

Greenwheel Insights

Feeding the Machine: Data centres, resources, and the environment

June 2024



Paul Drummond Climate and Environment Lead, Greenwheel

Executive Summary

- Data centres provide the physical foundation for the digital world. They use 1.5% of global electricity, and new demand driven mostly by non-Al uses is likely to grow data centre electricity use by 2.5x by 2030.
- Renewables covered 20% of data centre electricity demand in 2023, but more effort is needed to drive additional, rather than drawing on existing capacity.
- Grid constraints are likely to curtail new data centres in some regions. Along with the need for sufficient and reliable electricity supply, this may drive an increase in fossil fuel generation – particularly gas.
- Data centers usually have diesel generators to provide back-up electricity. New options, including batteries, are zero-carbon alternatives.

- Energy efficient cooling technologies are water intensive. A 'hyperscale' data centre can use as much water as a mid-size town.
- Hybrid air-liquid cooling technologies can balance electricity and water use, but total data centre water use is set to increase significantly. Closed loop circulation, rainwater harvesting and wider watershed management activities can reduce pressure on water resources.
- Data centres are also material intensive and use high volumes of copper. New data centres in North America alone drive 1% of global copper demand. Data centres can be designed for component and material reuse. Material that cannot be reused may be recycled, or e-waste can be responsibility disposed of.

The following questions form the foundation of a **framework for investors to** assess the environmental footprint of companies operating data centres.

- 1. Is the data centre sited in an area with suitable resources and connectivity?
- 2. Is the data centre designed to maximise the efficient and circular use of materials and components?
- 3. Is the cooling system designed to minimise both electricity and water intensity?
- 4. How is the company minimising its impact on freshwater abstraction and local watershed?
- 5. Is the use of zero-carbon electricity maximised, matched to demand and with additionality?
- 6. Does the data centre have a zero-carbon uninterruptable power supply?
- 7. Has consultation with local communities and other stakeholders been conducted?



CONTACT US

Please contact us if you would like to learn more about our strategies.

Background

Data centres provide the physical foundation for the digital world, and are resource hungry. Their use accounts for around 1.5% of global electricity consumption and associated CO₂ emissions, with data transmission networks accounting for the same again.ⁱ They also require significant volumes of water for cooling and can be material-intensive to build and refurbish.

There are over 5,000 data centres in the USA - around a third of the global total, and >10x more than any other country. Germany, the UK, China and Canada round out the top 5, with 300-500 data centres each.ⁱⁱ

As artificial intelligence (AI) grows in sophistication and use, demand for data centres to deliver it will accelerate. This could have significant implications for the development of electricity systems and their CO₂ emissions, water resource security and competition, and demand for key raw materials. These factors could in turn restrict data centre development and location.

This briefing provides an overview of the resource and environmental implications of data centres. It also **provides a framework for investors to assess the environmental footprint of companies operating data centres**.

Diverging demand for computation and electricity

Despite rapid growth in data centre capacity and workloads in recent years, data centre energy use has grown modestly (excluding cryptocurrency mining). This is due to a combination of three interrelated factors, illustrated below in Figure 1.

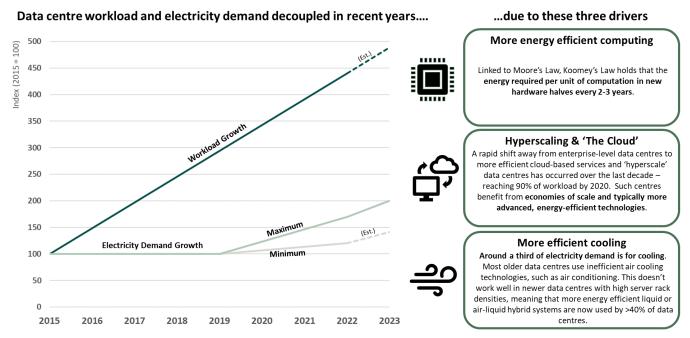


Figure 1 – Data centre workload and energy use decoupling, and its drivers (2015-2023). Data Sources: <u>IEA</u> (2023); <u>Goldman Sachs (2024)</u>. Workload growth and minimum electricity demand growth estimated for 2023, based on historic growth rate and maximum electricity growth rate, respectively. Graphic created by Greenwheel.



