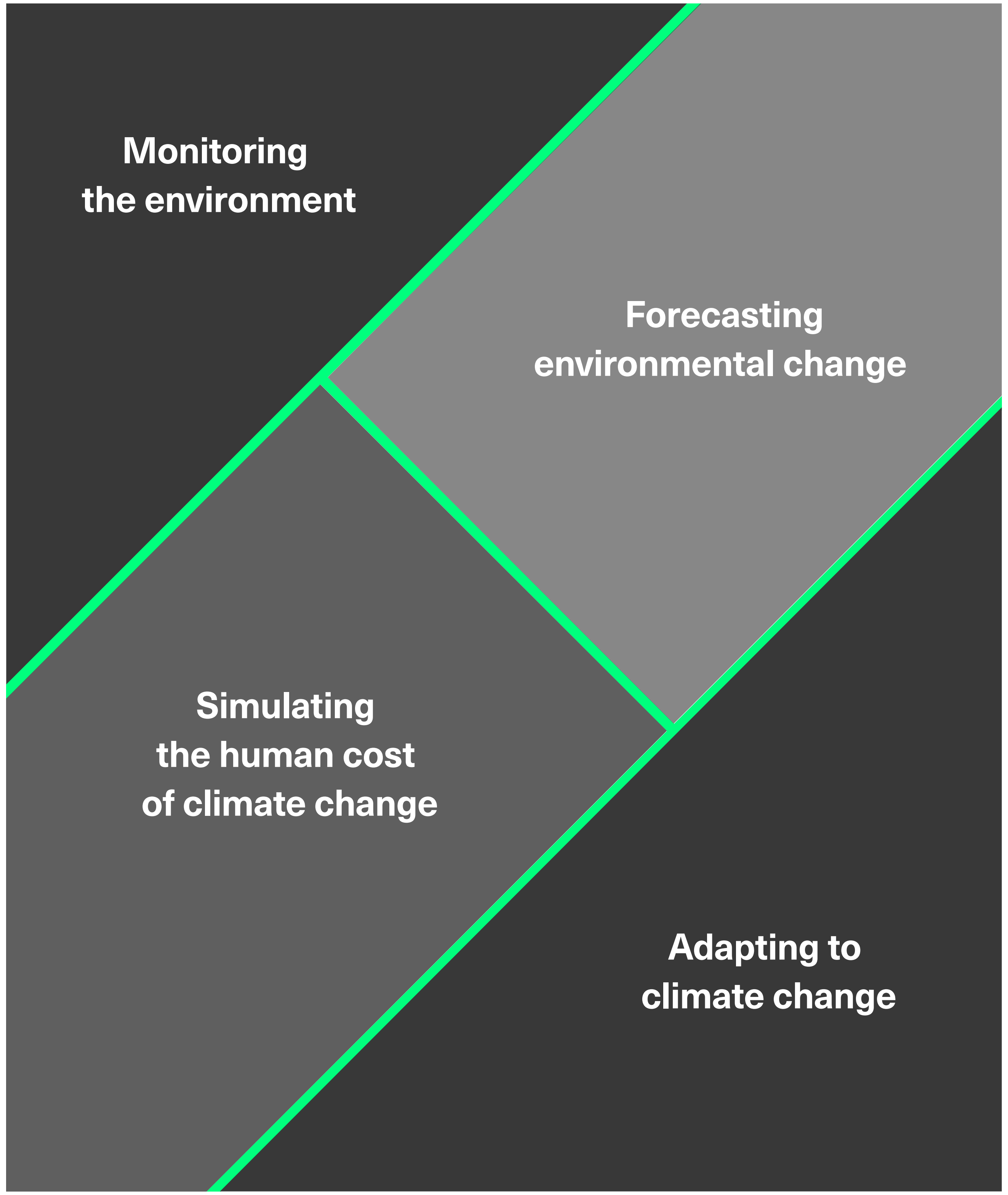
We believe that a truly interdisciplinary approach is key to addressing the climate problem. By bringing together methods and expertise from fields including environmental science, computer vision and predictive modelling, researchers will be able to develop adaptable, reproducible tools and digital frameworks that integrate data from disparate sources, providing novel insights that can inform policy-making.

