WHITE PAPER

The Impact of Digital Decarbonisation in Records and Information Management

Professor Thomas Jackson
Professor Ian Hodgkinson

Want to learn more?

Contact us: **digitaldecarbonisation@oasisgroup.com**

1. Foreword

At the heart of our service is the belief that the information we manage on behalf of our clients is more than just data. What we hold has an intrinsic value for the organisations that we work with. We treat everything with the utmost care and professionalism. Whether it's digitising records, storing and retrieving physical archives or providing specialist environments to scan and digitise delicate historical records or geographically significant sub surface core samples, we keep that value at the centre of everything we do.

We never forget that every piece of information has the power to have a real impact on the success of our clients' organisations. We're helping them to safeguard sensitive and critical information so that they can access and leverage the insight and intelligence for enhanced business operations and customer service.

Better Information, Better Decisions, Better Outcomes

The authors

Tom Jackson (BSc, PhD) is a Professor of Information and Knowledge Management at Loughborough Business School, Loughborough University. He has over 20 years' experience of research and industrial consultancy and secured over £40million of research funding. Known for the creation of EMOTIVE.AI which holds the world's best f-measure for fine grained emotion detection. A system that has been able to predict in real-time the outcome of events like US and UK elections through to mental health conditions. His research also includes the pioneering work on the dark side of digital decarbonisation and the impact on the environment; the first publicly available data carbon forecasting toolkit, which serves as a valuable resource for individuals and organisations embarking on their net-zero journey; the first model of information overload, and in his earlier career he was nicknamed 'Dr Email' by the BBC for his work on the impact of interrupts in the workplace. He is a Member of the OECD.AI Expert Group for Compute and Climate; Chair of the Technical Working Group for Ethics in the Government National Digital Twin Programme; and an Independent Scientific Adviser – BridgeAI Programme – The Alan Turing Institute.

Ian Hodgkinson (BSc, MSc, PhD) is a Professor of Strategy at Loughborough Business School, Loughborough University (UK). Ian has published extensively on digital transformation and the role of digital innovations for service value co-creation across private and public sector contexts, with this work featuring in a range of leading academic journals (e.g. Public Administration, Public Management Review, Technovation, Technological Forecasting and Social Change, Australian Journal of Public Administration). With a focus on real-world impact, this research has involved: telemedicine and international healthcare policy, public spending and service optimisation, and digital solutions to wicked problems. Alongside this work, Ian has advanced new knowledge on the inter-relationships between organisational learning, data, and the environment, which has been instrumental in guiding the digital decarbonization movement. As a core member of the Technical Working Group for Ethics, Ian co-developed the foundational ethics strategy for the UK's National Digital Twin Programme (NDTP), and, on the international stage, he is feeding into global AI policy development as a member of the OECD.AI Compute and Climate Expert Group.

Its 4

2. Contents Page

3. Executive Summary

As digital adoption continues to accelerate, organisations are increasingly relying on vast amounts of digital data to drive innovation, efficiency, and customer engagement. However, this surge in digitalisation also comes with a significant environmental cost. The generation, storage, and management of digital data require substantial energy, leading to increased carbon emissions that extend well beyond the initial production of hardware.

Unlike physical goods, digital data are intangible, making it easy to overlook its environmental impact. For years, digital data was mistakenly considered carbon-neutral, largely because its carbon footprint was not well understood. However, recent advancements have highlighted how the inefficient use and the 'store it all' mentality can significantly contribute to carbon emissions. This realisation has sparked the digital decarbonisation movement, which focuses on reducing the environmental impact of digital activities.

Purpose of the whitepaper

As global efforts to combat climate change intensify, decarbonisation has become a critical agenda for organisations. While technological advancements have played a central role in these efforts, they have also led to an explosion of digital data, which in turn contributes significantly to greenhouse gas (GHG) emissions. The whitepaper explores strategies for optimising data management to minimise digital-related GHG emissions, ultimately leading to reduced digital expenses.

Approach

The whitepaper outlines a structured approach to minimising the environmental impact of digital data by focusing on the classification, verification, and ongoing maintenance of data. The insights and findings generated are based on a comprehensive review of the burgeoning academic and grey literatures on the broad topic of digital decarbonisation. The phenomenon of digital decarbonisation is still in its infancy. With the limited awareness of the digital carbon footprint to date across organisations and industries, there is a huge opportunity for organisations to be pioneers of digital decarbonisation.

Key Findings and Recommendations

The paper provides a series of practical steps that organisations can implement to manage data more effectively and contribute to digital decarbonisation efforts. The recommendations offered in this paper are designed to help businesses navigate the complex data and information landscape while advancing their sustainability goals.

The proposed recommendations can be broadly classified under three core areas: (i) understanding data more effectively, (ii) utilising metadata to catalogue, manage, and optimise data use, and (iii) implementing best practices for data optimisation; a non-exhaustive list of the key findings and recommendations from this whitepaper are highlighted below:

Digital decarbonisation

Companies should prioritise digital decarbonisation for (i) Significant Reduction in Carbon Footprint; (ii) Alignment with Global Sustainability Goals; (iii) Substantial Cost Savings; (iv) Enhanced Operational Efficiency; (v) Potential for Reinvestment and Innovation; and (vi) Future-Proofing Against Regulatory Changes.

Data classification

It is important to classify sets of data and include a time dimension, as the data may be present in several repositories over time. A simple classification can be followed to aid responsible information and records management, and is outlined here:

- *Data that is core to business and created by the business.* These data are critical to the business, should be very well managed, reviewed periodically, should never be dark, and always be identifiable using available metadata.
- *Data that are consumed from third party and replicated into business system so the copy creates more data.* This could be data regarding product definitions from suppliers, social media data used to enhance marketing promotions, market driven data from external agencies. These are business to business (B2B) interactions that create additional data that must be constantly maintained as data changes at source.
- *Data that are external to the business and consumed as and when needed but never moved into the business.* For example, the use of APIs to access open-source data to obtain some insights into externally available information to enhance the business operations. Care should always be taken with this source of data to ensure its veracity and longevity.
- *Data that are required for backups, archives, and failover, which can become dark over time*. Backups and archives are frequently taken with old rules and policies and data retained that could possibly be deleted. These data are very important but should be reviewed periodically to check on its requirements. Sometimes it may be better to rebuild some (or all) of these environments or at least move them to cheaper forms of storage.

Data entity audit framework

- By integrating the concepts of data completeness and timeliness, the whitepaper introduces the Data Entity Audit framework. This framework enables organisations to audit and review their major data entities and attributes across their information landscape, identifying areas where improvements can be made to reduce GHG emissions and reduce their digital costs.

Metadata

- Accurately tagging and classifying new digital data with metadata allows for the systematic identification of data relevance and utility, ensuring that only valuable data is retained, and that unnecessary or outdated data is efficiently archived or deleted. This process not only optimises storage solutions but also significantly reduces the environmental impact of data management by minimising energy use associated with data storage and processing.
- For historical digital data, it is possible to scan for timestamps with operating system level tools to identify when data was created, last modified, its file size and so on. Gathering log files related to the data can then allow organisations to build a larger inferred picture around the data usefulness, for example, it may be possible to build a 'heat map' of data most used, less used, rarely used and never used; helping to identify digital waste for disposal.
- Other forms of metadata are required to classify information completely. These can only be added after the data are originally identified and as such are useful to prevent data that was 'bright' becoming dark over time. Tagging dark data (once found) with additional, suitable metadata to understand and use downstream can prevent that data becoming dark again.

Digital best practices for adoption

- Digital Records Lifecycle Management - Effective records management begins with a robust digital records lifecycle management strategy.
- Regular Audits and Disposal of Unnecessary Records - Conducting regular audits of digital records is essential to maintaining an efficient records management system.
- Efficient Storage Solutions - Implementing efficient storage solutions is another critical aspect of effective records management.
- Data Minimisation Strategies - Data minimisation is the practice of collecting only the data necessary for a specific purpose and retaining it only as long as needed.
- Centralised Data Management Systems - Implementing centralised data management systems allows organisations to consolidate their data storage and management efforts into a single, cohesive system.
- Use of Cloud Services with Green Certifications - Utilising cloud services that have green certifications can significantly reduce the environmental impact of data storage and processing.
- Avoiding Data Redundancy - Implementing strong data governance measures to prevent unnecessary duplication of data, which can lead to increased storage and energy costs.
- Managing Unused Data - Regularly review and manage dark data, ensuring it is either effectively utilised or responsibly discarded.

- Enhancing Knowledge Management - Develop robust systems for knowledge management that promote the efficient sharing and reuse of information.
- Staff Awareness and Training - Educate employees on the importance of responsible data usage and the environmental impact of digital operations.
- Responsible AI Integration - When implementing Artificial Intelligence (AI), carefully consider dataset quantity and location, and assess the necessity of using advanced technologies like Generative AI to avoid unnecessary energy consumption.

The challenges that lie ahead

1. Inconsistent Data Measurement Standards: A major challenge lies in the lack of a global, uniform approach to measuring data.
2. Comprehensive Assessment of GHG Emissions from Data Storage: Another obstacle is the insufficient consideration of all sources of GHG emissions related to data storage and processing.
3. Data Sharing Across Supply Chains: Ensuring seamless data sharing across different parts of the supply chain presents a significant challenge.
4. Underutilisation of Unstructured Data: The role of unstructured data, such as shipping documents, certificates of origin, and other relevant records, is often overlooked.
5. Growing Problem of Dark Data in Supply Chains: As data storage needs continue to expand, particularly across supply chains, the accumulation of dark data poses a significant challenge to digital decarbonisation.

4. Introduction

Definition of Digital Decarbonisation

Digital Decarbonisation is a transformative approach focused on the responsible and efficient use of knowledge and data within organisations to support sustainability goals. At its core, it advocates for the integration of digital best practices into sustainability strategies to significantly reduce data-related carbon emissions.

Explanation of digital activities contributing to carbon emissions

The rapid expansion of global data volumes poses a significant challenge to achieving net-zero carbon targets by 2050. As the digital economy grows, so too does the demand for data processing, storage, and transmission, all of which are energy-intensive activities. For instance, the International Energy Agency (IEA) has projected that the electricity consumption of data centres worldwide will more than double by 2026, surpassing 1,000 terawatt hours—equivalent to Japan's annual electricity consumption (Hodgson, 2024).