Despite a trebling in workload, electricity demand from data centres was held flat over 2015-2019 by these three factors. Electricity demand began to grow from 2020 as the efficiency gain from the shift from existing enterprise to hyperscale data centres (typically larger than 5,000 servers and 10,000 ft²) and cloud computing became largely exhausted.

By 2023, data centre electricity demand reached 300-400TWh (excluding crypto), equivalent to electricity demand in Italy at the lower end, with the addition of the Philippines at the higher end. **Compared to all other forms of data, workload and electricity demand from AI processes is currently negligible**.

Electricity demand for cryptocurrency mining – excluded from this brief - has grown from almost negligible in 2015 to equivalent to at least 30% of data centre electricity demand for all other purposes in 2022ⁱ, or the electricity consumption of the Netherlands.

Power Usage Efficiency (PUE) is the ratio between the total electricity used by a data centre and the power used by the IT equipment alone. **The more efficient the facility, the closer the value is to 1. The global average in 2007 was 2.5, reducing to 1.55 in 2022**^{III}, due to hyperscaling, cloud computing, and more efficient cooling technologies. **Most new hyperscale new data centres have PUEs of <1.2**^{III}, with some achieving <1.1.^{IV}

Can this divergence continue?

Little efficiency potential from the move to hyperscale data centres and cloud computing remains. However, demand for faster response times (low latency) for some applications, such as autonomous vehicle and smart grid management, may drive an increase in 'edge' computing – smaller data centres located near the end user. Data protection and sovereignty issues may also drive an increase in edge computing.

Edge computing is less scale-efficient but can reduce electricity consumption in four ways. First, edge data centres can operate when and to the extent required, including through shifting compute demand to optimise electricity use. By contrast, hyperscale and cloud data centres operate constantly. Secondly, a larger surface area around IT equipment allows for more efficient cooling. Thirdly, heat can be recovered for use in colocated industries or heating networks. Fourth, reducing distance to end users reduces electricity demand for data transmission.^v

More efficient cooling technologies (discussed below), **and other energy management approaches can continue to drive PUE values down toward 1. This will be a key driver for efficiency in the coming years**. Regulations around maximum PUE values have begun to enter force, but with highly varied ambition; as low as 1.2 in Germany and the Netherlands and 1.25 in China, but as high as 1.9 in Malaysia. The US has no regulations beyond federal government-operated data centres, and the EU requires only disclosure of energy performance from data centres from 2024.^{vi}

It is likely that the most significant factor driving the relationship between computation and electricity demand is 'Koomey's law', which observes that the



energy required per unit of computation in new hardware halves every 2-3 years, because of Moore's Law and wider improvements to system architecture.¹

Ultimately, it is likely that data centre workload demand growth will significantly outpace any gains in energy efficiency, increasing total electricity demand substantially.

This is partly driven by the likely growth in AI applications. **AI requires significantly more** electricity than non-AI applications, as it uses much more powerful hardware.

Al systems use Graphics Processing Units (GPUs) rather than traditional Central Processing Units (CPUs), due to their capacity for parallel processing. Although GPUs are less electricity intensive than CPUs, their significantly higher compute capacity means that they consume far more electricity in total.

For example, a ChatGPT query uses 6-10x more electricity than a Google search.^{viii} Al training is also highly electricity intensive. Estimates suggest that training ChatGPT-4 used as much electricity as 20,000 UK households use in a year.^{vii}

Although AI demand will grow significantly, the primary driver of data centre electricity demand growth to 2030 is likely to be from non-AI purposes.^{viii,ix}

Electricity demand for data centres is projected to grow 2.5x over 2023-2030 (Figure 2, excluding crypto) according to the IEA^{ix} and Goldman Sachs^{viii}, with this growth equivalent to electricity used in the UK and Germany combined. Data centres would account for 3-4% of global electricity demand in 2030, and 3-8% in the USA.

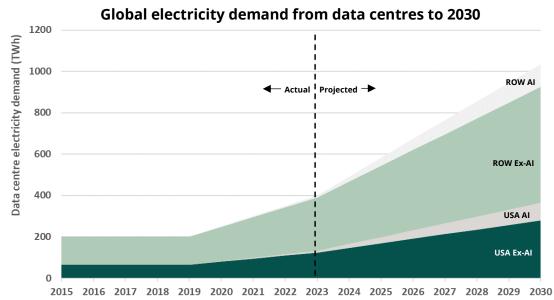


Figure 2 – Global electricity demand from data centres, 2015-2023 and estimated to 2030. Source: <u>Goldman</u> <u>Sachs (2024)</u>.

