National
Engineering
Policy Centre

February 2025

Engineering Responsible AI: foundations for environmentally sustainable AI

Royal Academy of Engineering

bcs The Chartered Institute for IT

IET The Institution of Engineering and Technology

Executive summary

Artificial intelligence (AI) can be leveraged to accelerate progress towards net zero carbon emissions and improvements in environmental sustainability. AI can be used to help optimise energy consumption, manage grid demand, reduce waste, monitor ecosystem health, and quantify the impact of climate change and adaptation strategies.[1] To maximise the environmental and societal benefits that AI can offer, however, **AI systems and services must be environmentally sustainable on a lifecycle basis**.

As AI systems and services are designed, built and used, they place demands on resources such as energy, water and critical materials, which can create new environmental harms or exacerbate existing harms. Moreover, by driving up energy demand, AI systems and services can increase the challenge of moving to a decarbonised electricity system. There are actions that can be taken now, across the AI value chain, to better understand and reduce this unsustainable resource consumption and the related environmental impacts.

This report proposes **five foundational steps to begin progress towards environmentally sustainable AI**:

1. **Expanding environmental reporting mandates**
2. **Addressing information asymmetries across the value chain**
3. **Setting environmental sustainability requirements for data centres**
4. **Reconsidering data collection, transmission, storage, and management practices**
5. **Leading the way with government investment**

This is the first report on AI sustainability from our engineering responsible AI programme. It follows a workshop and a series of interviews with experts from industry, academia, civil society and policymakers. **The foundational steps this report proposes can be taken by governments to advance AI sustainability, and reduce and monitor the environmental impacts of AI systems and services**. These internationally applicable foundational steps are accompanied by recommended actions for the UK government. The recommended actions aim to facilitate UK leadership in AI sustainability and minimise the risk of lock-in to systems that are not sustainable in the long term.

As this programme continues, we will work to identify further actions to support and advance environmental sustainability in AI which can build on these foundational steps. Future activities for this programme will consider long-term interventions to further support managing and reducing the environmental impacts arising from AI systems and services – with consideration for their potential to deliver societal benefits.

Definitions

Environmental sustainability: managing and reducing environmental impacts across the lifecycle of a system or service in a manner that ensures present needs are met without compromising the ability of future generations to meet their own needs.[2]

AI: a suite of technologies that encompasses machine learning systems that use both symbolic methods, which represent knowledge and reasoning using symbols and rules, and non-symbolic methods, which recognise patterns in raw data. These systems are supported by computational resource and data consolidation. In recent years, these systems have used computation and data at increasingly large scales.[3]

AI value chain: We use the term 'value chain' to refer to four key components that, when brought together, enable AI systems and services. While a 'supply chain' may refer to the "system and resources required to move a product or service from supplier to customer" a 'value chain' builds on this concept "to also consider the manner in which value is added along the chain, both to the product or service and the actors involved."[4]

For this report we have categorised the value chain as being comprised of four interdependent components:

AI compute

The logic, memory and interconnect components stored within a server or network.

Vipra and Myers, *Computational Power and AI*

IT infrastructure

"Facilities or physical plants provide space for networking hardware, servers and data centres. It also includes the network cabling in office buildings to connect components of an IT infrastructure together."

IBM *What is IT infrastructure?*

Algorithms and data

Data is "the raw material that is processed by compute" while **algorithms** encompass "the source code that defines everything from the architecture of AI models to the specific methodologies employed in the training."

Sastry et al. *Computing Power and the Governance of Artificial Intelligence*

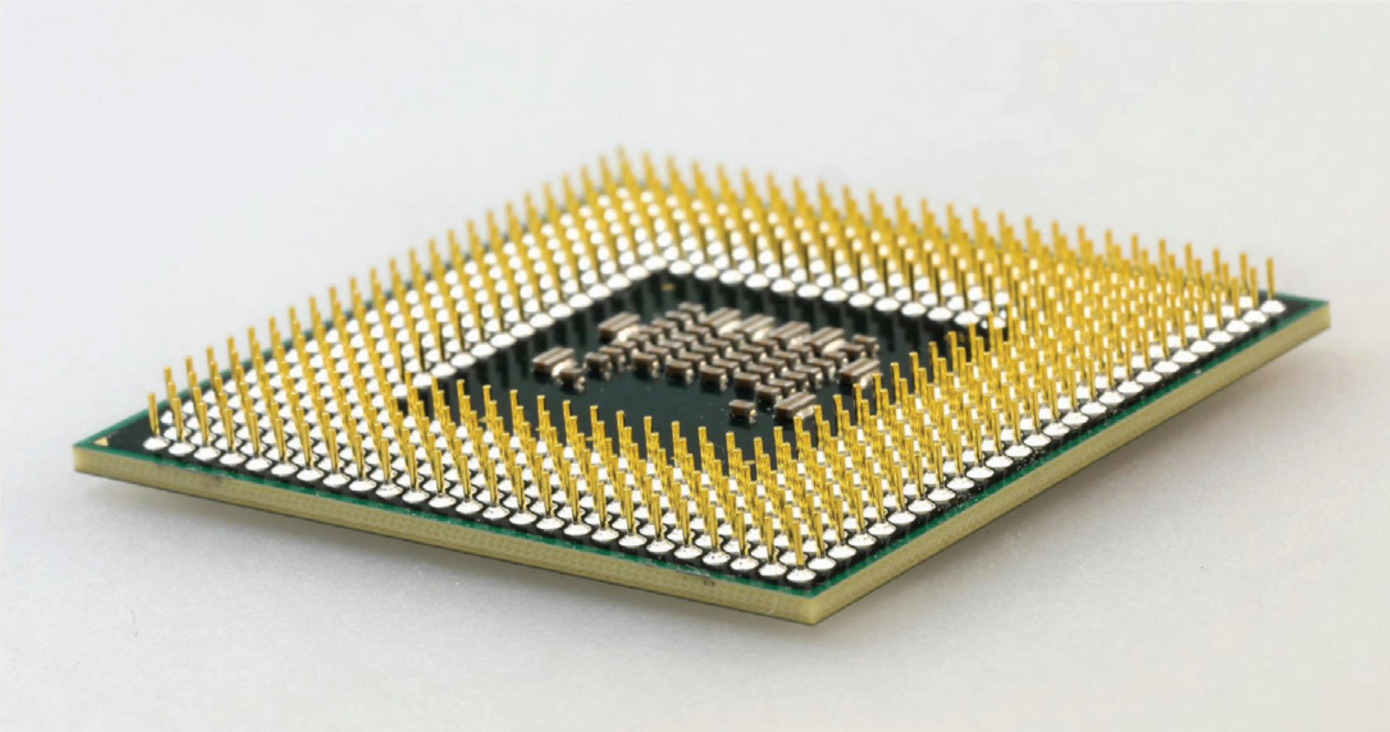
Interaction and use

Interaction and use encompasses "the impacts and opportunities created by the aggregated effects on societal structural changes by using (AI)." This includes procurement, use, and where AI indirectly influences organisational or individual behaviour.

ITU-T L1410

The environmental impacts of AI systems and services depend on interactions between components of the AI value chain:

- **AI compute**: the design of hardware components impacts the efficiency with which workloads are handled during operations. The production of, and the end-of-life strategy for, hardware components affects the lifecycle impacts of AI systems and services.
- **IT infrastructure**: where these facilities are located, draw their electricity, materials, water and other resources from, as well as how they are connected, built and managed at the end of their life, will affect the lifecycle environmental impacts of AI systems and services.
- **Algorithms and data**: approaches to data collection, transmission, storage, and management – as well as model design and training choices – can significantly reduce the demands of AI systems and services by reducing the need for state-of-the-art hardware, and resource consumption during training, deployment and inference.[5]
- **Interaction and use**: a considerable portion of a system's or service's impacts may be generated during inference (model use). The behaviour of end users therefore plays an important role in determining the overall environmental footprint of AI systems and services.



Introduction

AI has the potential to deliver significant benefits to the economy, society and the environment. To maximise the benefits delivered by AI systems and services, it is important to effectively manage environmental risks by actively monitoring and reducing negative environmental impacts.

The UK's target to reach net zero emissions by 2050 will require a concerted cross-economy decarbonisation effort. Use of AI and digital technologies is expected to play an important role in supporting businesses and transforming public services for the better, driving productivity and growth across the UK and enabling more rapid decarbonisation.[6]

In recent years, however, the rapid proliferation of AI has seen the environmental impacts associated with the development and use of these models and applications rise significantly.[7] Scaling IT infrastructures to support this demand has resulted in increased consumption of energy and water, increased use of critical materials and increased generation of heat.[8] For AI compute, the ongoing growth in semiconductor manufacturing is anticipated to continue, partially driven by demand for specialised hardware to support AI workloads.[9]

Ensuring AI contributes to a net reduction in environmental risks will require stakeholders involved in developing, deploying, and using AI systems and services to strive towards both efficiency and frugality. For the UK, this is critical to ensure AI demand does not contribute to total energy demand outpacing clean power generation, and does not place unsustainable demands on potable water and critical materials.

Published in January, the *AI Opportunities Action Plan*, which has received UK

> Ensuring AI contributes to a net reduction in environmental risks will require stakeholders involved in developing, deploying, and using AI systems and services to strive towards both efficiency and frugality

government endorsement, puts forward several recommendations intended to make the most of AI opportunities in the UK.[10] The government's response to the plan outlines intentions to develop a long-term compute strategy to help realise these opportunities – and, as part of this strategy, to assess environmental risks associated with AI infrastructure.