Since 2019, the Turing's AI for science and government (ASG) programme has been convening researchers from diverse disciplines to apply data science and AI to climate change and other environmental issues. This led to the launch of the programme's dedicated 'Environment and sustainability' (E&S) theme in 2021.

In this white paper, aimed at all researchers and funders who are exploring (or would like to explore) the intersections between data science, AI and environmental science, we present an overview (Section 2) of the four main climate change-related research areas that the E&S theme has been working across: (1) monitoring the environment; (2) forecasting environmental change; (3) simulating the human cost of climate change; and (4) adapting to climate change. For each of these research areas, we spotlight an E&S project that has demonstrated a particularly successful application of data science and AI.

Based on the learnings from our work, we also provide four recommendations for how researchers and funders can better harness the potential of data science and AI to tackle climate change (Section 3). These recommendations, which cover both operational and technical aspects of the research landscape, provide a roadmap to furthering climate science in the short- and long-term.

Data science and AI represent two of our most powerful assets in the fight against climate change. As computers and algorithms become faster and more efficient, the potential applications of these fields will continue to grow. Now is the time to reimagine the way we conduct climate science research and address the crisis head-on.




# 2. Using data science and AI to tackle climate change

Barely a day goes by without a reminder of the devastating impacts of climate change. Last year was one of record-breaking extreme weather events, from heatwaves in Europe to drought in the Horn of Africa and mass flooding in Pakistan. It was also the UK's warmest year on record, with a new record-high temperature of 40.3°C recorded on 19 July.

While no single weather event can be directly attributed to climate change, there is clear evidence that the changing intensity and frequency of extreme events cannot be explained by natural variability alone. Unless we find ways to limit global warming, these impacts are only going to worsen, leading to considerable suffering.

In this section, we outline four research areas in which data science and AI can play a key role in the scientific response to climate change, spotlighting successes from the ASG programme. First, in order to better understand climate change and its impacts, we need to **monitor the environment** (Section 2.1) using cutting-edge techniques that allow collection and analysis of high-quality data in real time across our planet's ecosystems. Second, we need to accurately **forecast environmental change** (Section 2.2), taking advantage of the predictive power of AI techniques such as deep learning.

Third, we need to use AI-based modelling to **simulate the human cost of climate change** (Section 2.3), such as the impacts on health, conflict, migration and economies. And finally, we need to develop tools that allow us to **adapt to climate change** (Section 2.4), such as AI-driven farming methods that can help to create more resilient food systems.

## 2.1 Monitoring the environment

In research stations across the world, environmental data is streaming in in real time from instruments including satellites, buoys and ground-based sensors. Combining different measurement techniques allows researchers to form a clear picture of complex environmental systems such as oceans, forests and volcanoes, and the advent of AI-based monitoring and analysis has furthered our ability to characterise these systems across multiple variables and scales. Importantly, AI monitoring technologies can also help to translate data into timely responses, improving our ability to protect ecosystems and communities vulnerable to natural hazards.

One crucial, AI-powered tool in environmental monitoring is computer vision, which automates the detection, classification, tracking and measurement of objects or regions in images. By processing images through pattern recognition, computer vision can efficiently analyse large datasets that would traditionally require a massive human effort to label all images manually. This technique can help to accelerate the pace of discovery in environmental science, empowering researchers to gain a deeper understanding of the ecosystems that require protection and the climate patterns that threaten them.


**Case study: Cutting-edge software for surveying ocean health**

Microscopic sea creatures called zooplankton play a critical role in the marine ecosystem and the global carbon cycle. Plankton populations are a sensitive indicator of ecosystem change, and the relative abundance of different plankton species can be measured as a proxy for ocean health. However, counting and classifying plankton manually is labour-intensive.

The research vessel (RV) *Cefas Endeavour*, a ship owned and operated by the Centre for Environment, Fisheries and Aquaculture Science (Cefas), houses a high-speed imaging instrument that Cefas uses to take pictures of seawater containing plankton. At first, Cefas scientists were unable to keep up with the number of images the instrument was generating: manual analysis required firstly removing the approximately 80% of images showing particles of marine detritus rather than plankton and then classifying the remaining images into two categories of plankton. Recognising that this was a task that could benefit from AI but not having the in-house expertise to develop the computer vision software themselves, they took the problem to an ASG-funded Data Study Group (DSG) – an intensive 'hackathon' that brings together data scientists from different disciplines to tackle a real-world problem.