Data centres, which are essential for cloud computing, big data analytics, and other digital services, require vast amounts of electricity to power servers, cooling systems, and other infrastructure. This energy demand is expected to escalate as data volumes continue to grow exponentially, driven by increasing internet usage, the proliferation of connected devices (IoT), and the expansion of AI and machine learning (ML) applications.

As digital infrastructure expands, so does its environmental footprint, creating a paradox where technological advancements that drive economic growth also exacerbate environmental challenges. Without substantial improvements in energy efficiency, the increased electricity consumption from digital activities could undermine global efforts to meet climate goals. This underscores the urgent need for sustainable data management practices, renewable energy adoption in the digital sector, and innovations in low-carbon technologies to mitigate the environmental impact of digital activities.

Urgency of addressing energy consumption in digital spaces

To contextualise the impact of digital activities, consider the average Instagram post by one of the world's most famous soccer players, Ronaldo. According to Channel 4 Dispatches, the power required to support the reach of such a post is equivalent to the energy needed to power 10 homes for an entire year. Scaling this example, with approximately 45,000 Instagram posts made every 60 seconds, if just two posts reach similar levels of engagement annually, the energy consumption would be enough to power 10.5 million homes—the equivalent of all the residences in Washington D.C., New York, and London.

So far, industrial decarbonisation efforts have predominantly focused on technological advancements to tackle the challenge. For example, the UK's Industrial Decarbonisation Strategy

(Department for Business, Energy and Industrial Strategy, 2021) highlights essential measures for transitioning "from fossil fuel combustion to low carbon alternatives such as hydrogen and electrification, deploying key technologies such as carbon capture, usage and storage, and supporting industrial sites to maximise their energy and resource efficiency" (p. 6). This underscores the crucial role of low-carbon technology and innovation in the process. Although technology and decarbonisation are often seen as closely linked—evidenced by a 3,750% rise in venture capital investment in Climate Tech from 2013 to 2019 (PWC, 2023)—there has been little consideration of the potential negative impact that digital practices themselves may have in this context.

Growing digital activity and its environmental impact

The impact of industrial data on GHG emissions and the cost to store, maintain, process, and exploit that data is therefore emerging as a critical concern for the signatories of the Paris Agreement and the commitment to net zero by 2050 (Jackson and Hodgkinson, 2024a). Within this context, the phenomenon of digital decarbonisation has been recently highlighted by Jackson and Hodgkinson (2023a), whose work is one of the earliest to explore how digital practices may in fact add to organisations' digital carbon footprints. The authors explore how, through the knowledge practices of organisations, an increasing volume of digital data is being generated, processed and stored, but then often forgotten or never reused, creating a huge and unnecessary demand on energy consumption.

To successfully manage data and contribute to digital decarbonisation initiatives through the reduction of data-related GHG emissions, there are three consideration that must be accomplished in the classification and verification of data use and the ongoing maintenance of classification of data:

(i) How best to understand data so it is meaningful and useful. This covers the number of attributes required in any dataset to be able to deal with the requirements the data is to be used for and the granularity of the data based on its frequency of being recorded.

(ii) How best to create additional metadata and use it to build a metadata repository to catalogue, use, manage and maintain data to prevent it becoming dark data (i.e. data which is being/has been captured but is either unknown to systems/users or embedded within data formats that make it/have made it difficult to extract useful information that a business can exploit).

(iii) How best to organise data and enable businesses, and 3rd parties to address the need to control data and to reduce GHG emissions from the storage and maintenance of this data.

In investigating these considerations, the whitepaper explains what forms data can take in organisations and presents a framework to help organisations assess the 'fitness' of data entities and which data can be removed.

5. The Digital Carbon Footprint

Understanding Digital Carbon Footprint

Organisations are generating, processing, and storing increasing amounts of digital data, much of which is often forgotten or never reused. This practice unnecessarily drives up energy consumption, a major contributor to global CO2 emissions. This issue is particularly relevant in the context of government policy. For example, in the UK, large businesses are required to report on greenhouse gas (GHG) emissions from direct business activities (scope 1), indirect activities (scope 2), and the entire value chain (scope 3), which is critical for achieving carbon net zero by 2050 (Carbon Reduction Policy paper, 2023). However, for net zero to be truly achieved, organisations must consider all significant contributors to GHG emissions across these scopes. Yet, the carbon footprint of digital data is notably absent from mandatory reporting, even though digitalisation already accounts for 4% of global GHG emissions (Teuful and Sprus, 2020). This is expected to rise as data generation grows, with predictions of 180 zettabytes globally by 2025 and over 2k zettabytes by 2035 (Statista, 2020).

Sources of energy consumption in digital activities (data centres, devices, networks)

The sources of energy consumption in digital activities span across three main areas: data centres, devices, and networks. Each of these areas contributes significantly to the generation of GHG emissions, impacting all three GHG Protocol scopes. These emissions are relevant not just for the data creator or owner but also for the data custodian, such as hyper-scale data centres. These centres, which store and process vast amounts of data, are particularly energy-intensive, contributing to substantial GHG emissions. ***Do you know the carbon footprint of where your data are stored and processed?***

Researchers have extensively tracked and verified these emissions throughout the supply chain. Studies by Acquaye et al. (2020) and Ciacci et al. (2020) have explored GHG emissions in data centres, shedding light on the environmental impact of digital data storage and processing. Additionally, there is a growing concern around "dark data"—data that is generated and used only ever once or indeed not at all (Gartner, 2022). This concept underscores the inefficiency within digital data management, as the energy consumed to create and store this data ultimately goes to waste if the data is never utilised again. ***Do you know how much dark data your organisation is storing, locally or in the cloud?***

Despite these insights, there has been limited research focusing on the integration of GHG emissions from digital activities across data centres, devices, and networks. Energy consumption in digital activities is a complex issue that impacts GHG emissions across various scopes and sectors. The integration of data from these different areas—data centres, devices, and networks—especially in industrial contexts, is crucial for developing strategies to reduce the carbon footprint associated with digital data.

Metrics and measurement of digital carbon emissions

Different stages of business operations and the supply chain produce a wide variety of data, ranging from small amounts of high-impact critical data (i.e., high data CO2) to vast quantities of smaller datasets that may rarely, if ever, be used (e.g., scanned documents, sensor data on minor processes). Despite their seemingly insignificant individual impact, these smaller datasets can collectively contribute to substantial data-related CO2. This can result in a "long tail" of diverse data types that unnecessarily generate GHG emissions through their creation, processing, and storage. Mapping the "data CO2 journey" is valuable for identifying redundant data while retaining only essential datasets.

To illustrate this, Figure 1 presents a data CO2 curve based on energy consumption (e.g., storage, processing) and the number of units (e.g., the physical environment for CPUs, memory, and processors in locations such as hyper-centres, local data centres, PCs, mobile devices, etc.). The curve is divided into four distinct zones, as described by Lockwood et al. (2023) and shown in Figure 1 below:

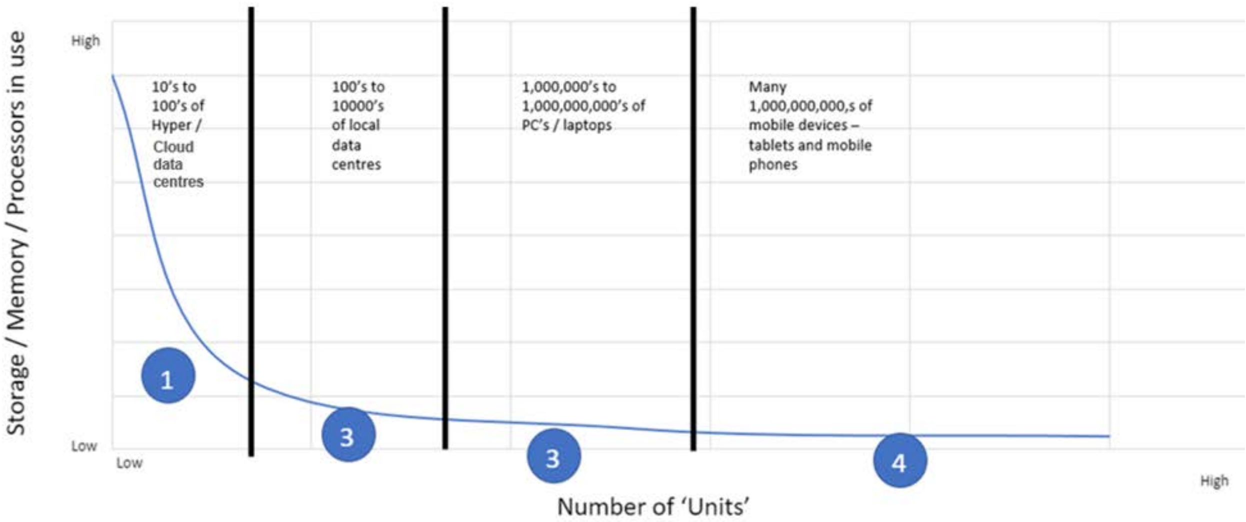

Figure 1. The data CO2 curve

- Zone 1 – Large Hyper/Cloud Data Centres: Companies like Amazon, Google, and Microsoft operate over half of the world's largest data centres. These facilities consume vast amounts of centralised energy, though some of this energy is sourced from green energy providers [Scopes 1 and 2].
- Zone 2 – Local Data Centres: Managed by individual companies rather than leased, these centres are significant because some businesses prefer to keep substantial data behind their own firewalls for security or cost reasons [Scope 2].
- Zone 3 – Desktops, PCs, Laptops: Individuals access data from Zones 1 and/or 2 or generate their own data, often replicating information from these zones. The GHG emissions here are

influenced by local factors, such as how energy is supplied to businesses and homes [Scope 3].

- Zone 4 – Tablets, Wearables, and Phones: While the energy consumption of each device is minimal, the sheer number of these devices globally leads to a considerable cumulative impact. As in Zone 3, GHG emissions are driven by local energy policies [Scope 3].