Presently, the ability of decision-makers to assess these risks is limited. Estimates of AI's environmental impacts vary due to a lack of data availability and consistent measurement methods. A lack of clarity on the boundaries between an AI system or service and the broader information and communication technology system further complicates quantifying these impacts. However, while the exact scale of environmental impacts may be uncertain, it is clear that current trends are unsustainable. This has been evidenced by companies such as Microsoft highlighting how AI uptake has impacted the achievability of their sustainability ambitions.[11]

This report from the National Engineering Policy Centre follows interviews and a workshop with experts from industry, academia, civil society and policymakers. Focusing on design, build and use choices, this report provides foundational steps and actions to improve the environmental sustainability of AI. To achieve AI sustainability at a systems level, steps and actions must be taken with consideration of social and economic factors, which will be addressed as this programme continues.

In the following sections, we provide context on the environmental impacts of AI systems and services, and opportunities for the UK to demonstrate leadership in AI sustainability internationally. We then propose foundational steps that can be taken by jurisdictions around the globe to minimise, and make more visible, the environmental impacts associated with AI. These foundational steps are supplemented by recommended actions that the UK can take today to begin progress towards environmentally sustainable AI.

Environmental impacts

This section provides an overview of the environmental impacts of AI systems and services, and how they relate to trends in development and uptake. A fuller analysis of the environmental impacts of AI systems and services can be found in Annex A.

Across their lifecycles, AI systems and services require energy, water, and critical materials. These resources are directly consumed during the production, use and end-of-life management of the compute and infrastructure hardware that underpins AI systems and services. They are also indirectly consumed during activities such as energy generation and distribution. Consuming these resources can cause air pollution, water pollution and thermal pollution, as well as creating solid waste. These impacts, in turn, negatively affect human and environmental health and can result in heightened risk of social inequity or exclusion – especially as the environmental impacts of AI are typically unevenly distributed.[12]

The software lifecycle includes "data collection and preparation, model development, training, validation, deployment, inference, maintenance and retirement."[13] Decisions made at each stage of the software lifecycle impact resource consumption by compute and infrastructure hardware.

In recent years, software development at the frontier of AI has tended to prioritise scale, size and capability, leading to increasing resource intensity across the software lifecycle – particularly impacting the training of models and their inference. This tendency has been compounded by the rapid expansion of both the development and uptake of AI.[14]

AI systems can take many forms, with different levels and profiles of resource intensity depending on their design and use. Resource consumption for an AI system or service can be understood as occurring in two stages: the training stage and the inference (or use) stage.[15] Depending on a system's design and the way it is used, the two stages may have different impacts on resource consumption. For example, an agent-based system may consist of multiple agents that collectively require fewer resources to train than a single large language model.[16] However, if these agents are designed to query one another via inference regularly, the system's total resource consumption may still increase.

While the long-term environmental impacts of different forms of AI systems and services remains to be seen, AI development and use up to this point has generally adhered to the following trends:

- **More resource-intensive training**: the scale of compute and data used to train models has been trending upwards. The training computation used for notable machine learning

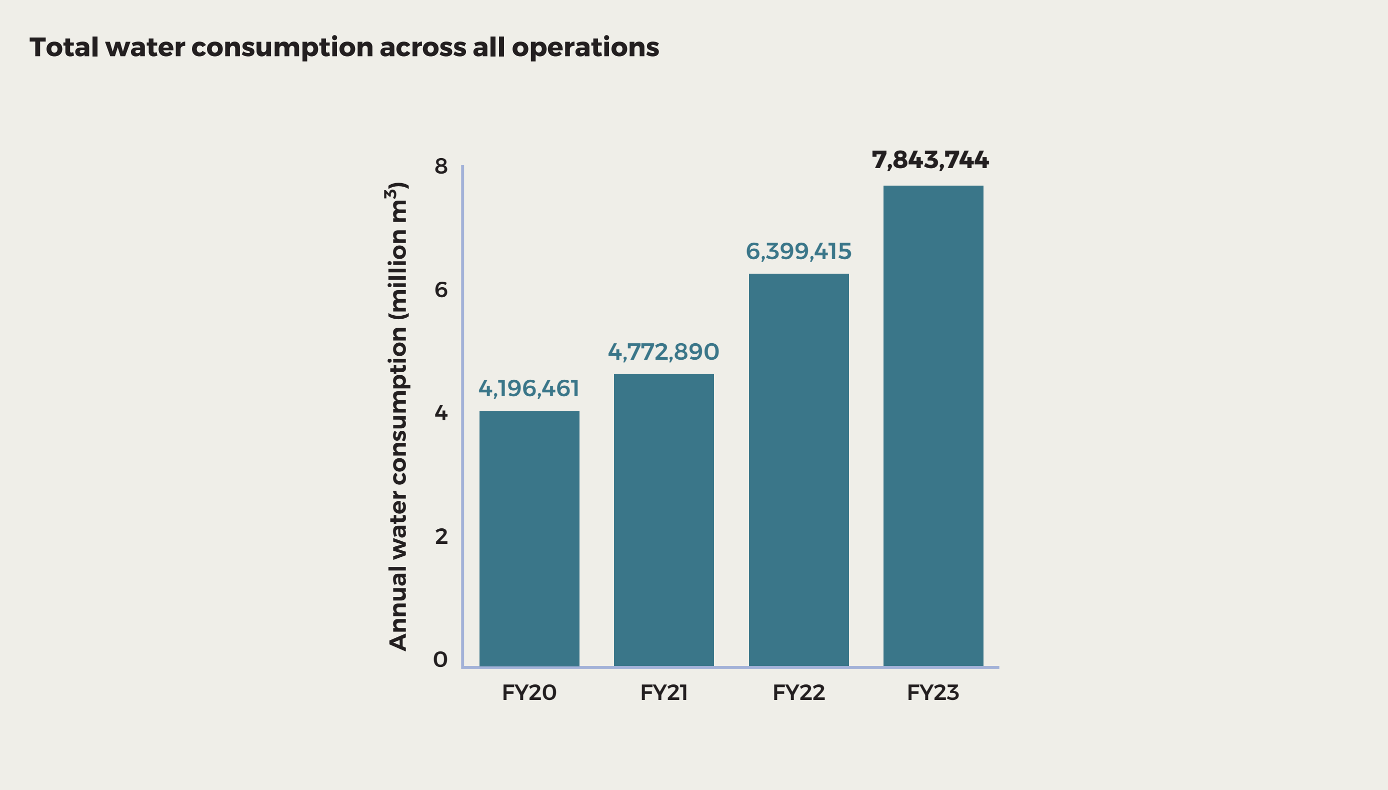


Figure 1 | Microsoft's total reported water consumption from financial year 2020 to 2023. Graph recreated from *2024 Environmental Sustainability Report*. Accessed 15 December 2024. © Microsoft

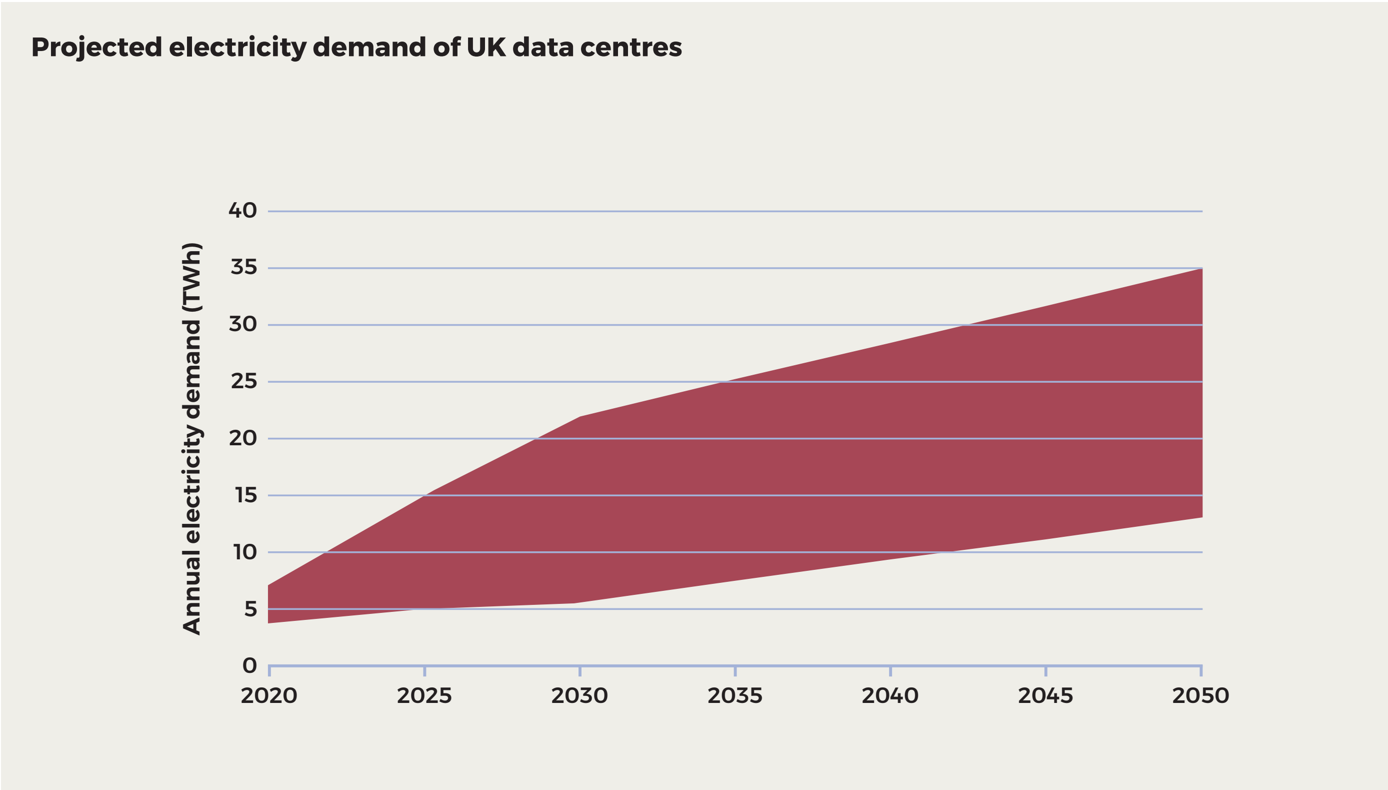


Figure 2 | Range of uncertainty of GB data centre electricity demand between today and 2050. Recreated from *Data Centres: What are data centres and how will they influence the future energy system?*, March 2022. © National Grid ESO

models has grown 4.4 times a year since 2010.[17] It has also been estimated that, if model training continues to follow current trends, that models could be trained on datasets roughly equal in size to the available stock of public human text data as early as 2026.[18]

- **Increasing demand for data storage**: data centre and endpoint storage capacity is predicted to grow from 10.1 zettabytes in 2023 to 21 zettabytes in 2027, partially driven by AI demand.[19] For context, one zettabyte is equal to one trillion gigabytes.