During two weeks in December 2021, the group developed an automated plankton classification algorithm. The current version counts and classifies 50 organisms per second, with accuracy rates of over 90%. Within six months, Cefas deployed the algorithm for use on the *Endeavour* and it is now being integrated with software that will allow remote access to data in real time, to help scientists guide and adapt sampling as it happens. Whilst the existing system is designed for plankton, the open-source code developed via the DSG could be adapted for use with other marine life, ushering in a new era of real-time ocean monitoring.

The speed with which the results were delivered was a product of the Turing's collaborative DSG format, which takes advantage of multidisciplinary expertise to brainstorm and innovate at a rapid pace. The project also benefitted from a Turing-developed computer vision toolkit called Scivision (see also page 18), which allowed DSG participants to quickly import data without the need for extensive data preparation, and to visualise and compare different model outputs. Feedback from the participants will contribute to future improvements in the Scivision platform.

## 2.2 Forecasting environmental change

For decades, researchers have developed physics-based models of the ocean and atmosphere to simulate environmental processes – work that has been instrumental in advancing our understanding of weather patterns and climate change. But even with the latest physics-based models, fed by large datasets and running on advanced computers, it can still be difficult to forecast extreme events or climate change impacts in remote environments (often in less developed countries) due to the sheer complexity of the processes involved and/or a lack of observational data. With AI-based models, however, we have the potential to rapidly advance forecasting techniques for the most difficult-to-predict phenomena. Deep learning, for example, is already being explored as a way to improve rainfall prediction and short-term weather forecasting.

One particular region that could benefit from AI-based modelling is the Arctic, which is warming four times faster than the rest of the world, resulting in a devastating decline in summer sea ice coverage. This puts the future of local ecosystems and indigenous communities in doubt, and has knock-on effects for the entire climate system. Combining the latest AI-powered predictive modelling techniques with detailed, long-timespan remote sensing data has the potential to significantly improve our ability to forecast conditions in the Arctic and other vulnerable regions.

**Case study: Faster, more accurate sea ice forecasting**

The future state of Arctic sea ice will have crucial implications at both a local and global level. But forecasting sea ice typically relies on physics-based models, which consume vast quantities of supercomputing power through lengthy run times, and rely on researchers developing complex, inexact, deterministic mathematical equations for the chaotic and small-scale processes that play out. AI promises faster, more accurate tools for predicting sea ice conditions.

Developed through a collaboration between Turing and British Antarctic Survey (BAS) researchers, IceNet is the first AI tool that can forecast the seasonal change in concentration of Arctic sea ice more accurately and efficiently than traditional physics-based approaches. Using deep learning models called convolutional neural networks (CNNs), trained on satellite observations and climate simulations, IceNet forecasts sea ice conditions for the next six months. Tested on unseen data after training, the tool is 92–97% accurate (dependent on the season). It outperforms state-of-the-art, physics-based forecasting systems in terms not just of accuracy but also speed, with researchers showing that it runs 2,000 times faster on a laptop than the physics-based models do on a supercomputer.

IceNet showcases the huge potential of AI for sea ice forecasting in the medium range (for the season ahead): a key timescale for decision-making. Improved sea ice predictions could underpin new early-warning systems designed to protect wildlife and coastal communities from dangerous sea ice conditions. BAS is already engaging with local governments and the World Wide Fund for Nature (WWF) to explore how IceNet can be applied in these contexts across the Arctic.

Meanwhile, the tool is being integrated with a virtual representation ('digital twin')

of the Royal Research Ship (RRS) *Sir David Attenborough*, operated by BAS, to help plot fuel-efficient routes through sea ice.

This Turing-BAS partnership exemplifies the kind of open, interdisciplinary collaboration that is needed to help us better understand and predict the consequences of climate change in one of the world's most fragile ecosystems. The goal for the next generation of IceNet is to go beyond monthly averaged Arctic forecasts to enable daily predictions of sea ice concentration in both the Arctic and Antarctic.