Current Trends and Statistics

The energy supply sector is recognised by the United Nations (2022) as the largest contributor of GHG emissions, accounting for 35% of total global emissions. Digital data represents a significant drain on the energy supply. “The cloud”, for instance, is made up of many data centres throughout the world and contains millions of servers that both process and store these data and manage IT systems. Power consumption of a single server running at 100% for a year is 319740-watt hours, according to Medium.com, and it is reported that data centres are responsible for 2.5% - 3.7% of all human-induced carbon dioxide and have a greater carbon footprint than the aviation industry (2.1%) (Climatiq, 2022). With data centres now among the highest consumers of electric power, it raises questions as to whether all of this data is actually needed, especially when one considers that up to 65% may comprise data never to be used again.

Environmental impact assessments

Based on available measures, we predict that an average small enterprise with 50 full-time employees will generate 2,295gb of data a day, this equates to 550,800gb of data generated a year and this will generate 1,102 tons of C02 per year. The carbon cost of data calculator available at digitaldecarb.org predicts the carbon footprint of data per year is the same carbon footprint as would be generated if you flew from London Heathrow to New York 1,281 times a year.

Summary

The chapter on "Understanding Digital Carbon Footprint" discusses the significant environmental impact of increasing digital data generation, much of which is unused ("dark data"). In their extensive examination of dark data, David J. Hand (2020) notes that while it is possible that unexamined collections of data might contain something valuable, there is also every chance they may not. Consequently, this data contributes to unnecessary energy consumption and GHG emissions, yet its carbon footprint is often overlooked in mandatory reporting.

Key sources of energy consumption include data centres, devices, and networks. Data centres, especially large-scale ones, are major contributors to GHG emissions. The chapter introduces the "data CO2 curve," highlighting how different data storage and processing zones contribute to emissions.

With digitalisation accounting for 4% of global GHG emissions, the chapter stresses the need for better integration of digital data emissions into carbon reduction strategies. It highlights the significant carbon footprint of even small enterprises' data usage.

6. Harnessing Dark Data for Sustainable Digital Transformation

Adopting sustainable digital practices

Dark data represents a largely untapped reservoir of information that organisations can leverage to reduce carbon emissions and enhance sustainability efforts. By identifying and analysing dark data, businesses can uncover inefficiencies in their operations, helping to minimise their carbon footprint and move closer to achieving net-zero emissions.

In today's digital age, businesses generate massive amounts of data, much of which goes unused, becoming "dark data." The scale of this issue is staggering; data centres alone consume more energy than the entire United Kingdom. This unused data contributes significantly to energy consumption and, consequently, to the overall carbon footprint. Effective management of dark data requires not only classifying and verifying its business value but also maintaining this classification to prevent data from becoming obsolete or "dark" again.

Successfully managing dark data begins with establishing a clear definition and framework for identifying it. While the concept of dark data is widely recognised, its definition varies across different sources. Some common interpretations include:

- New Data Capture: Data that is not currently being collected but would be valuable if captured. For instance, enhancing supply chain efficiency by installing additional sensors to monitor critical factors such as temperature during transportation or tracking the miles driven to deliver products. This involves capturing entirely new data.
- Data Misinterpretation: Data that skews analytical results, often due to being overlooked or improperly accounted for in data science models. These smaller, unnoticed data sets can lead to incorrect conclusions, such as confusing correlation with causation.
- Hidden Data: Information embedded within existing datasets that requires extraction or manipulation to become useful. Examples include documents, emails, videos, or audio files that need further processing to reveal their value.
- Lost Data: Data that is stored but effectively lost within data centres or devices due to poor labelling or organisation. This includes backup or archive data that hasn't been properly maintained, as well as log files that could provide valuable insights into data usage patterns but are difficult to locate.

Gartner (2019) defines dark data as "the information assets organisations collect, process, and store during regular business activities but generally fail to use for other purposes." Expanding on this, dark data can be described as "data that has been or is being captured but remains unknown to systems or users or is embedded in formats that make it difficult to extract and utilise." This definition focuses specifically on data that is already stored but has not been fully exploited.

Identifying dark data is crucial, as it can be either valuable or worthless (Hand, 2020). Only by recognising it and assessing its worth in context can organisations decide whether to utilise or discard it. A recent policy briefing (Jackson and Hodgkinson, 2024b) categorises dark data into four types:

1. Traditional Structured Data: This includes data that is manually input or generated into one system and subsequently used by other systems. Even if no longer actively used, this data might still hold value or be necessary for legal and audit purposes. It may contain Intellectual Property (IP), Personally Identifiable Information (PII), company data, and more.

2. IoT Data Streams: Data captured from IoT devices often lacks sufficient context, making it difficult to interpret later. Without proper tagging, these streams of data tuples can become virtually unusable.

3. Unstructured Data: Forms of unstructured data like videos, sound files, emails, web pages, or documents may hold valuable insights, but these insights can only be extracted after converting the data into a structured or semi-structured format.

4. System-Generated Data: This includes log files and other data created by systems that are rarely, if ever, used. For instance, BI/AI output data may be stored after a single use and never accessed again.

Dark data can be found at various levels within an organisation, including core data created by the business, data received from third parties, data consumed during specific processes and discarded, and data retained for continuity, security, or legal reasons. While there are open-source tools like DeepDive, Snorkel, and Dark Vision that assist in identifying dark data, the real challenge lies in continuously classifying and verifying its value to prevent it from becoming dark once again. This ongoing process is essential to ensuring that businesses not only reduce their carbon footprint but also operate more efficiently and sustainably.

Merits of Physical and digital support in Decarbonisation

The integration of physical and digital strategies plays a critical role in supporting decarbonisation and sustainability efforts, particularly in the context of managing dark data. Here are the key merits of combining physical and digital approaches to achieve these goals:

1. Enhanced Data Utilisation for Efficiency

- Digital: Digital tools and analytics can help identify, categorise, and extract value from dark data, turning it into actionable insights. This leads to more informed decision-making, optimising processes, and reducing waste.

- Physical: Implementing physical changes based on insights derived from dark data, such as optimising supply chains, reducing energy consumption in data centres, or improving resource allocation, can directly contribute to reducing carbon emissions.

2. Reduction in Energy Consumption

- Digital: By effectively managing dark data, businesses can decrease the energy required to store and maintain unnecessary or obsolete data. Digital tools can streamline data processes, reducing the energy footprint of data centres.
- Physical: Physical interventions, such as improving data centre cooling systems or transitioning to more energy-efficient hardware, can further minimise energy consumption. This combination leads to a significant reduction in the overall carbon footprint.

3. Sustainable Data Management Practices

- Digital: Continuous monitoring and classification of data through digital systems prevent data from becoming "dark" again. This ensures that all captured data remains useful and contributes to the organisation’s sustainability efforts.
- Physical: By maintaining sustainable practices in data storage and processing, such as using renewable energy sources to power data centres or implementing green IT strategies, organisations can further support their sustainability objectives.

4. Improved Operational Efficiency

- Digital: Digital tools can optimise operational workflows by analysing dark data to identify inefficiencies, such as redundant processes or underutilised resources. This can lead to more streamlined operations and lower energy consumption.
- Physical: Physically implementing these optimisations, such as reconfiguring supply chains, upgrading infrastructure, or reducing unnecessary physical assets, can further enhance efficiency and reduce the carbon footprint.

5. Innovation and Continuous Improvement

- Digital: Leveraging digital technologies, such as AI and machine learning, can help organisations continuously improve their data management strategies, leading to ongoing reductions in energy use and carbon emissions.
- Physical: Physical innovations, such as adopting energy-efficient technologies or redesigning workspaces to be more eco-friendly, complement digital advancements and contribute to long-term sustainability.

6. Regulatory Compliance and Risk Management

- Digital: Digital systems help ensure compliance with environmental regulations by providing accurate data tracking and reporting. This can help avoid penalties and improve the organisation's reputation for sustainability.
- Physical: Physical measures, such as waste reduction and proper disposal of obsolete equipment, support compliance with environmental standards and contribute to reducing the organisation's environmental impact.

7. Holistic Sustainability Strategy

- Digital: A digital-first approach to managing dark data ensures that all aspects of data lifecycle management are optimised for sustainability. This includes data collection, storage, processing, and disposal.
- Physical: Physical strategies, such as reducing unnecessary physical infrastructure or adopting circular economy principles, align with digital efforts to create a comprehensive sustainability strategy that addresses both digital and physical aspects of the organisation.

By combining physical and digital approaches, organisations can create a synergistic effect that maximises the benefits of each, leading to significant progress in decarbonisation and overall sustainability.

Summary

Adopting sustainable digital practices offers significant benefits for decarbonisation and sustainability efforts. Managing dark data, unused and often energy-consuming data, can help organisations reduce their carbon footprint by uncovering inefficiencies and optimising operations. Combining digital tools with physical interventions, such as improving data centre efficiency and utilising renewable energy, enhances energy savings and operational efficiency. This approach not only supports regulatory compliance but also fosters continuous innovation and a holistic sustainability strategy, ultimately contributing to significant progress toward net-zero emissions and overall sustainability goals.

7. Methods to Reduce Digital Carbon Footprint

What is Records Management?

Records management is the systematic control of an organisation's records throughout their lifecycle, from creation and receipt to maintenance, use, and disposal. It involves the classification, storage, retrieval, and archiving of records to ensure they are properly managed and accessible when needed. Effective records management ensures that essential records are preserved, while unnecessary or outdated records are appropriately disposed of.

Importance in Reducing Data Clutter and Energy Use

Records management plays a vital role in minimising data clutter and reducing energy consumption. As organisations accumulate vast amounts of digital records, the lack of proper management can lead to a buildup of redundant, obsolete, or trivial records, often referred to as "ROT data." This excess data consumes valuable storage space and increases energy use, particularly in data centres. By implementing effective records management practices, organisations can streamline their record-keeping processes, reduce unnecessary storage demands, and lower their overall energy consumption, contributing to more sustainable operations.