- **More resource-intensive inference**: increases in the size of inputs (represented by token limits) and the number of model parameters are resulting in more resource consumption at the use stage.[20] While a shift towards smaller models may impact this trend, where smaller models are assembled as part of an agent-based system, their overall impacts may be comparable to that of existing generative AI models.

- **More demand for data centres**: in 2023 there was 601 megawatts of new demand for data centres across the 14 largest markets in Europe, up from 546 megawatts in 2022. In the same period, 561 megawatts of new supply was delivered. This was the second time in five years that new demand exceeded new supply in Europe.[21] This, in turn, led to data centre vacancy rates in the top five European markets (Frankfurt, London, Amsterdam, Paris, Dublin), hitting an all-time low of 10.6% at the end of 2023.[22]

- **More demand for state-of-the-art hardware**: growth in semiconductor manufacturing is anticipated to continue, partially driven by demand for specialised hardware to support AI workloads. For example, generative AI chips are predicted to have reached more than $50 billion in sales for 2024. In the longer term, it is projected that AI chips could reach $400 billion in sales by 2027.[23]

As a result of these trends, the consumption of energy, water and critical materials attributable to the lifecycles of AI systems and services has been steadily increasing in recent years – and is projected to increase further (Figure 2). In the case of electricity consumption, some of the top-end estimates suggest that energy demand from AI could outstrip renewable generation capacity in some jurisdictions within the next 10 years.[24] Others suggest that surging power consumption from data centres driven by AI workloads could see total US energy demand outstrip generation capacity by 2028.[25] Even looking towards less dramatic projections – the consumption of energy and withdrawal of water from local networks poses a significant potential challenge, especially as data centres tend to be concentrated in certain localities.[26] Risks, at a local level, may be mitigated through a joined-up spatial strategy.[27]

Projected increases in resource consumption from AI do not only pose risks for the climate, the environment and human health. Where the consumption of resources to service AI workloads creates competition for, or even scarcity of, resources it will negatively impact other sectors of the economy. This, in turn, will create new challenges for industry to meet decarbonisation and growth goals – and will exacerbate cost-of-living pressures[28] and impact quality of life for affected communities.[29]

Opportunities for UK leadership

There is growing international concern over the environmental impacts of AI systems – among both jurisdictions[30] and industry.[31] In this context, there is an opportunity for jurisdictions such as the UK to demonstrate global leadership in AI sustainability, by developing and implementing pragmatic and effective regulation and guidance to monitor and reduce the environmental impacts of AI systems and services – as well as contributing to setting international standards where appropriate.

Incentivising environmental awareness from the producers, providers, partners and users of AI systems, as well as hardware manufacturers and data centre operators, may also create opportunities to grow the UK's relevance in the global AI market.[32] Opportunities include accelerating development and uptake of AI assurance technologies, alternative models and chip architectures, and smaller models.

- **AI assurance technologies**: AI assurance technology provides tools and techniques to measure, evaluate, communicate and mitigate the risks posed by AI – including environmental risks.[33] Creating regulatory requirements for compliance audits on environmental impacts and risks can drive uptake.[34] The UK AI assurance market has the potential to reach £6.53 billion (gross value added) by 2035 – a growth of over £5 billion from today.[35]
- **Alternative approaches to models and chips**: alternative models can drive reductions in data storage and compute requirements while delivering comparable, or even improved, performance on certain tasks. Alternative chip architectures, meanwhile, can potentially reduce processing time and power requirements for AI workloads. Combined, or in isolation, new models and chips may significantly reduce the environmental impacts that AI systems and services produce.

  Both alternative models and chips are active areas of research, within which several approaches with varying timelines for deployment are being developed. Should appropriate incentives be created for stakeholders in the AI value chain to prioritise environmental awareness, it is likely to increase appetites for alternative models and chips.
- **Smaller models**: smaller, task-specific models which can be trained from scratch, or developed by tuning or distilling a larger model, use smaller datasets and less compute than larger multi-purpose models, often without a significant drop-off in performance.[36] For example small language models (those not exceeding 7 billion parameters), are currently being deployed on end-user or edge devices. These small language models can deliver comparable, or even superior performance, to large language models.[37]

© National Grid

The UK ecosystem has existing strength in research and access to talent, making it well suited to meet rising demand for technologies and techniques designed to help monitor and reduce the environmental impacts of AI systems and services.[38] Where these technologies and techniques can be appropriately developed in an open-source environment, they also present an opportunity to further improve AI sustainability through resource sharing.

The proliferation of these technologies and techniques is also likely to benefit enterprise users of AI systems. In particular, small and alternative models can meet the needs of many businesses while reducing overhead costs. Office for National Statistics data suggests businesses are primarily adopting AI to improve cybersecurity (35%) and create efficiencies (35%).[39] These are use cases that, for many smaller enterprises, do not involve complex data analysis – and where adopting small models should be encouraged where appropriate.

Small models can have additional benefits for enterprise. For example, using smaller models reduces the risk of incurring technical debt, which is the "extra work required when developers choose an easy short-term solution over the best long-term approach."[40] Technical debt can be accrued because of insufficient investment in company culture, IT systems, data architecture, and security practices.[41] Larger models require data to be more diverse, plentiful and of higher quality – increasing the level of investment required by enterprise users. This places enterprises who use inappropriately large models at greater risk of incurring technical debt – which, in turn, creates a financial burden as resolving technical debt requires investment.[42] This is a growing risk internationally, with estimates suggesting current levels of technical debt in the US alone would require over $1.5 trillion to fix – indicating that, for enterprise users, there are notable economic benefits associated with adopting smaller models, beyond a reduced cost for computation.[43] In this context, including sustainability as a key criterion while procuring AI systems and services may be seen as a 'low regrets' decision.

Beyond creating opportunities for the development of specific technologies, incentivising environmental awareness via regulation, standards, and guidance can also improve the quality and trustworthiness of data available to decision-makers. This, in turn, could help improve forecasting capabilities – both in business and government.

This is especially important for government, as a lack of clarity on the precise scale of AI's environmental impacts currently hinders policymakers' ability to forecast and prepare for future risks, or opportunities. Presently, there is a high degree of uncertainty in forecasts of future energy, water and critical materials demands from AI.[44] This challenge is further compounded by uncertainty regarding the veracity of reported figures. An investigation from the *Guardian* in September 2024 found that emissions from Google's, Microsoft's, Meta's and Apple's data centres could be up to 662% higher than officially reported.[45]

It is an opportune moment to focus on better data – as several jurisdictions (including the UK) are currently developing internationally relevant mandatory reporting requirements for climate-related disclosures in line with the International Sustainability Standards Board (ISSB) standards.[46] The International Organization for Standardization (ISO) has developed standards to support these requirements, which are anticipated to improve data availability and have a substantial impact on industry. However, policymakers in these jurisdictions will need to determine how these reporting requirements explicitly apply to data centre operators, hardware manufacturers, and AI producers, providers, and partners.

In the UK, there is an ongoing assessment[47] regarding the adoption of IFRS S1 and S2.[48] Should this assessment end in a positive endorsement, this will result in the creation of UK Sustainability Reporting Standards which may then be implemented as part of the government's wider work on creating sustainability reporting requirements.[49]

The scale of the challenge and the risk of enterprises or governments 'locking in' to unsustainable systems demands a further acceleration of efforts. While quality and trustworthiness issues with currently available environmental data may hinder the ability for decision-makers to develop long-term action plans, there are several steps that can be taken in the short term to reduce impacts and improve monitoring capabilities.

> The scale of the challenge and the risk of enterprises or governments 'locking in' to unsustainable systems demands a further acceleration of efforts

Foundational steps

This report sets out five foundational steps that jurisdictions can take towards sustainable AI in the short-term. These steps are also critical to better understanding and managing the environmental impacts of AI systems and services in the longer term.

While these foundational steps are intended to be internationally applicable, we supplement these with practical actions the UK government can take in the near term. It is critical that action is taken now to ensure that near-term investment in infrastructure, research, and adoption aligns with our long-term net zero and environmental sustainability aspirations.

Positioning the UK as a leader in AI sustainability will require effective collaboration across several government departments and regulators. Government's intention to develop an AI Energy Council provides an important forum for collaboration on balancing the supply and demand for energy and compute. However, the AI sustainability challenge extends beyond energy – the usage of critical materials and water by systems and services, and the design and use of those systems and services, must also be considered.

Collaboration between departments should be led by the Department for Science, Innovation and Technology, who must set clear ambitions and criteria to guide the development and deployment of AI towards sustainability. This should be done in collaboration with the Department for Energy Security and Net Zero; Department for Business and Trade; and Department for Food and Rural Affairs, who will also play a critical role in improving reporting standards. Further collaboration must be sought with the devolved administrations; Department for Education; the Department for Digital, Culture, Media and Sport; the Ministry of Housing, Communities and Local Government; the Information Commissioner's Office (ICO) and standards development organisations to ensure environmental sustainability is embedded across the AI value chain.

1. Expanding environmental reporting mandates

Establish reporting mandates for meaningful AI-specific metrics to support the implementation of IFRS S2 reporting requirements and to improve understanding of the environmental impacts of AI systems and services.