## 2.3 Simulating the human cost of climate change

Climate change will impact humans in myriad ways, and is likely to disproportionately affect groups that are already vulnerable – exacerbating existing inequalities. A 2020 World Bank paper estimated that the climate crisis will push up to 132 million people into extreme poverty by 2030. There will be increased gender and economic inequality, and we are likely to see increased conflicts caused by pressure on food systems, competition over resources such as water and energy, and migration driven by extreme weather events.

The climate crisis will also profoundly affect human health, and poorer countries that are more exposed to climate risks will also bear the brunt of these impacts. The effects on health can be both direct (such as through exposure to unseasonably high temperatures or extreme weather events) and indirect (such as changes in climate-sensitive diseases such as malaria or cholera, or the physical and mental health consequences of climate-related poverty or conflict).

A promising computational modelling technique for understanding the impacts of climate change is microsimulation, which allows the effects of changes in the environment to be assessed at the individual, rather than aggregate, level. Microsimulation acknowledges differences within populations and provides a powerful tool for assessing how policy interventions affect individuals.

To provide privacy-preserving datasets for microsimulations that are both granular and large-scale, researchers can combine detailed individual data, such as that collected in a survey, with more complete population-level statistics, such as census data, to create a new 'synthetic population'. This enriched and anonymised representation of the actual population can then be combined with microsimulation to study the effects of a range of climate-related variables and allow the quantification of potential adaptation measures, such as those developed to address the increased intensity and frequency of heatwaves that are expected with the changing climate.

**Case study: Modelling the health impacts of extreme heat**

Climate change means that heatwaves, such as those that hit Europe in 2022, are becoming more common and more intense. This has knock-on effects for human health: the UK's Department of Health and Social Care cautions that severe heat can be fatal for vulnerable people.

Existing approaches to estimating the health impacts of environmental hazards such as high temperatures typically involve working at the aggregate level, i.e. looking at relationships between average levels of environmental exposures (in this case heat) and total counts of adverse health effects over a specified region. This results in an 'average' effect over the entire population and doesn't highlight potential differences in risk across different groups due to variation in people's exposure to the environmental hazard (e.g. how much time they spend outdoors, whether they have access to cool buildings) and individual factors that determine vulnerability to a given exposure (e.g. age, pre-existing medical conditions). This means that the effects of policy interventions designed to protect communities from extreme heat may differ across socio-economic groups and, in some cases, potentially increase inequalities. To develop more targeted interventions for the most vulnerable groups, we need modelling tools that can estimate the personal impacts of heat.

The Dynamic Microsimulations for Environments – Climate, Heat and Health (DyME-CHH) project aims to address this problem through the development of a microsimulation model, which allows simulation of the effects of heat at an individual scale. The model uses synthetic populations, provided by the Turing's Synthetic Population Catalyst project and including risk factors such as those listed above, to provide a detailed and representative dataset that contains no confidential information. The model incorporates data at multiple scales, allowing users to explore different scenarios at an individual, community, regional and national level.

The DyME-CHH team now plans to use its model to compare the effects of different policy interventions on climate-related health. By identifying the groups most at risk from heat exposure, the model will support local councils in developing effective ways to adapt to extreme heat. Data from the model will also feed into the University of Exeter's Local Climate Adaptation Tool (LCAT), which has been co-designed with local government authorities including Cornwall Council and Manchester City Council. LCAT provides decision makers with a user-friendly interface for exploring the predicted health effects of climate change, and suggests evidence-based recommendations for adaptation.

## 2.4 Adapting to climate change

While it is essential that we do all we can to limit greenhouse gas emissions, we will also need to find ways to adapt to both current and future climate change impacts. There are many ways in which AI could help with this, such as by creating predictive, interactive wildfire maps to help firefighters manage resources, or allowing businesses to predict and prepare for climate change-related supply chain or production issues.

Another area of promise for AI is in helping our food systems adapt to climate change. Increasing temperatures, shifting rainfall patterns and extreme weather events are already causing food insecurity, and the problem is only likely to worsen, with around 70% more food production required by 2050 to feed our planet's growing population. AI has potential applications across the agricultural sector, such as simulations for predicting crop yields, robotic machinery for carrying out large-scale tasks, and computer vision tools for automatically monitoring crop and soil health. According to one estimate, the market for AI in agriculture will be worth US$4.7 billion by 2028, up from US$1.7 billion in 2023.