¹ Between the 1950s and 2000 a halving in electricity intensity occurred every 1.6 years, slowing to 2.6 years after 2000 due to Moore's Law (the rate of doubling in the number of transistors in an integrated circuit) decelerating and the end of Denning Scaling (power density of transistors is constant with declining size).



Al-based electricity demand grows from negligible to around 20% of all data centre demand by 2030, in both projections. However, there is significant uncertainty around these projections, with the difference between the upper and lower bounds by 2030 equivalent to the electricity consumption of Brazil.

There are three key drivers behind such significant uncertainty (Figure 3).



Compute demand & costs

New AI hardware may have greater compute capacity and be more energy efficient, but require more electricity in total and be more expensive to install and operate. Demand for additional compute capacity and/or budgets may be limited, constraining growth.



Some AI applications will be new and drive additional demand for computation and data centres. Others will overlap with existing, non-AI applications, displacing existing computation and data centre demand. The extent of this overlap in the coming years is uncertain.



Infrastructure & resource constraints

Electricity networks around the world are already under substantial pressure, and restrictions on the connection of new, large data centres may persist and grow. Land and particularly water availability for cooling may also constrain development in some locations.

Figure 3 – Drivers of uncertainty behind electricity demand for AI data centres. Graphic created by Greenwheel.

A boon for fossil fuels or renewables?

Most data centres are powered by their local grids.. Their 'location-based' Scope 2 emissions are therefore driven by the local electricity generation mix, which is highly varied across the world.^x

...although some data centre operators are the largest corporate procurers of renewable energy in the world. Procurement is primarily through renewable Power Purchase Agreements (PPAs) and the purchase of 'renewable guarantee of origin' certificates. This reduces their 'market-based' Scope 2 emissions. Renewables covered around 20% of data centre electricity demand in 2023, from effectively zero in 2015.^{viii} Several major data centre owner and operators procure enough renewable electricity to cover all operational electricity demand.ⁱ

However, using PPAs or buying certificates does not necessarily match renewable supply to data centre demand in either time or location, and may draw on existing rather than drive development of new renewable capacity. This makes the realworld decarbonisation impact unclear.

A growing number of data centre operators, as signatories to the EU Climate Neutral Data Centre Pact, aim to procure zero-carbon electricity procurement for all demand matched by location and hour of demand by 2030.^{xi}

Data centres are usually equipped with uninterruptable power supplies (UPS) – most often diesel generators – to provide electricity if the grid fails. As grid failures in developed economies have been uncommon, they have usually only been used – and



produce emissions - in testing. Data centres in developing economies that experience greater risk of grid failure are more reliant on UPS systems, and use co-located renewables that attempt to mitigate some of this risk (e.g. South Africa).^{xii} However, **the scale and consistency of data centre electricity demand means that relying solely or mostly on co-located variable renewables is not possible** in most instances.

Some data centre operators have begun testing and using alternative, low-carbon fuels and technologies for UPS systems, while yet others have been proposed as options for the future. Figure 4 presents these options in order of maturity. Depending on their characteristics, some such options could be used in place of grid electricity, for at least some of the time.

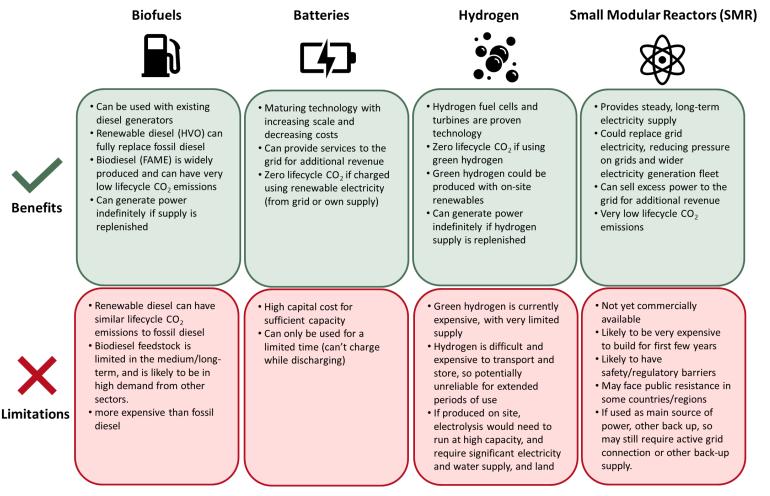


Figure 4 – Low carbon uninterruptable power supply options. Sources include: <u>Judge (2023)</u>; <u>Lepineux (2023)</u>; <u>Carlini (2023)</u>. Graphic created by Greenwheel.