This report recommends that the UK considers:

- **Using the ongoing consultation on a UK Green Taxonomy as an opportunity to consider the**

It is critical that action is taken now to ensure that near-term investment in infrastructure, research, and adoption aligns with our long-term net zero and environmental sustainability aspirations

development of an assessment framework for data centres, and for AI systems and services – which could further support improved clarity for entities on climate-related risks and opportunities.[50]

- **Making reporting on energy consumption, energy sources, water consumption and withdrawal, water sources, carbon emissions, e-waste recycling and the Power Usage Effectiveness (PUE) of data centres mandatory**. These requirements should be based on *Industry-based Guidance on implementing Climate-related Disclosures*.[51] They should also refer to emerging resource consumption data sharing practices such as those outlined in the proposed UK Net Zero Carbon Buildings Standard (UKNZCBS).[52]
- **Collaborating on new open and trustworthy data collection and reporting protocols and tools** to enable data sharing across the value chain for producers, providers and partners of all sizes. These could be shared through the proposed "AI Knowledge Hub".[53]
- **Introducing requirements for reporting on the reuse of server equipment**. A 2021 study indicated that 28% of companies globally reuse or repurpose IT hardware internally[54] – and initiatives such as the Circular Electronics Partnership (CEP) actively promote circular practice.[55] Introducing reporting requirements could improve policymakers' understandings of obstacles to reuse, or opportunities, enabling them to play a more active role in continually improving data centre circularity. Data centre circularity, beyond reducing environmental impacts, can also minimise exposure for UK-based data centre operators and enterprise to supply chain risk.
- **Introducing requirements to report on the balance of workloads at data centres**. The UK should collaborate with other jurisdictions to develop a measurement standard for assessing and reporting the balance of workloads in a data centre without compromising security or business sensitivity.
- **Implementing reporting requirements for the energy, carbon and water consumed during the development and use of AI services**. This could be achieved by implementing the requirements to create 'passports' for digital products, as described in the International Telecommunication Union's proposed standard, ITU-T L.1071 (a model for digital product passport information on sustainability and circularity).[56]

Enforcement of meaningful AI reporting requirements as part of IFRS S2 implementation is vital to ensuring the UK pursues AI opportunities with appropriate consideration of environmental risks. This data should be regularly reviewed by the AI Energy Council, and should inform future updates to the long-term compute strategy that UK government has committed to in their response to the *AI Opportunities Action Plan*.[57] This will require significant investment from government to build the necessary infrastructures to ensure compliance.

While additional reporting requirements may be seen as a burden to business, robust reporting will improve visibility of potential ESG risks and

© This is Engineering

opportunities. Robust reporting can also offer UK enterprises of all sizes improved access to capital as ESG funds continue to perform strongly[58] – and Nations and Regions Investment Funds increasingly emphasise sustainable investment.[59] Introducing additional reporting will also create opportunities for the UK to influence international standards for reporting and expand the AI assurance market.

2. Addressing information asymmetries across the value chain

Address information gaps along the AI value chain to ensure that all stakeholders have a clear understanding of the environmental impacts of AI systems and services. Providing information and education on the environmental impacts of AI systems and services, and responsible usage, can help encourage individual and enterprise users to find appropriate models for their use cases, and drive more environmentally sustainable practices across the AI value chain.

Educational efforts could be supported by ensuring:

- **Producers and providers develop and use appropriate communication tools to convey real-time information on the environmental impacts of AI systems to users**. These should be developed with educational and behavioural scientists and human factors experts to ensure they support informed decision-making from users about the appropriate tool for their computation request.
- **Design for environmental sustainability is embedded in AI, computer science and communications courses** in higher education, introduced in digital further education courses,

Effectively communicating the environmental impacts of AI systems and services to users can drive behaviour change from both the users and developers

and included in upskilling support for AI adoption.

- **As part of the curriculum review, environmental sustainability should be included in AI, ICT, and computing education provided in schools and colleges** to develop greater environmental awareness amongst future developers and users. This should be a component of a wider effort to educate on the environmental impacts of digital technologies.

- **Setting standards to ensure appropriate practices are used by providers, producers, and partners of systems and services**. Given that software development frequently occurs across national boundaries, these standards should be open where possible.

- **Mandate publishing of data and model cards** to inform users (both individual and enterprise) on the environmental impacts of systems, and to educate on the benefits of small and more sustainable AI.[60]

Effectively communicating the environmental impacts of AI systems and services to users can drive behaviour change from both the users and developers. Moreover, greater awareness and education will enable more effective collaboration between stakeholders across the value chain.

3. Setting environmental sustainability requirements for data centres

Setting requirements for new data centres of all sizes and establishing a timeline for legacy data centres to meet these requirements can support a reduction in data centre resource demands – and the environmental impacts of AI model training and inference.

The requirements should include:

- **Reduce potable water usage for all data centre activities and zero potable water for cooling**. This should be supported by best-practice guidance in ISO 46001:2019 (Water efficiency management systems)[61], and securing 'zero potable water for cooling' commitments from data centre operators.[62]

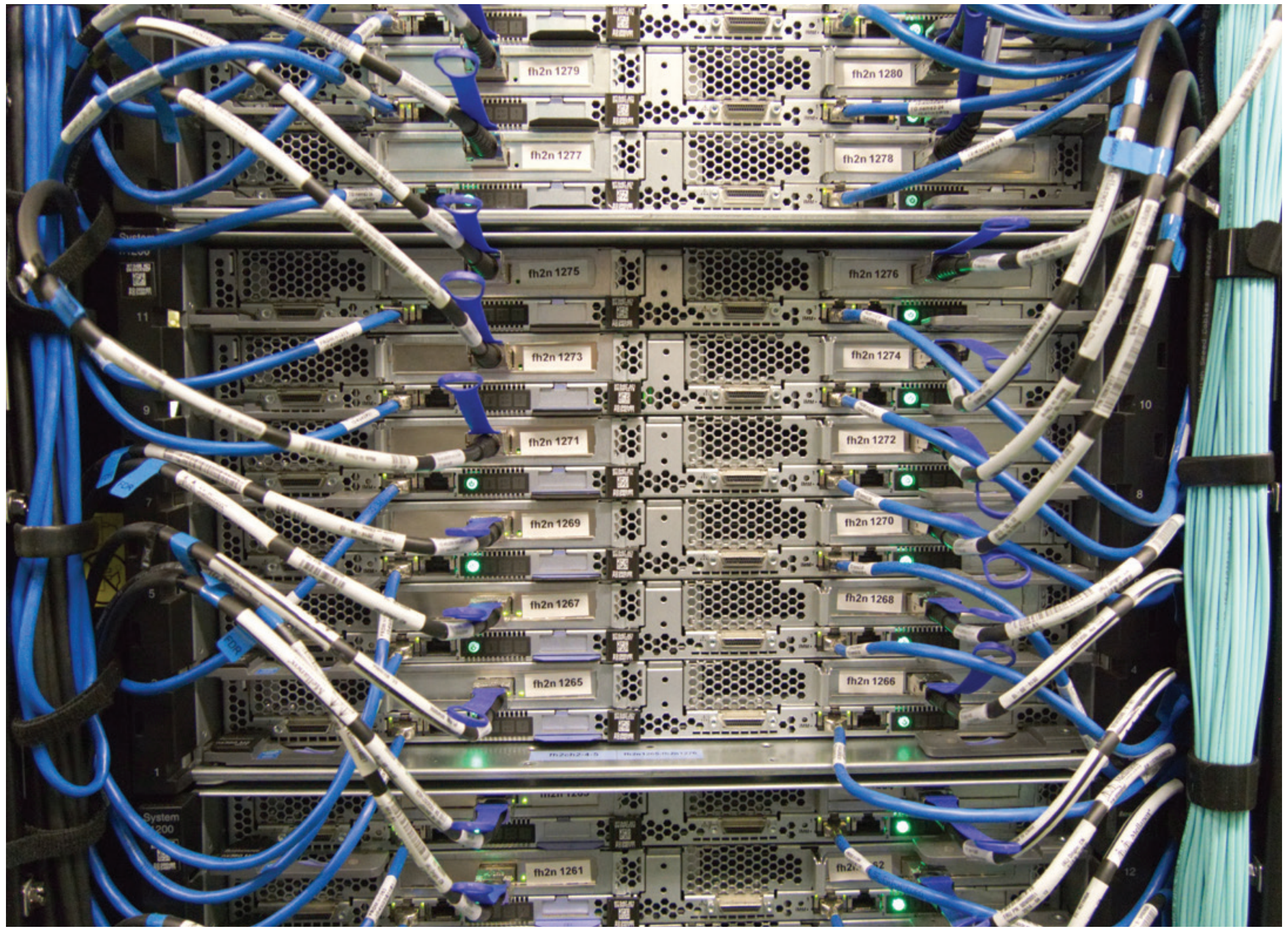
- **Maximise the reuse and recycling of server equipment**. This can be supported by referring to existing standards for the circular economy[63] and repair and refurbishment, as well as by developing targets for hardware reuse[64], encouraging the development of open-source hardware designs[65], and supporting data centre operators to follow best practice for IT asset management as outlined in ISO/IEC 19770-1.[66] To further support this goal, remanufacturers can be encouraged to follow best practice for reworking and remarketing of hardware as described in BS 8887-211.[67]

- **Recover waste heat and enable its reuse**. This can be supported by mandating quantities of waste heat be made available for reuse and through coordinated spatial planning. This should be informed by the proposed UKNZCBS, which outlines key considerations for sharing waste heat with district networks, and for monitoring and measurement[68] in a way that balances the challenges of data centre heat export.[69]