A tool that could prove especially useful in this space is digital twins – computer models that dynamically simulate a physical object or system by incorporating real-world sensor data. By analysing a digital twin of, say, a farm, a group of farms, or even a region's entire food system, fed with data streams from the real-world counterpart, researchers could better understand the physical version and make informed decisions to improve its functioning and produce food in a more sustainable, efficient way. As the costs of computing, live data sensing and wireless communication continue to drop, digital twins will become an ever more important tool in the scientific response to climate change.

## Case study: Optimising the world's first underground farm

In a disused World War II air raid shelter below London's Clapham High Street, Zero Carbon Farms offers a vision of how alternative farming methods could help address some of the food security impacts of climate change. This underground, hydroponic farm produces carbon-neutral salad greens for local stores and restaurants, and it benefits from a digital twin developed at the Turing called the Crop Research Observation Platform (CROP), which helps growers to optimise farm conditions so that crops are grown in the most efficient and environmentally friendly way.

CROP is the product of a collaboration between Zero Carbon Farms, researchers from the University of Cambridge and the Turing's data-centric engineering programme, and software engineers from the Turing's Research Engineering Group. The latest version of CROP takes data from sensors positioned around the farm to provide growers with remote access to current and historical information on variables including temperature, humidity, water quality and $CO_2$ levels. It also incorporates data on energy consumption and crop yields, allowing growers to pinpoint the optimal conditions for efficient crop growth.

As well as sensor data, the digital twin also integrates a physics-based model of the farm's environment that can forecast future conditions three days in advance, helping to guide proactive management of ventilation, lighting and other conditions to maximise crop yields and reduce losses – for example, in the case of a predicted heatwave. The physics-based model is calibrated by the incoming sensor data using a novel methodology developed by Turing researchers, ensuring it matches real-world conditions as closely as possible.

Turing researchers and software engineers continue to collaborate closely with the Zero Carbon Farms team, developing and tailoring new features to suit its needs. The extensive work carried out in improving the CROP platform stands to have benefits beyond just this farm: the project's open-source code can be used to build digital twins for other non-traditional farms seeking to maximise their yields and energy efficiency.

# 3. Recommendations

The future health, wealth and safety of our species is intimately linked to the fate of our planet – something that is encapsulated within the UN's Sustainable Development Goals (SDGs).

Data science and AI have huge potential for helping us to tackle the climate crisis and other challenges associated with environmental change, but their role is under-recognised and under-explored.

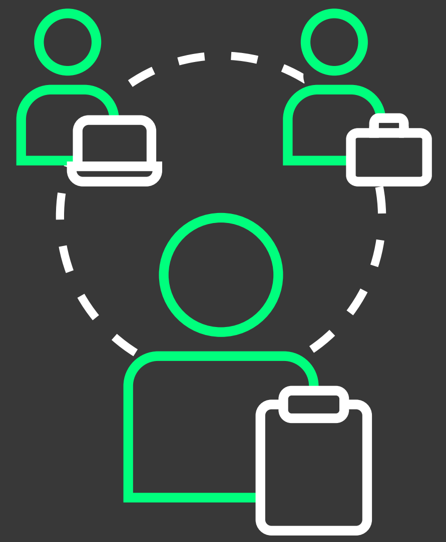
Based on the findings and learnings from the four projects detailed in Section 2, and the ASG programme as a whole, we present in this section four recommendations for how researchers and funders can better harness the potential of data science and AI across the environmental sector.