Grid connection queues and capacity constraints are likely to curtail new data centres in some regions, particularly as the required size of the connection increases for dedicated AI sites. For example, **a moratorium for new data centre grid connections has been in place in Dublin since 2022**. Data centres used 17% of electricity generated in Ireland in 2022 and are expected to consume a third by 2026.^{ix}



Attempts to overcome these constraints may impact local communities. For example, new grid infrastructure may be built without sufficient consultation or consent, or be completed at the expense of grid infrastructure required to supply other end users. It may also risk the stability of electricity supply to other end-users.²

The current inability of on-site low carbon technologies to satisfy a data centre's electricity demand, coupled with grid constraints and the need for a reliable electricity system means that **a rapid growth in data centres is likely to lead to growing electricity generation from fossil fuels – particularly gas.**

Gas power plants can be built quickly, can connect to existing gas pipelines, generate electricity on-demand, and produce half the CO₂ emissions of coal power plants. Existing gas plants on the grid may generate more for longer, and new gas capacity may be added either to the grid, or be co-located, owned and operated by data centre operators. For example, **Microsoft is constructing a gas plant to power a new data centre in Dublin**, due to the grid connection moratorium.

The extent to which these constraints put data centre renewables and emissions targets at risk, or whether these targets will curtail an increase in fossil fuel generation, is unclear. However, pressure from both directions is likely to be present.

Keeping cool is thirsty work

Traditional air-cooling approaches are energy intensive and generally unsuited to data centres with high computing loads and cooling demand, such as those supporting AI. Figure 5 (following page) illustrates key air, liquid and hybrid cooling technologies, and their implications for electricity and water use. **Current hyperscale and AI data centres largely draw on liquid and hybrid (evaporative) cooling systems**.

Water Usage Efficiency (WUE) measures the ratio of water used by the data centre to the power used by the IT equipment within the data centre. The lower the number, the higher the water efficiency. The global average is currently 1.8L/kWh.^{xiii}

Global average WUE values for major hyperscale data centre owners and operators range from <0.2L/KWh to >1.7L/kWh, due to differences in ambient temperatures and cooling technology required.^{vi,xiv} Those with the lowest values mainly use hybrid switchable systems, where the local environment is conducive.

This illustrates that while **hybrid (particularly switchable) systems can produce both water and electricity savings, their use and scale of savings available depends significantly on local environmental conditions** (e.g. cold ambient air).

Despite these low values, their large electricity and cooling demand means that **an average hyperscale data centre uses the same volume water per year as a town of around 35,000 people.**^{xv} At least two hyperscale data centre operators draw around a

² See Greenwheel Briefing: "Rage against the AI machine? Impact on Communities and End-Users" for more information



fifth of their freshwater from water stressed areas.^{xvi} Data centre cooling typically requires high quality water, meaning operators usually draw on potable resources.

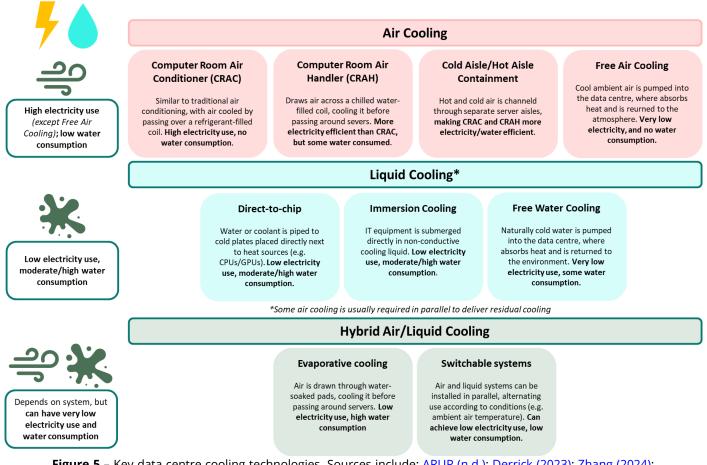


Figure 5 – Key data centre cooling technologies. Sources include: <u>ARUP (n.d.)</u>; <u>Derrick (2023)</u>; <u>Zhang (2024)</u>; <u>Sheldon (2024)</u>; Yang et al (2024). Graphic created by Greenwheel.