- **Match energy used with 100% carbon-free energy certificates**. Renewable Energy Certificates should be backed by clear, auditable, and time- and site-specific data. Certificates should prioritise emissionality, which is "the degree to which a new [renewable energy] project leads to actual [greenhouse gas] emissions reductions by displacing local fossil-fuel use".[70] They should also discourage carbon leakage or emissions offshoring.[71] Agreements should be regularly reviewed to ensure they deliver tangible benefits to local grids. Requirements could be supplemented by creating additional requirements for data centres to deliver year-on-year improvements on their Renewable Energy Factor as outlined in ISO/IEC 30134-3:2016.[72]

- **Improve PUE through ambitious but realistic targets**. This should be informed by the proposed UKNZCBS[73] and supported by best-practice guidance on reducing energy use outlined in ISO 50001:2018 (Energy Management Systems).[74] While PUE has been criticised as a flawed metric, where it is reported on alongside other metrics it can contribute to improved understanding of data centres' environmental impacts.[75]

Introducing these requirements now and applying them to all data centres, starting with those in AI Growth Zones, can reduce future pressures on the electricity grid and water sources, and offers opportunities for expanded heat reuse. They can also help to reduce data centre demand for critical materials – which is currently contributing to long lead-in times for data centre equipment[76], and may impact material availability for economy-wide decarbonisation efforts.[77]



© dlohner, pixabay

In the UK context, new requirements for data centres may have a negative impact on new data centres from locating in the UK in the short-term. However, many of these actions are mandated by the EU Energy Efficiency Directive[78] – and are also already being implemented by data centre operators internationally in a piecemeal way.[79]

4. Reconsidering data collection, transmission, storage, and management practices

Create incentives to reduce data collection, transmission and storage demands, and improve data management, to reduce the lifecycle environmental impacts of AI systems and services – and encourage frugality in AI procurement. These incentives could be delivered through:

- **Developing AI sustainability good practice** based on standards and best practice to embed consideration of environmental impacts associated with data collection, transmission, access, quality, minimisation, deletion, and storage for stakeholders across the value chain. Good practice should be applicable across the lifecycles of services and systems from problem specification to the long-term impact of data practices on AI inference costs.

- **Revising legislation mandating data retention** to go beyond existing 'data protection principles' and the ICO recommendation that organisations set up a data retention policy. This could be supported by use of ISO standards, such as the ISO 8000 series which provides guidance on data quality. This could be used to develop criteria for the long-term retention of data.[80]

- **More robust data minimisation, compression and storage limitation policy**, with more structured guidance to set up a data retention policy that encourages storage limitation, purpose limitation, and data minimisation based on ISO standards for information security, retention, and deletion (in particular, ISO/IEC 27555:2021).[81]

- **Using the development of the National Data Library as an opportunity to incentivise, demonstrate, and drive good practice in environmentally sustainable data management**, as part of its mission to enable efficient, effective, secure and acceptable access to data. Opportunities to expand the scope of assets included in the library to include private data, models and model components should be considered as this could significantly reduce the environmental impacts of development and storage.

Reducing data collection, transmission, and storage demands will reduce demands on energy, water and critical materials. This may reduce data availability; however, it may also drive smarter and more focused data collection from model providers, producers, and partners, as well as enterprise.

5. Leading the way with government investment

Embedding environmental sustainability in AI policy, procurement and support programmes will encourage proportionality and efficiency from data centre operators; hardware manufacturers; and AI providers, producers, partners and users.

Embedding sustainability into government's approach to investing in AI could involve:

- **Ensuring longevity and efficiency from public compute provisions without compromising performance**, using ISO 30134-5 and the ITEUsv metric to help minimise server waste by encouraging high utilisation levels.[82]

- **Embedding environmental sustainability in procurement guidelines for AI systems and supporting infrastructure** so the best, and most sustainable compute is available to UK government. This could be supported by referring to guidance on integrating sustainability in procurement decisions as outlined in ISO 20400.[83] Best practice lessons learned could be shared with industry via the proposed "AI Knowledge Hub".

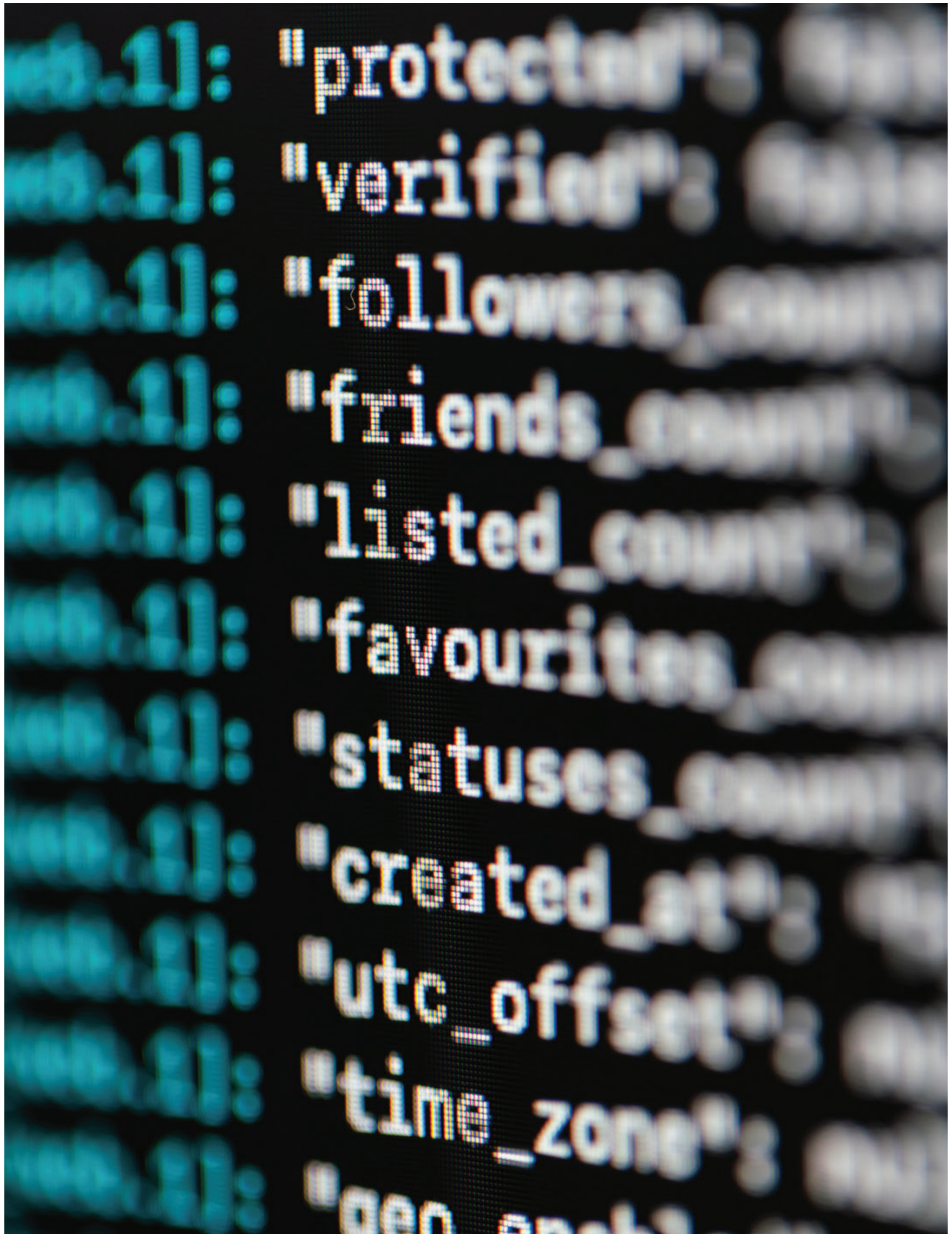
- **Establishing AI sustainability research as a key part of the UKRI AI strategy**, to further grow funding available for sustainability research to better understand and reduce environmental impacts across the value chain, including AI compute, IT infrastructure, algorithms and data, and interaction and use.

- **Developing and implementing an agreed approach to AI environmental impact reporting as part of government business cases and research funding bids** and integrate this approach into existing environmentally aware funding concordats and responsible innovation frameworks. This could be supported by the soon-to-be published ISO/IEC TR 20226[84] and CEN TR 18145[85] which will provide calculation methodologies for AI's carbon footprint, and guidance on frugal AI. Government should also consider lessons learned from the implementation of environmentally aware funding initiatives in other jurisdictions.[86]

- **Incentivising the use of smaller models, particularly task-specific models, as well as alternative model approaches and edge solutions**, where appropriate. Where edge solutions are used to move processing to devices nearer the user, they can reduce environmental impacts from data centres and data transmission. This can be supported by the development of smaller models, which can run on edge devices. The use of smaller, and alternative models, can be encouraged through adoption programmes to ensure government, researchers, enterprise and individual users have a diversity of choice so they can use the most appropriate AI models for their use case.

- **Embedding environmental sustainability as a core principle for the AI Growth Zones**, to ensure plans to accelerate the build-out of data centres appropriately address environmental risks. This should include consideration of how data centre location may impact local energy and water supplies – and how AI Growth Zones are influencing the lifecycle environmental impacts of AI systems and services by altering interaction within the AI value chain.

Embedding sustainability as a criterion for government investment can help to minimise environmental risks associated with expanding public compute provision, and creates incentives for frugal development and procurement. Moreover, it can also support the UK to develop a 'niche' and grow its advantage in the development and deployment of small and alternative models focused on solving well-defined problems, and assurance models designed to support regulatory compliance. Such an approach to investment can also create opportunities for alternative approaches to hardware, by offering incentives for investment and potentially generating new or alternative opportunities for AI deployment.



Conclusion

These foundational steps are intended to be implemented by jurisdictions in the short-term to monitor and reduce the environmental impacts of AI systems and services. Together, they can enable better understanding of environmental impacts, more sustainable infrastructure, and clear leadership from governments to accelerate change. These foundational steps are accompanied by actions that outline specific activities that could be undertaken by the UK government.