## 1

**Apply cutting-edge data science and AI to environmental decision-making**

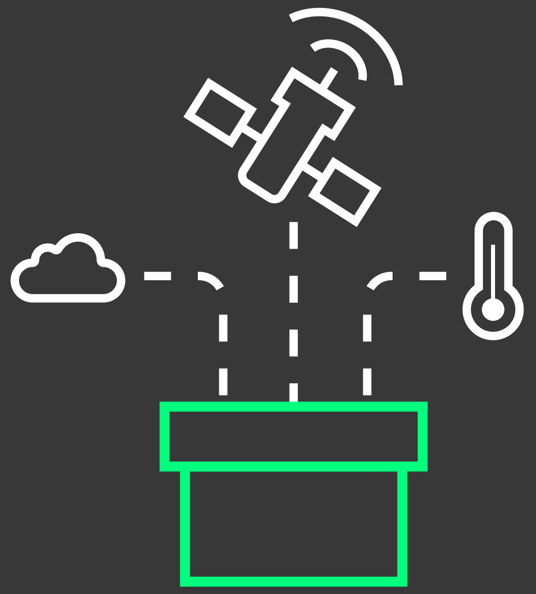
## 2

**Foster a community of AI specialists, environmental researchers and stakeholders**



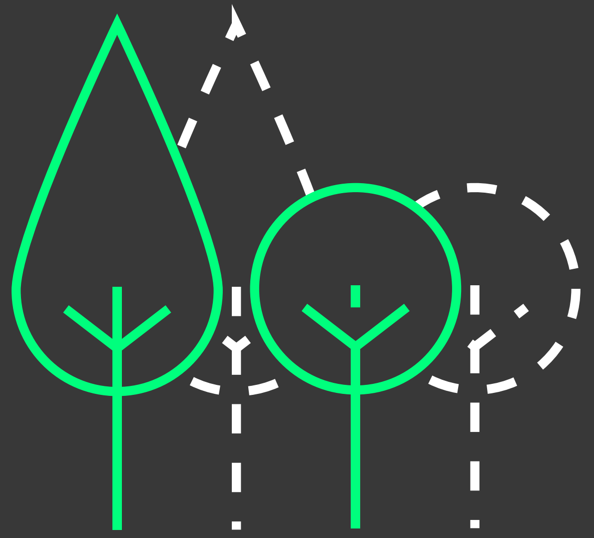
## 3

**Build robust digital pipelines**



## 4

**Develop digital twins to support decarbonisation**



## 1 Apply cutting-edge data science and AI to environmental decision-making

The sheer volume and complexity of data that needs to be analysed in environmental decision-making and in computing metrics (such as those for the SDGs) presents a golden opportunity for data science and AI. We need to capitalise on the combination of cutting-edge algorithms, machine learning, data-driven technologies and high-performance computing to deliver the information necessary for evidence-based decision-making.

The UN's 2030 Agenda for Sustainable Development – a plan of action for people, planet and prosperity – states that "quality, accessible, timely and reliable disaggregated data will be needed to help with the measurement of progress [towards the SDGs] and to ensure that no one is left behind". However, the availability of such information is currently a major issue, with the UN Environment Programme estimating that of the 93 environment-related SDG indicators, there is insufficient data to assess progress for 68% of them.

We believe that data science and AI hold great promise for filling information gaps such as these. In particular, these technologies can help to: monitor the resilience of environmental and social-ecological systems; provide early warning for environmental risks; inform the development of a new integrated assessment of sustainable pathways; and put people at the heart of sustainability science by leveraging the huge recent increase in social data.

We recommend that data science and AI methods currently used in other sectors be adapted for use in addressing environmental challenges, acknowledging the characteristics of multifaceted environmental data and incorporating the extensive existing knowledge of environmental systems. We also recommend that new methodologies are developed to directly answer environmental challenges and provide the meaningful insights that are required to drive intervention strategies. When implementing these recommendations, it must be recognised that data science and AI technologies can be energy- and resource-intensive, and so we recommend that all development takes place within the context of long-term environmental sustainability.

Ultimately, unlocking the full possibilities of data will have far-reaching impact within decision-making processes across a wide range of environmental areas, but will be especially crucial in ensuring that SDG targets and indicators are true representations of the world that the SDGs aim to protect.

## 2 Foster a community of AI specialists, environmental researchers and stakeholders

To create data science and AI tools that have real-world impact across multiple areas, we recommend a co-developmental approach that draws on expertise from diverse sectors and disciplines. ASG projects, for example, have been driven by collaboration across scientific fields and between partners from academia, industry and government agencies across the UK and beyond. Teams that collaborate in this way bring a range of perspectives and questions to a project, helping to ensure that any outputs are adaptable. Co-development also creates a solid foundation for open and reproducible work (a core theme of the ASG-funded online handbook *The Turing Way*), limiting duplication of effort and ensuring that others can build on the work.