Water use estimates increase when incorporating indirect (upstream) consumption. In the USA, around five times more water is used in the generation of electricity used by data centres than is used by data centres directly.^{vi} Pressure on water resources also occurs upstream in the supply chain, from semiconductor manufacture to the extraction and processing of other raw materials, such as copper.^{xvii}

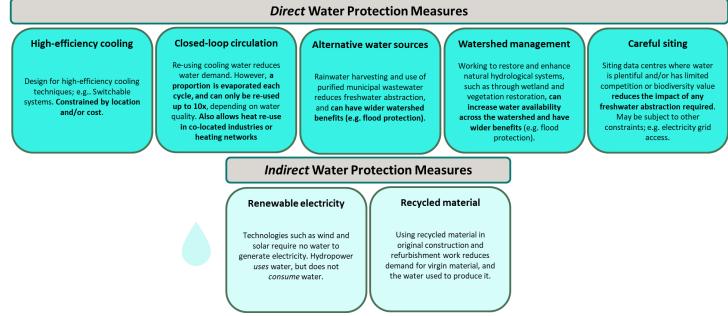
Demand for cooling is likely to increase competition for water resources. This can threaten the rights of different stakeholders, producing conflict which can in turn threaten the operation of a data centre, or prevent its construction.³ Significant protest movements against data centres have been seen around the world, including Ireland, the Netherlands, and Chile.^{xviii} Excessive water consumption is among the most common reasons for organisations to close data centers.^{vi}

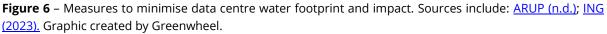
Water demand from a data centre and its impact on local water resources can be minimised through seven complementary measures, illustrated by Figure 6. **Some direct water efficiency measures are more difficult to retrofit into existing data centres**

³ See Greenwheel Briefing: "Rage against the AI machine? Impact on Communities and End-Users" for more information.



compared to new capacity, with retrofit potential specific to the data centre and its location. **However, both new and existing data centres can adopt measures to reduce their indirect water footprint**. From 2025, data centres in the EU must report water use and the reuse of waste heat in nearby facilities and heat networks.^{xix} Comparable regulations are not currently found elsewhere.





Many major hyperscale data centre operators have made commitments to reduce the water intensity of their data centres, and to use watershed management techniques to replenish at least as much water as they consume by 2030. However, progress is slow.^{xx}

Despite the range of measures available to moderate the water footprint of data centres, the **rapid growth in demand for new data centres means that water consumed by the sector** (both directly and in thermal power generation to supply it) **is projected to grow by >50%.** For AI demand alone, water abstraction may equal half that abstracted annually in the UK by 2027 (for both direct use and power generation).^{vi}

The material behind the digital

Data centres are material intensive. Significant volumes of steel and cement are used to construct their shells, the production of which produces GHG emissions equivalent to a third of the emissions from generating electricity to power the data centre.^{xxi} **This is despite a significant increase in building material efficiency due to the shift to hyperscale and cloud computing**.

A wide range of other materials comprise the IT equipment, communication networks and ancillary equipment that allow the data centre to function.

Demand for copper is particularly strong. Copper is mainly used for power distribution, but also grounding, communication and more general building services such



as cooling and plumbing systems. An average hyperscale data centre requires 900-2,200 tonnes of copper.^{xxii} Data centres constructed in 2020 in North America alone accounted for ~1% of global copper demand that year.^{xxiii}

Although building material intensity relative to compute capacity is likely to continue decreasing, copper intensity is likely to remain high.

Forecasts of additional annual global copper demand for data centre construction by 2030 vary significantly, from 200,000 tonnes^{xxiv} to 1 million tonnes per year^{xxv} – equivalent to ~1% and ~3% of current total global copper demand. The production of copper and other key raw materials in the data centre value chain are highly geographically concentrated, and many are associated with environmental and social risks.^{xxvi}

Material used in data centres eventually becomes waste. Data centre buildings may last 15-20 years but severs are typically replaced every 3-5 years. This means **the rate of e-waste generation is likely to become very substantial**. Although e-waste is well regulated in some jurisdictions (e.g. the EU), only around half of US States regulate e-waste. E-waste regulation is also largely absent from most emerging economies.^{xxvii}

Data centres can be designed for their equipment to be reused. This is particularly the case for hyperscale data centres, which use customised components. This equipment can either be retained and used for less demanding workloads, or sold on for the same purpose, also allowing residual value to be captured. Truly end-of-life equipment can be recycled to the extent possible, or otherwise disposed of responsibly.^{xxvii}

Although data centre buildings can have a large footprint, they are not sprawling. A hyperscale data centre defined as over 10,000 ft² – smaller than an Olympic size swimming pool. As hardware capacity continues to improve land intensity will decrease, but **the rate of demand growth will increase the total land required to support data centres**. Sites should be chosen to minimise their environmental and social impact.