Best Practices (Lifecycle, Audits, Storage)

- ***Digital Records Lifecycle Management*** - Effective records management begins with a robust digital records lifecycle management strategy. This involves defining the stages of a record's life—from creation, active use, and maintenance, to eventual archiving and especially disposal. By clearly outlining these stages, organisations can ensure that records are properly managed and stored according to their relevance and value. This approach helps in preventing the unnecessary accumulation of records and ensures that only important, active records occupy valuable storage space.
- ***Regular Audits and Disposal of Unnecessary Records*** - Conducting regular audits of digital records is essential to maintaining an efficient records management system. These audits help identify records that are no longer needed, outdated, or irrelevant. Once identified, these records should be systematically disposed of or archived according to established policies. This practice not only reduces data clutter but also helps in freeing up storage resources and reducing energy consumption associated with maintaining large volumes of unnecessary records.
- ***Efficient Storage Solutions*** - Implementing efficient storage solutions is another critical aspect of effective records management. This includes using centralised storage systems that allow for easy retrieval and management of records, as well as employing tiered storage strategies that allocate records to the most appropriate storage media based on

their usage frequency and importance. For example, frequently accessed records can be stored on high-performance storage systems, while less critical records can be moved to lower-cost, energy-efficient storage. Additionally, leveraging cloud storage services with green certifications can further enhance the sustainability of records management practices, ensuring that storage solutions are both effective and environmentally responsible.

What is Information Management?

Information management refers to the systematic process of collecting, storing, organising, and maintaining data and information in a way that ensures it is accessible, reliable, and useful for decision-making. It involves the governance, use, and protection of information assets within an organisation. Effective information management ensures that data is accurately captured, stored in a structured manner, and can be retrieved and used efficiently when needed.

Conversely, ineffective information management results in wasted time spent looking for information as well as a subsequent duplication of data and information. For instance, the IDC report that information workers in western Europe are losing 50% of their time every week searching for, governing, and preparing data (30%) and duplicating work (20%), building knowledge and information assets that already exist in the organisation (Jackson and Hodgkinson, 2023a). Not only is this an unnecessary drain on organisations productivity, but the consequence for the environment is also significant as these ineffective practices draw on energy resources to power knowledge that is not reused (Jackson and Hodgkinson, 2023a).

Developing policies that prioritise sustainable information use is a key best practice in information governance. These policies should include guidelines for data minimisation, ensuring that only necessary data are collected and retained. Policies should also address the classification and tagging of data with metadata to prevent the accumulation of dark data and promote the efficient use of storage resources. By setting clear standards for how data are managed, organisations can reduce their environmental footprint and improve data quality and accessibility.

Compliance with environmental standards is crucial for sustainable information governance. Organisations should adopt and adhere to standards that promote the responsible use of energy and resources in their data management practices. This includes using cloud services with green certifications, which ensure that data centres operate efficiently and use renewable energy sources. Additionally, organisations should regularly review and update their information governance practices staying in line with the latest environmental regulations and industry standards.

Employee training and awareness programs are essential for the successful implementation of sustainable information governance practices. These programs should educate employees on the importance of metadata management, the risks of dark data, and the organisation's environmental goals. Training should also cover best practices for data classification, storage, and disposal, ensuring that employees understand their role in maintaining data quality and reducing the organisation's carbon footprint. By fostering a culture of sustainability and responsibility, organisations can ensure that their information governance practices are consistently applied and effective.

Information management for Optimising Data Use and Storage

A robust information governance framework is essential for promoting sustainable information practices. A framework should involve the implementation of policies and procedures that guide how data are collected, stored, used, and disposed of, with a focus on minimising environmental impact. Sustainable information governance includes the use of metadata to ensure that data are correctly classified and managed, helping to prevent the creation of dark data. By tagging data with appropriate metadata, organisations can more effectively manage their information, reducing unnecessary storage and energy use. It also involves regular auditing of data entities to confirm their relevance and utility, thus minimising the accumulation of "dark data"—unused or unnecessary data that consumes storage resources. By categorising data (traditional, sensor, unstructured, system-created), organisations can apply specific strategies to manage each type effectively, ensuring that data are timely, complete, and stored at the appropriate level of granularity. This prevents the wasteful storage of irrelevant or redundant data, reduces storage costs, and enhances the overall efficiency of data processing systems.

Best Practices (collecting, centeralisation, Green cloud)

- ***Data Minimisation Strategies*** - Data minimisation is the practice of collecting only the data necessary for a specific purpose and retaining it only as long as needed. This approach reduces the risk of accumulating dark data and minimises storage costs. For example, in the context of sensor data, it is crucial to ensure that only relevant data are captured at the correct level of granularity. Collecting excessive data, such as recording sensor data too frequently when no significant changes occur, leads to unnecessary storage and processing, ultimately contributing to inefficiencies.
- ***Centralised Data Management Systems*** - Implementing centralised data management systems allows organisations to consolidate their data storage and management efforts into a single, cohesive system. This centralisation ensures that data are consistently categorised, audited, and maintained across the organisation, reducing the risk of data silos and redundant storage. Centralised systems also facilitate better access control, data sharing, and compliance with regulations, leading to more efficient and effective use of data assets.
- ***Use of Cloud Services with Green Certifications*** - Utilising cloud services that have green certifications can significantly reduce the environmental impact of data storage and processing. Green-certified cloud providers use energy-efficient data centres powered by renewable energy sources, which helps in lowering the overall carbon footprint associated with data management. Additionally, cloud services often offer advanced data management features, such as automated data archiving and deletion, which further

contribute to data minimisation efforts. By choosing cloud providers with a commitment to sustainability, organisations can ensure that their data management practices align with broader environmental goals and reduce their financial costs of storing data.

Knowing your data

A simplified view of data captured into any system can be disaggregated and categorised into four forms:

- ***Traditional data***: Data that are created within relational database management system (RDMB) often by manual means. This has been the core data capture method up to the 2000's and a great deal of historical data exists in this form. It is also often key data that the business uses to operate. It needs to be audited and managed to ensure its worth is still valid.
- ***Sensor data:*** Driven by the rapid growth in IoT solutions, sensor data is rapidly growing in volume. Care should be taken to ensure data captured is sufficient and necessary as often sensors are becoming embedded in devices and simply switched on by default. Care needs to be taken to ensure data are necessary to the business and are captured at the correct level of granularity.
- ***Unstructured data:*** Specific tools to extract meaningful structured data from unstructured sources needs to be undertaken to extract meaningful content. Artificial Intelligence (AI) in the form of Natural Language Programming (NLP) is often used to drive out data elements such as keywords, entities, concepts, categories, classification, sentiment, and emotion in document, email, text, Json type files. Other tools are used to extract information from unstructured data that is in the form of video, image, sound.
- ***System created data:*** This covers data created by operational and analytical systems that augment existing data. One area that has the potential to dramatically increase storage of data and potential dark data is around Large Language Models (LLM) which often depend on a huge corpus of data to work. In addition, backups / archives / failover sites are classed here. These data stores could be much bigger than the primary systems used by the business to operate.

The Role of Metadata in Sustainable Information Governance

Metadata plays a crucial role in managing dark data and supporting digital decarbonisation efforts, serving as a foundational element in the efficient and sustainable handling of information. As data about data, metadata provides detailed insights into the content, context, and structure of information, enabling organisations to make informed decisions about how data should be managed throughout its lifecycle. This includes determining whether data should be retained, archived, or discarded, and ensuring that it is stored in the most efficient and sustainable manner possible.

In the broader context of digital decarbonisation, metadata management is a key strategy for reducing the carbon footprint of data centres and IT operations. By facilitating better data governance, metadata helps organisations align their data practices with sustainability goals, ensuring that information is not just managed effectively, but also in an environmentally responsible manner. As organisations increasingly recognise the importance of sustainable data practices, investing in robust metadata management systems and tools becomes essential for both operational efficiency and environmental stewardship.

Daniel and Daniel (2012) describe many forms of metadata that can be useful, as briefly described below:

- ***Business metadata*** must be defined by the business first and then tagged to data items. It is the effort to tag/classify this data that is costly and prevents many businesses from taking on this labour. Tools are available to help with this manual effort. Recently AI approaches have been developed that can help this by scanning the data and using supervised training models to help define what the metadata elements could be.
- ***Governance actions and processes*** refers to the steps a business takes when it manages its data so this may be metadata that defines security policy or audit related information, an example might be around PII (Personally Identifiable Information) where the metadata defines who can see these data, who can see and use an obfuscated or modified version of it, allowing testing / developing but no view of the actual data or finally who simply cannot access that data in any form.
- ***Classification schemes / hierarchies*** allow data to be seen as part of a wider corpus of information, and example could be the individual parts of a car, through to sub-assemblies to the actual finished product. This is useful in accessing if information not considered could be useful when considered in a wider context.
- ***Feedback and Grouping metadata*** could be useful if users have used this form of metadata to feedback on how useful the data was to them, which could be a measure of its worth to the business.

Ideally, all ***technical metadata*** would be automatically generated and captured when data are ingested into any system, offering a comprehensive description of the data's characteristics. However, in practice, this is often not the case. Technical metadata, such as physical schemas, are frequently incomplete or outdated, which can lead to significant challenges in managing data effectively. Without accurate and up-to-date metadata, organisations struggle to maintain data quality metrics, making it difficult to assess the true value of their data. This lack of clarity can result in the unnecessary retention of data, contributing to data clutter, increased storage costs, and higher energy consumption, particularly as data centres work harder to manage and store vast amounts of information.

Furthermore, the absence of a comprehensive ***data dictionary***, a centralised repository that defines all data elements, compounds these issues, making it harder for organisations to maintain a clear understanding of their data assets. This can lead to the accumulation of dark data, or data that are stored but not actively used, which not only wastes storage resources but also hinders an organisation's ability to extract valuable insights from its data.

Once such metadata are in place, it must then be ***maintained*** utilising human or automated resources. This is increasingly being recognised by vendors through tooling solutions to help businesses identify and then maintain this metadata. It may be that any Chief Information Officer (CIO) should consider investing in teams and/or tools to ensure such rigour takes place as this will assist in several areas: Enabling accurate operational and analytical insights; Only maintaining essential data; Moving data from high-cost storage mediums to lower cost options; and, deleting unnecessary data. The ultimate goal here is to build a comprehensive data catalogue of all data assets, and ensure they are kept up to date.