Demonstrating leadership in monitoring and reducing the environmental impacts of AI systems and services also presents opportunities to align environmental goals with broader social and economic aspirations. The transition towards environmentally sustainable AI must address not only environmental concerns, but also ensure social and economic equity and inclusivity, and should therefore appropriately reference Just Transition frameworks.[87]

This report is the first on AI sustainability from our engineering responsible AI programme, which will continue through 2025, building on these foundational steps to produce further advice for decision-makers. This will include additional consideration of how the environmental impacts of AI systems and services can be monitored and reduced – including direct impacts from use and indirect impacts arising from behavioural changes precipitated by use.

There is a unique opportunity at the present moment, while AI remains an emerging technology, to embed environmental sustainability as a key requirement for future systems and services. However, finding the appropriate means by which to promote sustainability via regulation, standards, and guidance over the longer term requires policymakers to have greater clarity on where and how the environmental impacts of AI systems and services arise.

No jurisdiction will be able to solve AI sustainability on its own, but there is an opportunity now for leaders to establish themselves and facilitate collaboration on this global challenge. We welcome opportunities to collaborate with likeminded organisations who are passionate about engineering responsible AI and minimising the environmental risks to collectively drive change.

Annex A
Environmental impacts

This section provides further evidence on how AI systems are impacting demand for energy, water and critical materials.

Energy

Estimates of AI's future energy demands vary and usually focus on the energy demand from data centres. However, while estimates of future demand have varied significantly, they consistently indicate that increasing demand from AI will impact electricity grids on a local level at a minimum.

The divergence between the IEA's *Electricity 2024* and *World Energy Outlook 2024* exemplifies how significantly estimates can vary. The former suggests data centre energy consumption could expand by as much as 540 terawatt-hours by 2026,[88] an amount greater than Germany's total electricity consumption in 2021.[89] The latter suggests data centre electricity demand could grow by 223 terawatt-hours by 2030.[90] This is comparable to the total electricity consumption of Spain in 2021.[91] While significantly less than the projection offered by *Electricity 2024*, the *World Energy Outlook 2024* scenario still projects data centre demand to negatively affect electricity markets at the local level.

In some jurisdictions, this may mean bringing non-renewables online to service the energy demands of AI. It has been estimated that, "US data centres' increasing energy demands will lead to additional gas demand of between 3 and 6 billion cubic feet per day by 2030."[92] For context, the whole of the UK consumed an estimated 6.14 billion cubic feet per day in 2023.[93]

In the UK, 2024 brought concern over future energy demands from data centres with John Pettigrew, Group CEO of National Grid, suggesting that total energy demand from data centres in the UK could increase sixfold in the next 10 years.[94] These concerns may be heightened in the future as agent-based systems gain greater popularity. Agent-based systems require a higher level of inference than more traditional generative AI systems, which may cause energy costs to spiral – as, already, inference has been estimated to account for as much as 80% of AI computational demand.[95]

This increasing demand adds to the systems-level challenge of achieving grid decarbonisation. According to the *Clean Power 2030 Action Plan*, current projections indicate that for the UK to reach its goal of at least 95% of energy generation coming from clean power, current wind and solar capacity must increase by approximately 2.7 times.[96] As such, where AI continues to add to electricity demand, it will increase the scale of the clean energy generation challenge.

Water

Water consumption from data centre operators continues to rise year on year according to major provider reports. For instance, both Google and Microsoft have reported year-on-year increases in operational water consumption since 2020 (Figure 1).[97] Many of these water withdrawals come from potable water sources: Google reported that approximately 78% of their global water withdrawals came from potable sources in its 2023 fiscal year.[98]

However, while much of the publicly accessible data cites global figures for water withdrawals and consumption, water is a local issue. Data centres, generally, withdraw and consume water from nearby sources, meaning that businesses and communities located in regions with a high concentration of data centres could be exposed to the risk of water scarcity.

In the UK, should data centre consumption of potable water continue to rise, water scarcity issues may too arise in several regions. According to analysis from the Environment Agency, total public water supply demand deficit in England is estimated to reach 3040 million litres per day in 2040 and 4860 million litres per day in 2050.[99] By 2030, seven regions in England are predicted to be severely water stressed.[100] London, which accounts for 77.3% of the UK's data centre market, is amongst those regions.[101]

Water, however, is not exclusively consumed during the use phase of an AI system. Significant quantities of water are consumed in the manufacturing of compute components. An average chip manufacturing facility today can use over 37 million litres of ultrapure water per day – as much water as is used by 33,000 US households every day.[102] The global semiconductor industry uses an estimated 1.2 trillion litres of water every year.[103] Water is also consumed in significant quantities during electricity generation, "through cooling at thermal power and nuclear plants and expedited water evaporation caused by hydropower plants."[104]

Critical materials

The consumption of critical materials across the AI lifecycle creates both environmental and strategic risks. The hardware components, including both general-purpose and specialised AI components, that sit within data centres, networks, and user terminals, are made up of a variety of critical materials including antimony, gallium, indium, silicon, and tellurium.

Where recycled critical materials are not available, they must be sourced via mining. Mining operations cause environmental harms such as the loss of land (including carbon sinks), loss of habitat and biodiversity through direct displacement and chemical pollution, and drought and freshwater pressures that can impact local ecosystems.[105]

Demand for these materials is rising and specialised AI components are a particularly strong driver of demand. Taiwan Semiconductor Manufacturing Company Limited (TSMC) has pledged to double its production to meet this demand.[106] This is reportedly driven in no small part by NVIDIA, which has announced plans to increase the rate at which it rolls out new products.[107]

This increasing demand poses a strategic risk, as the complex supply chains enabling compute production are highly internationalised. Semiconductor chip production inputs cross around 70 international borders.[108] Disruptions to these international supply chain flows can cause shortages, which can in turn cause significant economic disruption. For example, a silicon supply shortage that started in 2020 was estimated to have disrupted global GDP growth by 1% in 2021 and affected 169 sectors.[109]

Recycling and reuse can both reduce the demand for critical materials and reduce strategic risk associated with exposure to supply chain disruption. The UK performs particularly poorly in this area, generating the second highest amount of e-waste per capita at 23.9 kg in 2020.[110] Between 2008 and 2022, IT and telecoms e-waste almost doubled in the UK.[111]

While there is relatively little data available on total e-waste generation attributable to the use of AI systems, one study focused on generative AI found that total global accumulation of e-waste from these systems could reach between 1.2 and 5.0 million tonnes during 2020 to 2030, and that implementing circular economy strategies along the generative AI value chain could reduce e-waste generation by between 16 and 86%.[112]

Industry has begun to act on e-waste – with some of the larger stakeholders in the value chain creating initiatives to improve reuse and recycling rates. These initiatives have produced improvements, albeit not consistently. For instance, Microsoft has seen its reuse and recycling rates for servers and components fluctuate from 86.7% in fiscal year 2020, to 76% in fiscal year 2021, to 82% in fiscal year 2022, to 89.4% for fiscal year 2023.[113]

Annex B
Acknowledgements

The report has been developed by the Royal Academy of Engineering in partnership with the Institution of Engineering and Technology and BCS, the Chartered Institute of IT, under the National Engineering Policy Centre (NEPC).

Working Group

This report was delivered by a National Engineering Policy Centre Working Group made up of the following experts.

Chair: Professor Tom Rodden CBE FREng FRS FBCS, Pro-Vice-Chancellor of Research & Knowledge Exchange and Professor of Computing, University of Nottingham

Professor Adisa Azapagic MBE FREng FRSC FIChemE, Professor of Sustainable Chemical Engineering, University of Manchester

Alex Bardell FBCS, Founder, SDAdvocate

Professor Ana Basiri, Professor of Geospatial Data Science, University of Glasgow

Professor Mandy Chessell CBE FREng FBCS HonFIED, Founder of Pragmatic Data Research Ltd and President of the Institute of Engineering Designers

Dame Dawn Childs DBE FREng FICE FIMechE FRAeS, CEO of Pure Data Centres Group

Professor Steve Furber CBE FREng FRS, Professor Emeritus in the Department of Computer Science, University of Manchester

Professor Sarvapali (Gopal) Ramchurn FIET, Professor of Artificial Intelligence, University of Southampton and CEO of Responsible AI UK

Professor Stephen Roberts FREng FIET, Professor of Machine Learning, University of Oxford

Dr Eve Schooler, Royal Academy of Engineering Visiting Professor of Sustainable Computing, University of Oxford

Professor Shannon Vallor, Baillie Gifford Chair in the Ethics of Data and Artificial Intelligence, University of Edinburgh

Dr Kommy Weldemariam, Director of Sustainability Science and Innovation, Amazon

Academy Staff

Eliot Gillings, Policy Advisor, Digital and Physical Infrastructure

Dr Alexandra Smyth, Head of Policy – Infrastructure and resilience

References

1 Sophia Falk and Aimee van Wynsberghe, *Challenging AI for Sustainability: What Ought It Mean?*, AI and Ethics 4, no. 4, 31 July 2023.

2 This definition has been adapted from: *ISO Guide 82:2019*.

3 Pascal Hitzler et al., *Neuro-Symbolic Approaches in Artificial Intelligence*, National Science Review 9, no. 6, 1 June 2022.

4 Timea Walker, *What Is a Value Chain? Definitions and Characteristics*, Cambridge Institute for Sustainability Leadership (CISL), 5 July 2021.

5 Boyuan Chen et al., *Model Cascading for Code: Reducing Inference Costs with Model Cascading for LLM Based Code Generation* arXiv, 24 May 2024.

6 *AI Expert to Lead Action Plan to Ensure UK Reaps the Benefits of Artificial Intelligence*, GOV.UK, 26 July 2024.

7 *AI Has an Environmental Problem. Here's What the World Can Do about That*, United Nations Environmental Programme, 21 September 2024.