To fully realise the potential of data science and AI in this area, environmental researchers, policy makers and other stakeholders need not only access to the latest tools, but also the knowledge and opportunity to be able to use them effectively. We recommend that: environmental researchers are upskilled to become more familiar with what is possible with data science and AI methods; those who work across research domains are championed; the lengthy process of developing new collaborative networks (such as the Turing's 500-strong interest group in environment and sustainability) is supported; truly interdisciplinary research is funded; and a pool of research software engineers (RSEs) is developed to ensure that digital technologies are developed sustainably to underpin both scientific discovery and policy decisions.

A strong example of co-development from the ASG programme is Scivision: an open-source toolkit that curates pre-trained computer vision models and scientific datasets, enabling researchers without computer vision expertise to find and use specialist software for image classification and feature identification. The Scivision development team comes from a range of scientific backgrounds, so the tool was built with adaptability in mind. As well as being used in the Cefas plankton classification project (Section 2.1), Scivision is already being applied in plant phenotyping, satellite-based landscape monitoring, and object detection in electron microscopy images.

Central to the success of the four ASG projects profiled in Section 2 are the Turing's Research Application Managers (RAMs) – a role created within the ASG programme to drive development of tools aimed at solving real-world problems. RAMs work directly with researchers to consider the long-term goals of a project; identify potential external stakeholders; encourage the creation of accessible, generalisable and reproducible tools; and – ultimately – maximise a project's impact and sustainability. We believe that environmental research can benefit greatly from these types of roles.

## 3 Build robust digital pipelines

It is clear that a wide variety of areas related to environment and sustainability can benefit from the increasing availability of large and complex datasets from diverse sources including environmental monitoring, remote sensing, bio-logging, climate modelling, social media and contributions from citizen science. This richness of data provides us with an exceptional opportunity to transform our knowledge of the effects of environmental change and to develop solutions to environmental challenges.

However, for data science and AI to effectively use this data, we must design, develop and deploy robust digital infrastructure to not only handle the immense volume of data being collected, but also to efficiently connect those data pipelines to powerful computing resources that allow us to harness machine learning, statistical analysis, numerical models, and other data science and AI methods.

This will require the development of standards for metadata and data architecture to ensure that datasets are accessible and reusable; the automation of methods for ingesting large datasets of different types; and the creation of frameworks in which data can pass freely between different computational systems. There is also a need for usable software libraries, and reproducible analyses and workflows.

The Turing-supported Pangeo community platform is a good example of a resource developed to address the challenges of handling big data in geosciences research, providing the types of software and infrastructure needed to enable environmental researchers to glean insights from large and complex datasets.

To help build the resources required for more robust digital pipelines and reproducible research, we recommend that RSEs are fully integrated within research teams and that software-based outputs are developed using best principles, encouraging reuse and applicability of software beyond the original project.

This connection between research and applications is key to the delivery of impactful research. An example of this is the pioneering Research Engineering Group at the Turing – a team of RSEs who work directly with scientists throughout the lifetime of a project, and often beyond. This way of working ensures that tools that are developed within individual research projects are useful and applicable to a wider range of areas.

IceNet (Section 2.2) is one project that has benefitted from close involvement from RSEs at the Turing and BAS, who ensured that the sea ice forecasting framework was robust and extendable to other applications. IceNet's data pipeline is now being used to inform the information management framework for environmental digital twins proposed by the Natural Environment Research Council (NERC).

## 4 Develop digital twins to support decarbonisation

The UK government's Net Zero Strategy details how greenhouse gas emissions will be reduced across all sectors of the economy to comply with the UK's commitment to meeting its net zero target by 2050. Data science and AI have a crucial role to play in providing information and tools that will enable the successful development of policies and interventions for decarbonisation, helping to facilitate change across energy, transport, agricultural and other environment-related systems.

Digital twins represent one of the most powerful tools for helping us to achieve this. We envision a wide range of digital twins of both natural and built environments, allowing researchers to optimise systems and policies with the aim of minimising greenhouse gas emissions. For example, a digital twin of an energy system could allow decision makers to explore different policies around renewable energy generation or support schemes for home energy upgrades. The Turing has already been involved in a demonstrator project in this area with Energy Systems Catapult and the government's former Department for Business, Energy and Industrial Strategy (BEIS).