Data ENTITY Audit Framework

Across different data forms, data entities must be assessed and evaluated for appropriateness. We present the data entity audit framework, a tool to ensure easy access, understanding and use of data entities to ensure data are fit for now and for the future. There are two key aspects to the framework when considering each data entity and its attributes required by the business:

1. Firstly, are **data *complete***, are the data necessary to satisfy all the needs the business has of it. This can be considered at entity level in a data enterprise architecture. This does not only concern itself with any attributes within an entity that is absent, but potentially attributes that are stored or planned to be stored that have no business use. This could become a source for dark data later.

Examples of data being incomplete is common, often seen in customer related data where key attributes for that customer are missing, incomplete or incorrect. This can result in analytics delivering poor offers to customers in the retail sector, incorrect diagnosis of a patient in the medical sector or wrong / poor products and services being offered in the banking and insurance sector. Identifying if data are complete is supported by its metadata, this is key to the approach.

2. Secondly, are these **data *timely***, is it captured or planned to be captured at the correct level of granularity. This can be 'too little' or 'too much': data could be recorded too infrequently such that valuable insights are missing in the time between records being changed; the opposite is true, and data could be captured far too frequently when it is in fact not changing and offers no additional value.

Examples of data being too coarse could be seen in medical records where data are only captured every time a patient visits, however in the meantime several issues may have arisen with the patient that is not seen and consequently managed by any medical practitioner. Data being recorded at too fine a level can often be seen in IoT (Internet of Things) scenarios, where sensors are added to machinery and streams of data are dumped to storage, *recording temperature, frequency, or pressure where no changes are occurring. If responses are not needed within seconds, then a longer time between record capture, or to only capture data when a change is detected would reduce useless data that is destined to become dark over time.*

Integrating the above two dimensions, ***data completeness*** and ***data timeliness,*** the **Data Entity Audit Framework** in Table 1 below shows how this approach could be embedded within businesses to audit and review major data entities across the information landscape:

Data Entity Audit Framework				
Is the data *Complete*?	YES	NO	NO	YES
Is the data *Timely*?	NO	YES	NO	YES
Description:	*Data are complete for use, but the data are recorded at the wrong level of time-based granularity.*	*Data are incomplete for use but data are being stored at correct time-based granularity.*	*Data are incomplete for use and are not sufficient regarding time-based granularity.*	*Data are complete for use and are recorded at correct level of time-based granularity, and they have metadata.*
ACTION:	**Resolve data timeliness issue(s)**	**Complete the dataset *or* discard data.**	**Rebuild the dataset *or* discard data.**	**Maintain data and evaluate over time.**

Table 1 – Data Entity Audit Framework

Summary

This chapter focused on methods to reduce the digital carbon footprint through effective records and information management. It emphasises the importance of systematic records management, which involves controlling the lifecycle of records from creation to disposal. Proper management practices, such as regular audits and efficient storage solutions, help minimise data clutter, reduce energy consumption, and optimise storage use. This not only preserves essential records but also ensures that unnecessary or outdated records are disposed of, thereby contributing to more sustainable operations.

The chapter also highlights the role of robust information management in promoting sustainability. It advocates for data minimisation strategies, centralisation of data management systems, and the use of green-certified cloud services. By implementing these best practices, organisations can reduce the accumulation of redundant or "dark" data, which consumes unnecessary storage, energy and costs. Additionally, the effective use of metadata is emphasised as a key strategy for sustainable data governance, helping organisations classify and manage data efficiently, thereby aligning their data practices with broader environmental goals.

8. Benefits of Digital Decarbonisation

In the global push to combat climate change, digital decarbonisation has emerged as a crucial strategy with far-reaching environmental and economic benefits. As organisations strive to reduce their carbon footprints, digital decarbonisation offers a path not only to sustainability but also to enhanced operational efficiency and cost savings.

Environmental Benefits

The integration of digital decarbonisation into broader organisational strategies aligns with global sustainability goals, such as those outlined in the Paris Agreement and the United Nations Sustainable Development Goals (SDGs). By reducing the environmental impact of digital infrastructure, organisations contribute to the achievement of these goals, particularly those related to climate action, responsible consumption and production, and sustainable industry innovation. The systematic reduction of digital carbon emissions not only aids in meeting regulatory requirements but also positions organisations as leaders in the global transition towards a low-carbon economy.

Economic Benefits of Digital Decarbonisation

One of the most immediate economic benefits of digital decarbonisation is the potential for significant cost savings through reduced energy consumption from on-premises offerings and a reduction in the use of data centre services. Digital decarbonisation also brings about efficiency gains by streamlining data management processes. When organisations systematically identify and eliminate redundant or obsolete data, they reduce the complexity of their digital systems. This streamlining not only cuts down on the number of required digital services used but also enhances the overall efficiency of IT operations. The resulting efficiency gains can lead to faster data processing, improved decision-making, and better resource allocation. Additionally, the cost savings achieved through digital decarbonisation can be reinvested into further sustainability initiatives or used to drive innovation and growth within the organisation.

So why invest in Digital Decarbonisation?

Investing in digital decarbonisation is not just a response to growing environmental concerns, it's a strategic business move that can deliver substantial benefits across multiple fronts. Here is why companies should prioritise this investment:

1. **Significant Reduction in Carbon Footprint** - Digital decarbonisation directly addresses the environmental impact of your company’s digital operations. By optimising data storage, processing, and management, companies can significantly reduce the energy consumption of their IT infrastructure, particularly in on-premises environments and reduce their data centres costs. This reduction in energy use leads to a decrease in greenhouse

gas (GHG) emissions, helping companies meet their sustainability targets and contribute to global climate action efforts.

2. **Alignment with Global Sustainability Goals** - As the world increasingly focuses on achieving sustainability milestones, aligning your company's strategies with global initiatives like the Paris Agreement and the United Nations SDGs is essential. Digital decarbonisation helps your company meet these goals by lowering the environmental impact of its digital infrastructure. This not only ensures compliance with emerging regulations but also positions your brand as a leader in sustainable business practices, enhancing your reputation among customers, investors, and partners.
3. **Substantial Cost Savings** - One of the most immediate and tangible benefits of digital decarbonisation is the potential for significant cost savings. By reducing energy consumption through more efficient data management, companies can lower their operational costs. These savings can be substantial, particularly in industries with large-scale IT operations. Reducing reliance on data centre services and optimising on-premises digital infrastructure can translate into lower energy bills and ***reduced costs associated with data storage and processing***.
4. **Enhanced Operational Efficiency** - Streamlining data management processes through digital decarbonisation leads to greater operational efficiency. By identifying and eliminating redundant or obsolete data, companies can reduce the complexity of their digital systems, resulting in faster data processing and improved decision-making. This efficiency gain not only boosts the overall productivity of your IT operations but also frees up resources that can be redirected to more strategic initiatives.
5. **Potential for Reinvestment and Innovation** - The cost savings achieved through digital decarbonisation provide opportunities for reinvestment. Companies can allocate these savings towards further sustainability initiatives, research and development, or innovation projects. This reinvestment can drive long-term growth, enhance competitiveness, and support your company's journey toward becoming a more resilient and future-ready organisation.
6. **Future-Proofing Against Regulatory Changes** - As governments and regulatory bodies increasingly focus on environmental sustainability, the ability to demonstrate reduced digital carbon emissions will become a critical factor in compliance. Investing in digital decarbonisation now enables your company to stay ahead of regulatory changes, avoid potential penalties, and ensure that your operations remain in line with evolving environmental standards.

Summary

Investing in digital decarbonisation is a smart business decision that offers a powerful combination of environmental stewardship and economic advantages. By reducing your company's carbon footprint, aligning with global sustainability goals, and unlocking significant cost savings and efficiency gains, digital decarbonisation positions your organisation as a leader in the transition to a low-carbon economy. As the world moves towards a more sustainable future, companies that embrace digital decarbonisation will not only contribute to the fight against climate change but also secure a competitive edge in an increasingly eco-conscious market.

9. Challenges and Considerations

In the quest for sustainability, digital decarbonisation has emerged as a critical priority for companies seeking to reduce their environmental impact. As digital activities increasingly contribute to carbon emissions, organisations must adopt strategies that not only enhance efficiency but also minimise their digital carbon footprint. This involves a holistic approach that includes data management, energy-efficient storage solutions, and responsible data retention practices. However, achieving digital decarbonisation is not without its challenges, such as managing the growing volume of data, balancing operational needs with environmental goals, and overcoming barriers to implementing sustainable practices. By addressing these challenges through strategic planning and innovative solutions, companies can make significant strides toward reducing their carbon emissions and contributing to a more sustainable digital future.

Numerous challenges and obstacles hinder the effective use and management of data by businesses, individuals, and society, all of which must be addressed to adopt a sustainable and responsible approach:

1. **Inconsistent Data Measurement Standards**: A major challenge lies in the lack of a global, uniform approach to measuring data. While some countries and regions have established various standards, these are not universally applied, leading to inconsistencies in how data is managed and understood across borders.
2. **Comprehensive Assessment of GHG Emissions from Data Storage**: Another obstacle is the insufficient consideration of all sources of greenhouse gas (GHG) emissions related to data storage and processing. This includes not only Cloud data centres but also on-premises data centres (which are prevalent in hybrid and multi-cloud environments) and even local devices where replicated data may contribute to increased emissions.
3. **Data Sharing Across Supply Chains**: Ensuring seamless data sharing across different parts of the supply chain presents a significant challenge. This requires the development of data standards and metadata that can maintain the coherent context of shared data, determine what can be shared while maintaining security, verify the accuracy of the data, and establish a common logical data model to facilitate collaboration.
4. **Underutilisation of Unstructured Data**: The role of unstructured data, such as shipping documents, certificates of origin, and other relevant records, is often overlooked. Extracting valuable information from these documents is challenging but essential, with approaches like Natural Language Processing (NLP) offering potential solutions to unlock the data's value for analysis.
5. **Growing Problem of Dark Data in Supply Chains**: As data storage needs continue to expand, particularly across supply chains, the accumulation of dark data poses a significant challenge. This unused and often overlooked data can become a substantial drain on storage resources, exacerbating inefficiencies and increasing the environmental impact.