8 Generated heat, if properly handled, can be reused – availing benefits to surrounding homes and businesses. [Amber Jackson, *Using Data as Power: Data Centre Heat Reuse Technology*, Data Centre Magazine, 15 January 2024.] However, if improperly handled, generated heat can cause thermal pollution. [Andrey Semenov and Ekaterina Oganesyan, *Data Centers Environmental Impact Assessment Features*, E3S Web of Conferences 311, 19 October 2021.]

9 Semiconductor manufacturing energy consumption is projected to double from current levels, partially due to AI demand – consuming 237 terawatt hours (TWh) of electricity globally by 2030, roughly equivalent to Australia's 2021 electricity consumption. [*Invisible Emissions: A Forecast of Tech Supply Chain Emissions and Electricity Consumption by 2030*, Greenpeace East Asia, April 2023.]

10 *AI Opportunities Action Plan: Government Response*, GOV.UK, 13 January 2025.

11 "In 2020, we unveiled what we called our carbon moonshot. That was before the explosion in artificial intelligence," Microsoft president Brad Smith said in an interview with Bloomberg. "So, in many ways the moon is five times as far away as it was in 2020, if you just think of our own forecast for the expansion of AI and its electrical needs." [Justine Calma, *Microsoft's AI Obsession Is Jeopardizing Its Climate Ambitions*, The Verge, 15 May 2024.]

12 Shaolei Ren and Adam Wierman, *The Uneven Distribution of AI's Environmental Impacts*, Harvard Business Review, 15 July 2024.

13 United Nations Environment Programme, *Artificial Intelligence (AI) End-to-End: The Environmental Impact of the Full AI Lifecycle Needs to Be Comprehensively Assessed – Issue Note*, 21 September 2024.

14 Generative AI has been adopted at a faster pace than PCs or the internet [Alexander Bick, Adam Blandin, and David Deming, *The Rapid Adoption of Generative AI*, Federal Reserve Bank of St. Louis, 23 September 2024.]

15 Alexandra Sasha Luccioni, Yacine Jernite, and Emma Strubell, *Power Hungry Processing: Watts Driving the Cost of AI Deployment?* arXiv, 28 November 2023.

16 An agentic AI system "uses sophisticated reasoning and iterative planning to autonomously solve complex, multi-step problems… Agentic AI systems ingest vast amounts of data from multiple sources to independently analyse challenges, develop strategies and execute tasks." [Erik Pounds, *What Is Agentic AI?*, NVIDIA Blog, 22 October 2024.]

17 Compute as measured in FLOP. [Jaime Sevilla and Edu Roldán, *Training Compute of Frontier AI Models Grows by 4–5x per Year*, Epoch AI, 28 May 2024.]

18 This estimate was made on the assumption that current model training and data generation trends continue. It also assumes that "10% and 40% of deduplicated web data can be used for training without significantly compromising [model] performance." [Pablo Villalobos, *Will We Run Out of Data? Limits of LLM Scaling Based on Human-Generated Data*, Epoch AI, 6 June 2024.]

19 *Data Centers 2024 Global Outlook*, JLL Research, January 2024.

20 *Exponential Growth of Parameters in Notable AI Systems*, Our World in Data, 8 January 2025.

21 *Demand for European Data Centres Outstripped Supply Last Year Despite Rising Rental Rates*, CBRE, 16 February 2024.

22 Dan Robinson, *Europe's Datacenter Dilemma Is That Hyperscalers Are Hogging Them All*, The Register, 20 February 2024.

23 *2024 Semiconductor Industry Outlook*, Deloitte United States, 2024.

24 *AI Is Wreaking Havoc on Global Power Systems*, Bloomberg.Com, 21 June 2024.

25 *Utilities Must Reinvent Themselves to Harness the AI-Driven Data Center Boom*, Bain & Company, October 2024.

26 *World Energy Outlook 2024 – Analysis*, International Energy Agency, October 2024.

27 *National Planning Policy Framework*, GOV.UK, 12 December 2024.

28 *The Cost to Life: How Soaring Living Costs Affect People's Health and Wellbeing*, Senedd Research, 19 January 2023.

29 Narayan Ammachchi, *Water Scarcity Jeopardizes Data Center Projects in Chile and Uruguay – Nearshore Americas*, Nearshore Americas, 16 November 2023.

30 United Nations Environment Programme, *Artificial Intelligence (AI) End-to-End: The Environmental Impact of the Full AI Lifecycle Needs to Be Comprehensively Assessed – Issue Note*, 21 September 2024.

31 Haley Zaremba, *AI's Energy Appetite Sparks Global Power Grid Concerns*, OilPrice.com, 6 December 2024.

32 *AI and Sustainability Policy Position Paper*, The Institution of Engineering and Technology, October 2024.

33 *Assuring a Responsible Future for AI*, GOV.UK, 6 November 2024.

34 *Ensuring Trustworthy AI: The Emerging AI Assurance Market*, The Digital Regulation Cooperation Forum, 16 July 2024.

35 *Unlocking the Growth Potential of the UK's AI Assurance Market*, Frontier Economics, October 2024.

36 Alexandra Sasha Luccioni, Yacine Jernite, and Emma Strubell, *Power Hungry Processing: Watts Driving the Cost of AI Deployment?* arXiv, 28 November 2023.

37 Shengding Hu et al., *MiniCPM: Unveiling the Potential of Small Language Models with Scalable Training Strategies* arXiv, 3 June 2024.

38 *The Global AI Index*, Tortoise Media, accessed 13 December 2024.

39 *Understanding AI Uptake and Sentiment among People and Businesses in the UK*, Office for National Statistics, 16 June 2023.

40 *Technical Debt*, Techopedia, 13 June 2024.

41 Justus Bogner, Roberto Verdecchia, and Ilias Gerostathopoulos, *Characterizing Technical Debt and Antipatterns in AI-Based Systems: A Systematic Mapping Study*, 2021.

42 *Technical Debt Is Hindering AI Development*, PA Consulting, 23 October 2024.

43 Christopher Mims, *The Invisible $1.52 Trillion Problem: Clunky Old Software*, Wall Street Journal, 2 March 2024.

44 *AI Has an Environmental Problem. Here's What the World Can Do about That*, United Nations Environmental Programme, 21 September 2024.

45 Isabel O'Brien, *Data Center Emissions Probably 662% Higher than Big Tech Claims. Can It Keep up the Ruse?*, The Guardian, 15 September 2024.

46 *IFRS – ISSB Issues Inaugural Global Sustainability Disclosure Standards*, International Financial Reporting Standards, 26 June 2023.

47 *UK Sustainability Reporting Standards*, GOV.UK, 25 November 2024.

48 *IFRS – IFRS S1 General Requirements for Disclosure of Sustainability-Related Financial Information*, International Financial Reporting Standards, 2023. *IFRS – IFRS S2 Climate-Related Disclosures*, International Financial Reporting Standards, 2023.

49 *Framework for Developing UK Sustainability Reporting Standards*, GOV.UK, 16 May 2024.

50 *Open Consultation – UK Green Taxonomy*, GOV.UK, 6 February 2025.

51 *IFRS S2 – Industry-based Guidance on implementing Climate-related Disclosures*, International Financial Reporting Standards, 2023.

52 *Pilot Version | UK Net Zero Carbon Buildings Standard*, UK Net Zero Carbon Buildings Standard, September 2024.

53 *AI Opportunities Action Plan: Government Response*, GOV.UK, 13 January 2025.

54 *Supermicro's Third Annual Data Center & The Environment Survey*, Supermicro, March 2020.

55 *Who We Are*, Circular Electronics Partnership, accessed 8 January 2025.

56 *L.1071 : A Model for Digital Product Passport Information on Sustainability and Circularity*, International Telecommunication Union, 6 November 2024.

57 *AI Opportunities Action Plan: Government Response*, GOV.UK, 13 January 2025.

58 *ESG Funds Continue to Thrive and Outperform Traditional Funds across Equity and Fixed-Income Asset Classes*, The Institute for Energy Economics and Financial Analysis, 10 June 2024.

59 *Nations and Regions Investment Funds*, British Business Bank, accessed 16 December 2024.

60 Margaret Mitchell et al., *Model Cards for Model Reporting*, in Proceedings of the Conference on Fairness, Accountability, and Transparency, 14 January 2019.

61 *ISO 46001:2019 – Water Efficiency Management Systems — Requirements with Guidance for Use*, July 2019.

62 Microsoft, for instance, have already made such a commitment. [Steve Solomon, *Sustainable by Design: Next-Generation Datacenters Consume Zero Water for Cooling*, The Microsoft Cloud Blog, 9 December 2024.]

63 *BS 8001:2017*, May 2017.

64 AWS Data Centres, for instance, regularly reuse retired compute components. [*A Look inside the AWS Lab Where Retired Data Center Hardware Gets a Second Chance*, AWS, 6 June 2023.]

65 *How to Contribute*, Open Compute Project, accessed 16 January 2025.

66 *ISO/IEC 19770-1:2017 – Information technology — IT asset management*, December 2017.

67 *BS 8887-211:2012 – Design for manufacture, assembly, disassembly and end-of-life processing (MADE) – Specification for reworking and remarketing of computing hardware*, August 2012.

68 *Pilot Version | UK Net Zero Carbon Buildings Standard*, UK Net Zero Carbon Buildings Standard, September 2024.

69 *Warming Up to Efficiency: Understanding the Potential Benefits and Pitfalls of Data Centre Heat Export in the UK*, techUK, February 2024.

70 Katrin Sievert et al., *Expanding Renewable Electricity Use in Global Corporate Supply Chains*, Environmental Research: Energy 1, no. 3, 17 July 2024.

71 Offshoring describes geographically relocating a company's high greenhouse gas emitting activities to another country. [*Offshoring*, The Chancery Lane Project, 22 December 2022.]

72 *ISO/IEC 30134-3:2016 – Information Technology – Data Centres – Key Performance Indicators – Part 3: Renewable Energy Factor (REF)*, April 2016.

73 *Pilot Version | UK Net Zero Carbon Buildings Standard*, UK Net Zero Carbon Buildings Standard, September 2024.