What does this mean for investors?

Almost all businesses rely on data centres, but in many cases this reliance is indirect. The data centres that underpin generic digital services are generally not considered in methodologies that measure the environmental impact of business that use them. Their ability of businesses to engage and influence providers of generic digital services, or change provider, is also likely to be limited.

However, **business that build, own, operate or procure dedicated capacity in data centres have greater influence over their design and operation**, and the environmental impacts they carry. For these businesses and their investors seeking to minimise and manage these impacts, **the following framework of questions may be applied** (with associated questions in *italics*):



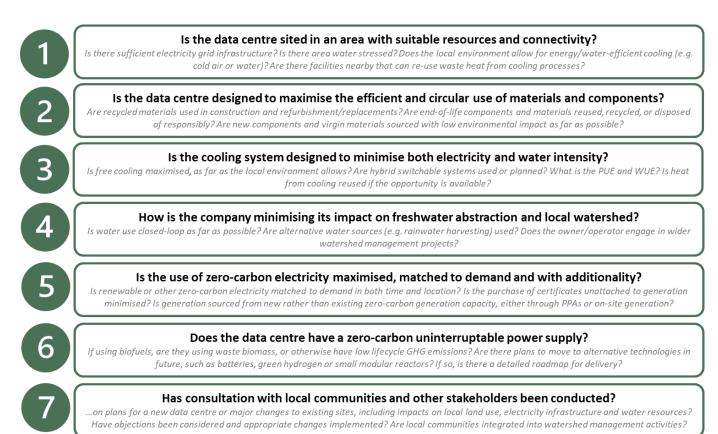
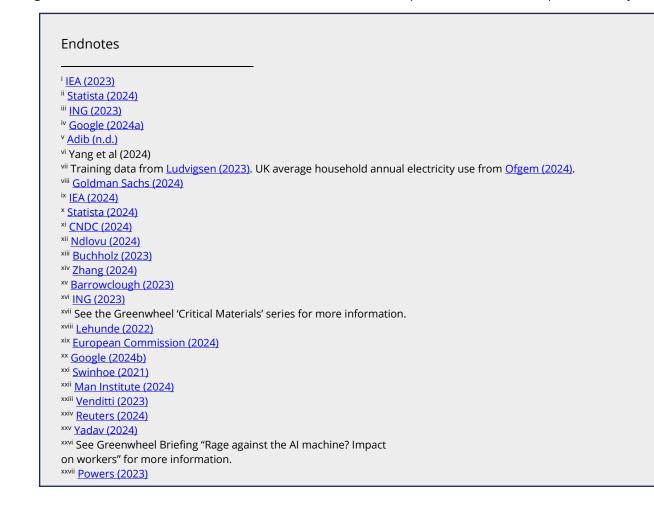


Figure 7 – Framework for investors to examine the environmental impact of data centres. Graphic created by Greenwheel.





Key Information

No investment strategy or risk management technique can guarantee returns or eliminate risks in any market environment. Past performance is not a guide to future results. The prices of investments and income from them may fall as well as rise and an investor's investment is subject to potential loss, in whole or in part. Forecasts and estimates are based upon subjective assumptions about circumstances and events that may not yet have taken place and may never do so. The statements and opinions expressed in this article are those of the author as of the date of publication, and do not necessarily represent the view of Redwheel. This article does not constitute investment advice and the information shown is for illustrative purposes only. Whilst updated figures are not available for all sources, we have performed further analysis and believe that this data has not significantly changed and is reflective for 2024.