Addressing these challenges requires a concerted effort to standardise practices, enhance data sharing, better utilise unstructured data, and manage the growing issue of dark data. Only by overcoming these obstacles can businesses and society move toward a more sustainable and responsible approach to data management.

Mitigation Strategies

Given the context outlined above we consider various mitigation strategies that could be adopted to overcome many of the challenges and pitfalls faced by organisations. We disaggregate mitigation strategies by three levels: functional, team, and corporate strategies:

FUNCTIONAL STRATEGIES

Effective metadata management is vital for understanding and optimising data within organisations, serving as a cornerstone for sustainable digital practices. Metadata not only provides context and meaning to raw data but also enables organisations to classify, tag, and manage information more efficiently. This capability is crucial for minimising the environmental impact of digital activities, a key focus of digital decarbonisation. However, despite its potential, the adoption of metadata practices faces several challenges, including the complexity of tag options, the costs associated with tagging data, and the difficulties in applying metadata to older datasets. Overcoming these obstacles is essential for improving data management, reducing storage costs, and advancing digital sustainability.

To overcome the challenges associated with metadata adoption, organisations can employ a range of solutions designed to enhance efficiency and sustainability. One such solution is the use of **automated tools and frameworks** that simplify the tagging process. These tools can significantly reduce the costs associated with metadata management while ensuring consistency across both new and legacy datasets. By automating the application of metadata, organisations can improve data accuracy and usability without incurring prohibitive costs or expending excessive energy resources.

In addition to technological solutions, fostering a **culture of responsible data management** within the organisation is essential. This involves educating employees on the importance of metadata and its role in digital decarbonisation. Regular training sessions and awareness programs can help employees understand how their actions impact the organisation's carbon footprint and why it is crucial to adopt sustainable digital practices. Encouraging a mindset of sustainability among staff will make it easier to implement and maintain effective metadata practices.

In the context of digital decarbonisation, implementing strategic approaches to metadata management is crucial. One fundamental strategy is **data minimisation**. By prioritising the collection and retention of only essential data, organisations can significantly reduce the generation of unnecessary information. This not only cuts down on storage requirements but also lowers the energy consumption associated with data management, directly contributing to a reduced carbon footprint. Data minimisation, when applied rigorously, ensures that the digital infrastructure is leaner, more efficient, and environmentally friendly.

Optimised storage solutions are another critical strategy. Ensuring that data is stored in an energy-efficient manner involves selecting the right technologies and regularly reviewing storage needs to eliminate redundant data. This approach not only conserves energy but also minimises the environmental impact of data centres, which are notorious for their high energy consumption. As data continues to grow, optimising storage becomes increasingly important in the fight against climate change. For data that must be retained for legal, regulatory, or historical reasons, **deep storage** offers a sustainable solution. Deep storage involves archiving data in a secure environment where it is not frequently accessed, thereby minimising the energy required to maintain it. This approach is particularly suitable for information that does not require regular access but must be preserved for the long term. By utilising deep storage, organisations can significantly reduce the energy costs associated with maintaining infrequently accessed data on more active storage systems, aligning with digital decarbonisation goals.

Deep storage facilities are designed to be energy-efficient, often located in areas with low environmental impact, such as cooler climates, which naturally reduce the need for energy-intensive cooling systems. This makes deep storage not only a practical solution for data preservation but also an environmentally responsible choice.

Clear and well-defined **data retention rules** are crucial for preventing unnecessary storage, whether physical or digital. These rules should specify the timeframes for retaining different types of data, as well as the criteria for determining when data should be securely deleted. By implementing such guidelines, organisations can avoid the accumulation of ROT data, which not only clutters digital storage systems but also increases energy consumption and carbon emissions.

Furthermore, establishing **strategies for timely data destruction** is essential. Automated processes should be set up to identify and delete data that has outlived its usefulness. By ensuring that data is destroyed in a timely manner, organisations can prevent the buildup of digital waste, thereby reducing their carbon footprint and contributing to a more sustainable digital environment.

TEAM STRATEGIES

Forecasting of the environmental (and cost) implications of planned digital practices is a core strategy to reducing the environmental and financial footprints of new digital projects and should be embedded organisation wide. Digitaldecarb.org has developed a comprehensive, complementary, and open Digital Decarbonisation Toolkit aimed at assisting organisations in realising their net-zero goals. This toolkit empowers organisations to assess the environmental footprint associated with data acquisition and its lifecycle within the organisation, focusing primarily on project-level evaluation (with potential extension to legacy data and projects).

The primary guiding principle behind this toolkit is the data carbon ladder, which encompasses a set of formulas driving the outputs. As described by Jackson and Hodgkinson (2023b), the journey through the data carbon ladder consists of multiple steps, and we have highlighted a few of them below:

- **Data aggregation** options include importing new datasets (creating copies that contribute to the carbon footprint), using datasets at the host server (interrogated remotely, reducing the carbon footprint), and aggregating datasets with others to fulfil project requirements (each dataset having its own carbon footprint).

- The carbon footprint of data is further determined by the **dataset size**, with a carbon score directly calculated based on its magnitude.
- **Data velocity** is another factor to consider, with options ranging from real-time data (continuously collected and analysed) to near real-time data (providing snapshots of historical information) to batch data (processed in large amounts at regular intervals) to static data (unchanging after collection).
- The rate of size increase of the dataset (**cadence**) over time (measured in megabytes, gigabytes, terabytes, or petabytes over a month) adds another CO_2 factor to consider.
- Storage methods are critical, including scenarios where data is not stored separately from the host server, stored in data centres for shared IT operations, or stored on-premises using local hardware.
- The **type of data application** used for analytics, ranging from low carbon impact (descriptive analytics) to moderately carbon-intensive (prescribed, prescriptive analytics) to highly carbon-intensive (cognitive analytics), impacts the overall data exploitation.

Moreover, when managing data and knowledge assets, employees should be able to answer the following six questions:

- *Where* are (existing) digital data being stored?
- *Why* are new data and information needed?
- *Which* information governance structures enable new knowledge to be codified?
- *Who* is responsible for sharing new data and information with others?
- *What* are the opportunities for data and knowledge reuse?
- *When* are digital data evaluated and digital waste disposed of?

Addressing each of these six questions will help to mitigate against poor digital practices and minimise the creation of dark data, single-use knowledge, and redundant, obsolete, and trivial data - collectively reducing the carbon cost of digital data. Embedding these questions into organisations digital sustainability practices is a simple but effective way to reduce data CO_2 and promote digital decarbonisation.

CORPORATE STRATEGIES

Leadership is pivotal in advancing digital decarbonisation through effective data management. Leaders must not only champion the importance of these practices but also commit to providing the necessary resources and support for their implementation. This includes investing in cutting-edge technologies, ensuring ongoing staff training, and setting clear expectations for data management across the organisation. Leadership should also emphasise the integration of digital decarbonisation into the broader corporate strategy, making it a core component of the company's sustainability goals.

By articulating a clear vision for digital decarbonisation and embedding it into the corporate culture, leaders can inspire a collective effort toward reducing the environmental impact of data management. This top-down approach ensures that all levels of the organisation are aligned with the goal of achieving sustainability through responsible data practices.

The BASE (sustainable data & ethics) framework outlined by Ian Hodgkinson and Tom Jackson (digitaldecarb.org) underscores the critical importance of sustainable growth, fairness, transparency, security, and accountability for a digital decarbonisation corporate strategy. Strategic alignment to the principles of the framework will assist in addressing the environmental impact of data infrastructure, reducing unnecessary data proliferation, and mitigating risks associated with data misuse. The BASE framework prioritises inclusive growth and sustainable development of digital capabilities, whilst helping to mitigate the environmental footprint of digital operations to ensure equitable and environmentally conscious digital behaviours. The framework is presented below (Table 2) and is informed by the following source documents: OECD, "Good Practice Principles for Data Ethics in the Public Sector", (2020); OECD, "The State of Implementation of the OECD AI Principles Four Years On", October (2023); OECD, "Common Guideposts to Promote Interoperability in AI Risk Management", November (2023); and OECD, "Recommendation of the Council on Digital Government Strategies", (2014), respectively.

Table 2 – The BASE (sustainable data & ethics) framework

DATA CONSIDERATIONS				
Sustainable growth & development	*Purpose & fairness*	*Transparency & explainability*	*Security & safety*	*Assurance & accountability*
Engage in activities that can contribute to inducing inclusive growth, sustainable development, and wellbeing.	*Respect the rule of law, human rights, and democratic values throughout data lifecycle activities,*	*Commit to responsible disclosures to provide information to foster stakeholders' understanding of data use.*	*Ensure traceability and apply systematic risk management approaches to mitigate, among others, safety and security risks.*	*Be accountable for the proper functioning of data systems and for the respect of data roles and the data context.*
• Consolidate research networks and collaborative platforms for data reduction. • Enable, guide, and foster access to, use and re-use of, data and evidence. • Reduce the potential environmental impact of data infrastructure. • Avoid the proliferation of unnecessary, redundant, or overlapping datasets. • Manage and reduce dark data volume. • Monitor and control the quality, suitability,	• Conduct human rights impact assessments, where appropriate, to analyse effects of data input activities on rights-holders. • Employ initiatives to reduce bias that may feature in data. • Protect privacy over data lifecycle. • Ensure the availability of diverse teams collaborating around data projects to help mitigate biases. • Publish data governance and management policies,	• Disclose information about historical and future use of data. • Involve expected data users to allow adjustments to data needs for successful scaling of data projects. • Reuse data based on assessment of existing data assets to increase efficiency over time. • Gather and record information on data system(s) functioning over time for evaluation. • Assess quality of data inputs.	• Maintain records of data characteristics for traceability. • Adopt and uniformly apply standards, guidelines, codes for data procurement. • Perform regular and random data audits to assess data input quality and if data is fit for purpose. • Identify and assess data risks through risk management approaches. • Communicate residual risks, data accuracy, & serious data incidents.	• Adopt clear terms of who should be held responsible for data and in which circumstances. • Establish a system for "check and balances" of decisions on spending on data and related technologies. • Establish independent oversight bodies to audit the use of data and data practices. • Use tools and processes to document data system decisions and to ensure accountability.