74 *ISO 50001:2018 – Energy Management Systems*, August 2018.

75 Loïc Lannelongue, Jason Grealey, and Michael Inouye, *Green Algorithms: Quantifying the Carbon Footprint of Computation*, Advanced Science 8, no. 12, 2 May 2021.

76 *Data Centre Cost Index 2024*, Turner & Townsend, October 2024.

77 Yang Qiu et al., *The Impacts of Material Supply Availability on a Transitioning Electric Power Sector*, Cell Reports Sustainability 1, no. 10 (25 October 2024).

78 *Energy Efficiency Directive*, European Commission, October 2023.

79 For instance, Google has pledged that all energy matching will be done in a bundled fashion by 2030. [*24/7 Clean Energy – Data Centers – Google*, Google Data Centers, accessed 15 December 2024.]

80 Peter Benson, *ISO 8000: A New International Standard for Data Quality*, Data Quality Pro, accessed 15 December 2024.

81 *ISO/IEC 27555:2021 Information Security, Cybersecurity and Privacy Protection – Guidelines on Personally Identifiable Information Deletion*, October 2021.

82 *ISO/IEC 30134-5:2017 Information Technology – Data Centres – Key Performance Indicators – Part 5: IT Equipment Utilization for Servers (ITEUsv)*, November 2017.

83 *ISO 20400:2017 Sustainable Procurement – Guidance*, April 2017.

84 *ISO/IEC DTR 20226*, accessed 16 January 2025.

85 *FprCEN/CLC/TR 18145 – Environmentally Sustainable Artificial Intelligence*, accessed 16 January 2025.

86 Loïc Lannelongue, Juliette Fropier, and Even Matencio, *How to Include Environmental Sustainability Criteria in National AI Funding Schemes? Reflecting on the Example of France and the Green Algorithms Tool*, 7 January 2025.

87 *Guidelines for a Just Transition towards Environmentally Sustainable Economies and Societies for All*, International Labour Organization, 2 February 2016.

88 *Electricity 2024 – Analysis*, International Energy Agency, January 2024.

89 *International Total Electric Power*, U.S. Energy Information Administration (EIA), accessed 21 January 2025.

90 *World Energy Outlook 2024 – Analysis*, International Energy Agency, October 2024.

91 *International Total Electric Power*, U.S. Energy Information Administration (EIA), accessed 21 January 2025.

92 *Data Centers: More Gas Will Be Needed To Feed U.S. Growth*, S&P Global, 22 October 2024.

93 *UK Natural Gas: Consumption, 1965 – 2024 | CEIC Data*, CEIC, accessed 20 January 2025.

94 *Data Centre Power Use 'to Surge Six-Fold in 10 Years'*, BBC News, 26 March 2024.

95 Joseph McDonald et al., *Great Power, Great Responsibility: Recommendations for Reducing Energy for Training Language Models*, in Findings of the Association for Computational Linguistics: NAACL 2022, 19 May 2022.

96 *Clean Power 2030 Action Plan*, GOV.UK, 13 December 2024.

97 *2024 Environmental Sustainability Report*, Microsoft Sustainability, 2024.

98 *2024 Environmental Report*, Google Sustainability, July 2024.

99 *A Summary of England's Revised Draft Regional and Water Resources Management Plans*, GOV.UK, 20 December 2024.

100 *Seven Regions in England Will Face Severe Water Stress by 2030 as Brits Significantly Underestimate Their Daily Water Usage*, Kingfisher and Cebr2, May 2023.

101 *Tech Tourism: Inside a Data Centre*, The Chartered Institute for IT, 15 February 2024.

102 Kirsten James, *Semiconductor Manufacturing and Big Tech's Water Challenge*, World Economic Forum, 19 July 2024.

103 Justin Winter, *Quenching the Semiconductor Industry's Thirst*, Impax Asset Management, 23 January 2024.

104 Shaolei Ren, *How Much Water Does AI Consume? The Public Deserves to Know*, OECD.AI Policy Observatory, 30 November 2023.

105 *Critical Materials – Reducing Demand and Ensuring Sustainability*, National Engineering Policy Centre, October 2024.

106 John Werner, *TSMC To Double Production Based On Nvidia Numbers And Overall Demand*, Forbes, 7 November 2024.

107 Matt Ashare, *Nvidia CEO Promises to Deliver 'a Mound of Chips' as AI Development Speeds Ahead*, CIO Dive, 23 May 2024.

108 *Globality and Complexity of the Semiconductor Ecosystem – GSA – Global Semiconductor Alliance*, Global Semiconductor Alliance, March 2020.

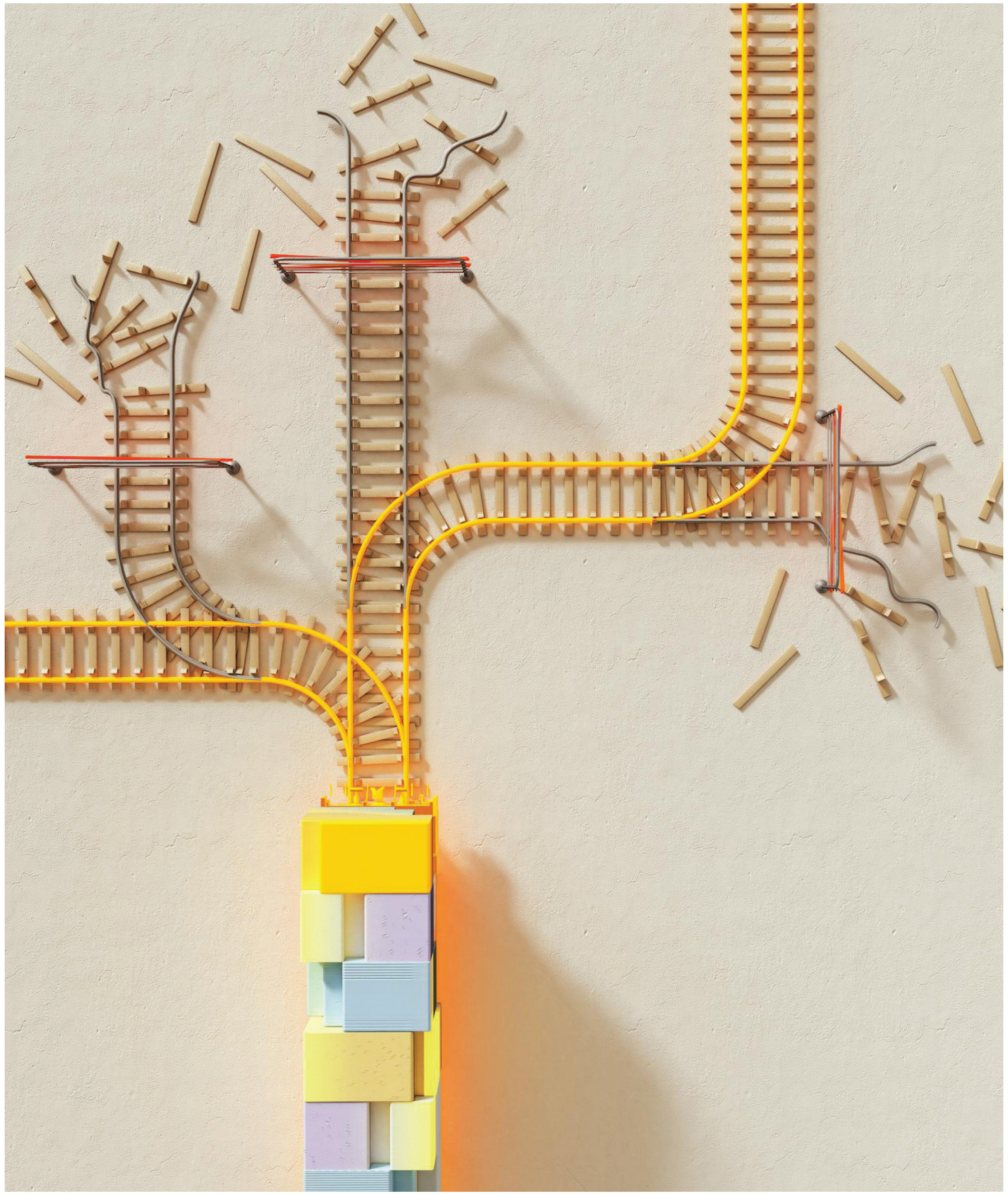
109 Debbie Woods and Devyani Gajjar, *Supply of Semiconductor Chips*, Parliamentary Office of Science and Technology, 13 December 2024.

110 Peter Dennis, *UK Generated 2nd Largest Amount of E-Waste as a Country in 2022*, Circular Online, 19 January 2023.

111 Ibid.

112 Peng Wang et al., *E-Waste Challenges of Generative Artificial Intelligence*, Nature Computational Science 4, no. 11, 28 October 2024.

113 *2024 Environmental Sustainability Report*, Microsoft Sustainability, 2024.



© Google DeepMind

Notes

National
Engineering
Policy Centre

Royal Academy
of Engineering

THE ROYAL ACADEMY OF ENGINEERING

The Royal Academy of Engineering is harnessing the power of engineering to build a sustainable society and an inclusive economy that works for everyone.

In collaboration with our Fellows and partners, we're growing talent and developing skills for the future, driving innovation and building global partnerships, and influencing policy and engaging the public.

Together we're working to tackle the greatest challenges of our age.

NATIONAL ENGINEERING POLICY CENTRE

We are a unified voice for 43 professional engineering organisations, representing 450,000 engineers, a partnership led by the Royal Academy of Engineering.

We give policymakers a single route to advice from across the engineering profession.

We inform and respond to policy issues of national importance, for the benefit of society.

Royal Academy of Engineering
Prince Philip House
3 Carlton House Terrace
London SW1Y 5DG

Tel 020 7766 0600 www.raeng.org.uk
@RAEngNews

Registered charity number 293074