Most existing digital twins (such as CROP, Section 2.4) represent relatively simple, closed systems, but future digital twins will represent vast, interconnected systems using multimodal data from a broad range of sources. These technically complex simulations will require advances in data processing and AI methods in order to integrate and analyse highly diverse data, and we must consider how current research outputs will work as components of future digital twins. These themes are explored further in the ASG programme's companion white paper on 'ecosystems of digital twins'.

To realise the full potential for digital twins in the environmental sector, we recommend that experts in data science, AI, software engineering, computer science and environmental research are given the time, space and funding to work together to co-design and co-develop methods that transcend the traditional and often siloed domain areas. Digital twins must also be designed with interoperability in mind, i.e. the ability to couple environmental, social and economic systems, incorporating a wide variety of data sources. To enable all of this, we recommend that: digital twins are developed in the open using best practice; guidance and demonstrators are provided to democratise the tools; and researchers and organisations communicate effectively how digital twins can benefit environmental science, policy, society and the natural world. These recommendations will help to inform the new Turing Research and Innovation Cluster in Digital Twins, which builds on the wider ASG portfolio of digital twin research and innovation.

# 4. Conclusions

The climate emergency is the defining crisis of our time. Governments around the world must do all they can to limit greenhouse gas emissions and secure a safe and hospitable planet for future generations. Meanwhile, science and technology also have a vital role to play in understanding, mitigating and adapting to the problem.

In this white paper, we presented a whistlestop overview of some of the ways in which data science and AI can help to tackle climate change, drawing on successful projects from the 'Environment and sustainability' theme within the Turing's ASG programme. From classifying plankton and forecasting Arctic sea ice to modelling the health impacts of extreme heat and monitoring an underground farm, underpinning all of these ambitious projects is a drive to use reproducible, ethical, multidisciplinary and collaborative research to benefit our planet and the people that call it home. We believe that the community building and convening ability of the ASG programme have helped to break boundaries between environmental research and data science and AI.

Our recommendations for researchers and funders are four-fold. We need to use the latest data science and AI techniques to support decision-making across the environmental sector. We need to encourage collaboration across domains and sectors to ensure that outputs from environmental research are adaptable, reproducible and ready to use in real-world situations. We need to build robust digital infrastructure that allows researchers to share, store and analyse the vast amounts of data being collected about the environment. And we need to develop digital twins to model and understand the systems that will help us to decarbonise our society.

Used in the right way, data science and AI are incredibly powerful techniques that can offer huge societal benefit. Let's continue to harness their potential and develop the solutions that the future of our planet depends on.

# 5. Acknowledgements

The authors would like to thank the following people for their input during the development of this paper and for their work on the profiled ASG projects:

Sebastian Ahnert, Tom Andersson, Ceren Barlas, Nick Barlow, Arielle Bennett, Hayley Bennett, Mark Birkin, Robert Blackwell, Emma Bland, Ruth Bowyer, Alice Budden, James Byrne, Ed Chalstrey, Ryan Chan, Dave Chapman, Ruchi Choudhary, Alejandro Coca-Castro, Zack Xuereb Conti, Evangeline Corcoran, Jess Dicken, Jennifer Ding, Zeynep Engin, Marjan Famili, Isabel Fenton, Beatriz Costa Gomes, Markus Hauru, Katy Henderson, Kasra Hosseini, Aoife Hughes, Zaynab Ismail, Mariya Iqbal, David Salvador Jasin, Alan Lowe, Ben MacArthur, Harriet McCann, Aida Mehonic, David Sarmiento Pérez, Sophie Pitois, James Robinson, Jon Rowe, James Salter, Camila Rangel Smith, Oliver Strickson, Hannah Sweeney, Louisa van Zeeland, Rebecca Ward, Peter Yatsyshin.

We also thank the creative communications company Soapbox for designing the paper and the recommendations icons.

This work was supported by  Wave 1 of The UKRI Strategic Priorities Fund under the EPSRC Grant EP/T001569/1 and EPSRC Grant EP/W006022/1, particularly the ‘AI for science’ and ‘Environment and sustainability’ themes within those grants and The Alan Turing Institute.

## Image credits

Front and back cover: Marcel Poncu
Page 5: Everett Collection
Page 8: Crown copyright
Page 10: Kaspars Grinvalds
Page 12: Nelson Antoine
Page 14: Paul Marc Mitchell

UK Research and Innovation

The Alan Turing Institute

turing.ac.uk
@turinginst