sustainability, and impartiality of data inputs by defining and deploying data management rules and practices. • Deploy communication tools to help facilitate engagement of appropriate stakeholders, or representatives, as and when appropriate.	practices, and procedures. • Be user-driven and place users' needs and their concerns at the core of data project design, implementation, and monitoring. • Communicate to relevant stakeholders, or representatives, about the use of data and its focal purpose. • Protect privacy of legacy data. • Protect right to freedom of expression, association, and personal autonomy. • Avoid the emergence of new forms of "digital exclusion" in the workplace, communities, and society. • Protect whistle-blowers reporting wrongdoing.	• Define a formal process for relevant parties to challenge the use of data. • Be transparent, open, and clear about data inputs and machine / human processes that led to final decisions. • Ensure the processing of personal, personal sensitive or community data by third parties in the context of public-private partnerships is transparent and comply with and adhere to applicable policy and legislation.	• Establish compliance measures where appropriate. • Agree on trustworthy data management practices across departments. • Identify user, intended data use and reasonably foreseeable data misuse (hazard identification). • Adopt impact mitigation planning (IMP) for social & environment impacts. • Track efforts to reduce and address risk(s) from data use. • Manage digital security risks and the safety of connected products and services.	• Establish codes of ethical conduct and practical technical tools for data use. • Acknowledge that the type and use-context of data determine the relevant principles, rules and norms bearing on its use. • In the case of a negative outcome, take action to ensure a better future outcome. • Articulate the value proposition for all data projects, above a certain size. • Understand potential sanctions to intended or unintended data abuse and mismanagement. • Create safe havens for reporting data misuse, negative outcomes, and early warnings.

Summary

By implementing responsible data management practices, and embedding these throughout the organisation, it is possible to significantly reduce the environmental impact of data. Key strategies include:

- **Selective Data Collection:** Focus on gathering and storing only the most critical data, minimising the generation and retention of unnecessary information.
- **Optimised Storage Solutions:** Ensure data is stored in an energy-efficient manner, regularly assessing and removing redundant data to maintain optimal storage practices.
- **Timely Data Deletion:** Establish clear protocols for the prompt deletion of data that is no longer needed, reducing the energy consumption associated with storing obsolete information.
- **Avoiding Data Redundancy:** Implement strong data governance measures to prevent unnecessary duplication of data, which can lead to increased storage and energy costs.
- **Managing Unused Data:** Regularly review and manage dark data, ensuring it is either effectively utilised or responsibly discarded.
- **Enhancing Knowledge Management:** Develop robust systems for knowledge management that promote the efficient sharing and reuse of information.
- **Staff Awareness and Training:** Educate employees on the importance of responsible data usage and the environmental impact of digital operations.
- **Responsible AI Integration:** When implementing AI, carefully consider the dataset' quantity and location, and assess the necessity of using advanced technologies like Generative AI to avoid unnecessary energy consumption.

By embedding these practices into the organisational culture, leveraging deep storage where appropriate, and fostering a commitment to sustainability, companies can effectively manage their data, reduce their carbon footprint, and contribute to a more sustainable digital future.

10. Future Outlook

Technological Advancements Supporting Digital Decarbonisation

The future of digital decarbonisation is being shaped by several emerging technological advancements. Key among these is the development of more energy-efficient data centres, powered by renewable energy sources and optimised through advanced cooling technologies and AI-driven energy management systems. Innovations in data storage, such as solid-state drives (SSDs) and the use of low-power storage solutions, are also playing a critical role in reducing the energy consumption of IT infrastructure. Additionally, the integration of AI and ML is enabling more precise and efficient data management, particularly in identifying and minimising dark data, which contributes to unnecessary energy use. Edge computing, which processes data closer to its source, is another trend that reduces the need for energy-intensive data transmission and storage in central data centres.

However, technology alone will not solve the challenge of digital decarbonisation. Organisations need to take proactive control of their digital carbon footprint, as relying solely on tech giants to manage this aspect will not sufficiently reduce their specific carbon emissions. Without a comprehensive approach to digital decarbonisation, organisations risk missing out on substantial efficiency gains, cost savings, and the potential for enhanced innovation. By actively engaging in digital decarbonisation strategies, organisations can better position themselves to achieve sustainability goals while also unlocking new opportunities for growth and improved operational performance.

Future Policies and Regulations

As concerns about climate change and sustainability grow, we can expect to see more stringent policies and regulations aimed at promoting digital decarbonisation. Governments and international bodies are likely to introduce standards that require companies to report on their digital carbon footprints and adopt sustainable IT practices. Regulations may also mandate the use of energy-efficient technologies and the reduction of electronic waste, pushing organisations to adopt circular economy principles in their IT operations. In addition, there could be increased incentives for companies to invest in green technologies, such as tax breaks or grants for implementing energy-efficient systems or using renewable energy sources for their data centres. These policies will likely drive greater transparency and accountability in how digital activities impact the environment.

Long-Term Impact - Projected Benefits for Organisations and the Environment

The long-term impact of digital decarbonisation initiatives is expected to be profound, offering significant benefits for both organisations and the environment. For organisations, adopting sustainable digital practices will lead to reduced operational costs, and more efficient use of resources. Efficient data management and the reduction of dark data will enhance data quality and accessibility, leading to better decision-making and innovation. Companies that proactively embrace digital decarbonisation are also likely to see reputational benefits, as consumers and investors increasingly favour environmentally responsible businesses.

From an environmental perspective, the widespread adoption of digital decarbonisation practices will contribute to a substantial reduction in GHG emissions associated with IT operations. This will be particularly impactful as global data usage continues to grow. By minimising the energy required to store, process, and transmit data, and by ensuring that the energy used comes from renewable sources, digital decarbonisation efforts will play a critical role in achieving global climate goals. Additionally, the reduction in electronic waste through more sustainable IT practices will help mitigate the environmental impact of technology on landfills and ecosystems.

The future of digital decarbonisation is poised to bring about significant technological, regulatory, and environmental advancements. Organisations that invest in these emerging trends will not only enhance their operational efficiency but also contribute meaningfully to the global effort to combat climate change, securing long-term benefits for themselves and the planet.

Summary

This chapter highlights the role of emerging technologies in digital decarbonisation, focusing on energy-efficient data centres, AI-driven data management, and edge computing to reduce carbon footprints. It stresses that technology alone is insufficient, organisations must actively manage their digital carbon footprint to unlock cost savings, enhance efficiency, and drive innovation. Future regulations will likely mandate sustainable IT practices, making proactive digital decarbonisation essential for long-term business and environmental benefits.

11. Conclusion

Digital decarbonisation is becoming increasingly vital as the world confronts the challenges of climate change and the growing environmental impact of digital technologies. The vast amounts of data generated and stored by organisations contribute significantly to energy consumption and GHG emissions. By adopting digital decarbonisation strategies, organisations can play a crucial role in reducing their carbon footprint while also optimising their operations.

Overview of Methods and Benefits

Several methods have been highlighted as effective in driving digital decarbonisation. These include the use of energy-efficient data centres, the implementation of metadata management to reduce dark data, and the adoption of cloud services with green certifications. Regular audits, data minimisation strategies, and the use of AI and machine learning for better data management are also essential practices. The benefits of these approaches are clear: reduced energy consumption, lower operational costs, improved data quality, and a stronger reputation for sustainability. Additionally, these practices support compliance with emerging environmental regulations and contribute to the global effort to mitigate climate change.

Call to Action - Encouragement for Organisations to Adopt Sustainable Digital Practices

As the digital landscape continues to evolve, it is imperative for organisations to take proactive steps in adopting sustainable digital practices. The shift towards digital decarbonisation is not just a trend but a necessary transformation for ensuring long-term business viability and environmental stewardship. Organisations are encouraged to assess their current digital operations, identify areas for improvement, and implement strategies that reduce energy use and minimise waste. By doing so, they will not only enhance their operational efficiency and cost-effectiveness but also contribute positively to the environment and society as a whole. Now is the time for businesses to lead by example and commit to sustainable digital practices that will benefit both their bottom line and the planet.

12. Appendices

Glossary of Terms

AI-Driven Energy Management Systems: Advanced systems that use artificial intelligence to optimise energy consumption, particularly in data centres, helping to reduce carbon footprints.

Artificial Intelligence (AI) and Machine Learning (ML): Technologies that enable systems to learn from data, identify patterns, and make decisions with minimal human intervention. These can optimise energy use and improve data management efficiency, contributing to digital decarbonization efforts.

Carbon Footprint: The total GHG emissions caused directly or indirectly by an individual, organisation, event, or product, typically expressed in equivalent tons of carbon dioxide (CO_2).

Circular Economy: An economic system aimed at eliminating waste and the continual use of resources, where materials and products are reused, refurbished, and recycled to extend their lifecycle and reduce environmental impact.

Circular Economy Principles: An economic system aimed at eliminating waste and the continual use of resources through recycling, reuse, and regeneration of materials.

Data Carbon Ladder: A framework used to calculate and assess the carbon footprint associated with data acquisition, storage, and processing within an organisation.

Data Centres: Facilities used to house computer systems and associated components, such as telecommunications and storage systems, that consume significant amounts of energy, contributing to an organisation's carbon footprint.

Data CO2 Curve: A conceptual model illustrating how different types of data storage and processing (e.g., in data centres, local devices) contribute to carbon emissions, with varying levels of impact based on their energy consumption and usage.

Data Completeness: A measure of whether all necessary data attributes are present and correct, ensuring the data satisfies all business requirements.

Data Decarbonisation: The process of reducing carbon emissions associated with digital activities, including data storage, processing, and transmission.

Data Entity Audit Framework: A tool to ensure data entities are complete and timely, helping organisations assess and evaluate the appropriateness of their data for current and future needs.

Data Misinterpretation: The incorrect analysis or understanding of data, often due to overlooked or improperly accounted for data, leading to inaccurate conclusions.

Data Minimisation: The practice of collecting only necessary data and retaining it only as long as needed to reduce storage costs and prevent the accumulation of dark data.

Data Timeliness: A measure of whether data is captured at the correct level of granularity, ensuring it is neither too coarse nor too fine for its intended purpose.

Dark Data: Data that is collected and stored by organisations but not used, which can lead to inefficiencies and increased environmental impact.

Deep Storage: A method of archiving data in a secure and energy-efficient environment where it is infrequently accessed, helping to reduce energy consumption.

Descriptive Analytics: A type of analytics focused on summarising historical data to understand what has happened, typically with a lower carbon impact than other forms of analytics.

Digital Carbon Footprint: The specific carbon emissions generated by digital activities, including data storage, processing, and the operation of IT infrastructure.

Digital Decarbonisation: A strategic approach focused on reducing the environmental impact of digital activities by optimising data management and minimising energy consumption associated with data storage, processing, and transmission.

Digital Decarbonisation Toolkit: A set of tools and guidelines designed to help organisations reduce the carbon footprint of their digital activities.

Edge Computing: A distributed computing model that processes data closer to its source, reducing the need for energy-intensive data transmission to central data centres.

Energy-Efficient Data Centres: Data centres designed to minimise energy consumption through the use of renewable energy sources, advanced cooling technologies, and other optimisation strategies.

Energy-Efficient IT Practices: Techniques and technologies that reduce the energy consumption of IT infrastructure, including the use of renewable energy, advanced cooling systems, and energy-efficient hardware.

Environmental Impact Assessment: A process for evaluating the potential environmental effects of a proposed project or action, including the assessment of digital data's contribution to GHG emissions.

GHG (Greenhouse Gas) Emissions: Emissions from gases, such as carbon dioxide and methane, that contribute to the greenhouse effect and global warming.

Green Cloud Services: Cloud services provided by data centres that operate with energy efficiency and renewable energy sources, contributing to lower carbon footprints.

Greenhouse Gas (GHG) Emissions: The release of gases, such as carbon dioxide (CO_2), into the atmosphere that contribute to global warming and climate change. Digital activities, particularly in data centres and IT infrastructure, significantly contribute to GHG emissions.

Hidden Data: Information embedded within existing datasets that require extraction or manipulation to be useful.

Hyper-Scale Data Centres: Large-scale data centres operated by major cloud providers like Amazon, Google, and Microsoft, which manage vast amounts of data and require significant energy resources, making them critical targets for digital decarbonisation efforts.

Information Management: The systematic process of collecting, storing, organising, and maintaining data to ensure its accessibility, reliability, and usefulness for decision-making.

IoT Data Streams: Data captured from Internet of Things (IoT) devices that often lack sufficient context, making them difficult to interpret and potentially unusable.

Lost Data: Data that are stored but effectively lost within data centres or devices due to poor labelling, organisation, or maintenance.

Metadata: Data that provides information about other data, such as how and when it was created, accessed, and used. Effective metadata management helps reduce digital waste and optimise data use.

Metadata Management: The process of handling metadata, which provides context and meaning to raw data, enabling more efficient data classification, tagging, and management.

Natural Language Processing (NLP): A field of artificial intelligence that focuses on the interaction between computers and human language, often used to extract information from unstructured data.

Net-Zero Emissions: The balance between the amount of greenhouse gases produced and the amount removed from the atmosphere, resulting in no net increase in atmospheric greenhouse gas levels.

Operational Efficiency: The ability of an organisation to deliver products or services in the most cost-effective manner without compromising quality, often enhanced through digital decarbonization practices.

Physical and Digital Support: The combination of physical and digital strategies to optimise data utilisation, reduce energy consumption, and enhance sustainability in managing dark data.

Records Management: The systematic control of an organisation’s records throughout their lifecycle, ensuring proper classification, storage, retrieval, and disposal to minimise data clutter and energy use.

Renewable Energy: Energy derived from natural resources that are replenished constantly, such as solar, wind, and hydro power. Using renewable energy in data centres is key to reducing the digital carbon footprint.

Renewable Energy Sources: Energy sources that are naturally replenished, such as solar, wind, and hydropower, often used in energy-efficient data centres to reduce carbon emissions.

ROT Data: Redundant, Obsolete, or Trivial data that accumulates without proper management, consuming valuable storage space and increasing energy consumption.

Scope 1, 2, and 3 Emissions: Categories used to define different sources of GHG emissions. Scope 1 refers to direct emissions from owned or controlled sources, Scope 2 to indirect emissions from the generation of purchased electricity, and Scope 3 to all other indirect emissions that occur in a company's value chain, including digital activities.

Storage Optimisation: The practice of storing data in the most energy-efficient manner, regularly assessing and eliminating redundant data to reduce environmental impact.

Sustainable IT Practices: Practices that minimise the environmental impact of IT operations, including reducing energy consumption, managing electronic waste, and adopting renewable energy sources.

System-Generated Data: Data created by systems, such as log files or output data from business intelligence (BI) and artificial intelligence (AI) processes, that are often underutilised.

Traditional Structured Data: Data that is manually input or generated into one system and used by other systems, potentially containing valuable information or required for legal purposes.

Unstructured Data: Data in formats like videos, audio files, emails, and documents that require conversion to a structured or semi-structured format to extract valuable insights.

Velocity of Data: The speed at which data is generated, processed, and analysed, which can impact the carbon footprint depending on whether it is real-time, near real-time, or batch-processed.

References

Acquaye, A., Genovese, A., Barrett, J. and Koh, S.L. (2014), "Benchmarking carbon emissions performance in supply chains", *Supply Chain Management: An International Journal*, 19(3), pp. 306-321.

Ciacci, L., Fishman, T., Elshkaki, A., Graedel, T.E., Vassura, I. and Passarini, F. (2020), "Exploring future copper demand, recycling and associated greenhouse gas emissions in the EU-28", *Global Environmental Change*, 63, pp. 1-10.

Climatiq. (2022), "Measuring greenhouse gas emissions in data centres: the environmental impact of cloud computing" available at: https://www.climatiq.io/blog/measure-greenhouse-gas-emissions-carbon-data-centres-cloud-computing (Accessed on 20 August 2024).

Daniel, L.E. and Daniel, L.E. (2012), *Digital Forensics for Legal Professionals*, Elsevier Inc.

Department for Business, Energy and Industrial Strategy. (2021), *Industrial Decarbonisation Strategy*. London: Her Majesty's Stationery Office (HMSO).

Gartner. (2022), "Dark data" available at: https://www.gartner.com/en/information-technology/glossary/dark-data (Accessed on 14 August 2022).

Hand, D.J. (2020), *Dark Data: Why What you Don't Know Matters*, Princeton University Press.

Hodgson, C. (2024), "Booming AI demand threatens global electricity supply", *Financial Times*, April available from: https://www.ft.com/content/b7570359-f809-49ce-8cd5-9166d36a057b (Accessed on 18 April 2024).

Jackson, T and Hodgkinson, I.R. (2023a), "Keeping a lower profile: How firms can reduce their digital carbon footprints", *Journal of Business Strategy*, 44(6), pp. 363-370.

Jackson, T. and Hodgkinson, I.R. (2023b), "Is there a role for knowledge management in saving the planet from too much data?", *Knowledge Management Research & Practice*, 21(3), pp.427-435.

Jackson, T.W. and Hodgkinson, I.R. (2024a), "Debate: The data threat to 2050 net zero—public administrations' responsibility for the 'data-scape'", *Public Money and Management*, 44(3), pp. 182-184.

Jackson, T.W. and Hodgkinson, I.R. (2024b), *Why, where and when dark data affects greenhouse gas emissions*, Policy Briefing: International Advisory Group, Academy of Social Sciences.

Lockwood, S., Hodgkinson, I.R., Jackson, T.W., Jackson, L., and Barbour, E. (2023), "Data decarbonisation, greenhouse gas emissions, and industrial business", In British Academy of Management Conference, *BAM2023 Proceedings*, UK.

Medium.com. (2022), “How much power do the servers consume?” available at: https://medium.com/@zodhyatech/how-much-power-do-the-servers-consume-8f9620be71a (Accessed on 28 August 2024).

PWC. (2023), “Accelerated action on climate change requires all hands on deck” available at: https://www.pwc.com/gx/en/services/sustainability/publications/accelerated-action-on-climate-change-calls-for-all-hands-on-deck.htmlhttps://www.pwc.com/gx/en/services/sustainability/publications/accelerated-action-on-climate-change-calls-for-all-hands-on-deck.html (Accessed on 07 March 2023).

Riley, J. (2017), *Understanding Metadata*, A Primer Publication of the National Information Standards Organization, Washington DC, United States.

Statista. (2020), “Data volume of internet of things (IoT) connections worldwide in 2019 and 2025” available at: https://www.statista.com/statistics/1017863/worldwide-iot-connected-devices-data-size/ (Accessed on 09 March 2022).

Teuful, B. and Sprus, C.M. (2020), “How digitalization acts as a driver of decarbonization” available at: https://www.ey.com/en_ch/decarbonization/how-digitization-acts-as-a-driver-of-decarbonization (Accessed on 27 February 2022).

The OECD. (2023), *The State of Implementation of the OECD AI Principles Four Years on*, OECD Publishing.

The OECD. (2023), *Common Guideposts to Promote Interoperability in AI Risk Management*, OECD Publishing.

The OECD. (2023), *Recommendation of the Council on Open Government*, OECD Publishing.

The OECD. (2020), *Good Practice Principles for Data Ethics in the Public Sector*, OECD Publishing.

The OECD. (2014), *Recommendation of the Council on Digital Government Strategies*, OECD Publishing.

United Nations. (2022), “Act now: facts and figures” available at: https://www.un.org/en/actnow/facts-and-figures [Accessed on 18 August 2024].

Yokoi, R., Watari, T. and Motoshita, M. (2022), “Future greenhouse gas emissions from metal production: gaps and opportunities towards climate goals”, *Energy & Environmental Science*, 15(1), pp.146-157.

Want to learn more?

Contact us: digitaldecarbonisation@oasisgroup.com

Want the latest updates and insights?

Scan the QR code to sign up and stay informed!