Disclaimer

Redwheel ® is a registered trademark of RWC Partners Limited ("RWC"). The term "Redwheel" may include any one or more Redwheel branded entities including RWC and RWC Asset Management LLP, each of which is authorised and regulated by the UK Financial Conduct Authority and, in the case of RWC Asset Management LLP, the US Securities and Exchange Commission ("SEC"); RWC Asset Advisors (US) LLC, which is registered with the SEC; RWC Singapore (Pte) Limited, which is licensed as a Licensed Fund Management Company by the Monetary Authority of Singapore; Redwheel Australia Pty Ltd is an Australian Financial Services Licensee with the Australian Securities and Investment Commission; and Redwheel Europe Fondsmæglerselskab A/S which is regulated by the Danish Financial Supervisory Authority.

Redwheel may act as investment manager or adviser, or otherwise provide services, to more than one product pursuing a similar investment strategy or focus to the product detailed in this document. Redwheel seeks to minimise any conflicts of interest, and endeavours to act at all times in accordance with its legal and regulatory obligations as well as its own policies and codes of conduct. This document is directed only at professional, institutional, wholesale or qualified investors. The services provided by Redwheel are available only to such persons. It is not intended for distribution to and should not be relied on by any person who would qualify as a retail or individual investor in any jurisdiction or for distribution to, or use by, any person or entity in any jurisdiction where such distribution or use would be contrary to local law or regulation.

This document has been prepared for general information purposes only and has not been delivered for registration in any jurisdiction nor has its content been reviewed or approved by any regulatory authority in any jurisdiction.

The information contained herein does not constitute: (i) a binding legal agreement; (ii) legal, regulatory, tax, accounting or other advice; (iii) an offer, recommendation or solicitation to buy or sell shares in any fund, security, commodity, financial instrument or derivative linked to, or otherwise included in a portfolio managed or advised by Redwheel; or (iv) an offer to enter into any other transaction whatsoever (each a "Transaction"). Redwheel bears no responsibility for your investment research and/or investment decisions and you should consult your own lawyer, accountant, tax adviser or other professional adviser before entering into any Transaction. No representations and/or warranties are made that the information contained herein is either up to date and/or accurate and is not intended to be used or relied upon by any counterparty, investor or any other third party.

Redwheel uses information from third party vendors, such as statistical and other data, that it believes to be reliable. However, the accuracy of this data, which may be used to calculate results or otherwise compile data that finds its way over time into Redwheel research data stored on its systems, is not guaranteed. If such information is not accurate, some of the conclusions reached or statements made may be adversely affected. Any opinion expressed herein, which may be subjective in nature, may not be shared by all directors, officers, employees, or representatives of Redwheel and may be subject to change without notice. Redwheel is not liable for any decisions made or actions or inactions taken by you or others based on the contents of this document and neither Redwheel nor any of its directors, officers, employees, or representatives (including affiliates) accepts any liability whatsoever for any errors and/or omissions or for any direct, indirect, special, incidental, or consequential loss, damages, or expenses of any kind howsoever arising from the use of, or reliance on, any information contained herein.

Information contained in this document should not be viewed as indicative of future results. Past performance of any Transaction is not indicative of future results. The value of investments can go down as well as up. Certain assumptions and forward looking statements may have been made either for modelling purposes, to simplify the presentation and/or calculation of any projections or estimates contained herein and Redwheel does not represent that that any such assumptions or statements will reflect actual future events or that all assumptions have been considered or stated. There can be no assurance that estimated returns or projections will be realised or that actual returns or performance results will not materially differ from those estimated herein. Some of the information contained in this document may be aggregated data of Transactions executed by Redwheel that has been compiled so as not to identify the underlying Transactions of any particular customer.

No representations or warranties of any kind are intended or should be inferred with respect to the economic return from, or the tax consequences of, an investment in a Redwheel-managed fund.

This document expresses no views as to the suitability or appropriateness of the fund or any other investments described herein to the individual circumstances of any recipient.

The information transmitted is intended only for the person or entity to which it has been given and may contain confidential and/or privileged material. In accepting receipt of the information transmitted you agree that you and/or your affiliates, partners, directors, officers and employees, as applicable, will keep all information strictly confidential. Any review, retransmission, dissemination or other use of, or taking of any action in reliance upon, this information is prohibited. Any distribution or reproduction of this document is not authorised and is prohibited without the express written consent of Redwheel.

Funds managed by Redwheel are not, and will not be, registered under the Securities Act of 1933 (the "Securities Act") and are not available for purchase by US persons (as defined in Regulation S under the Securities Act) except to persons who are "qualified purchasers" (as defined in the Investment Company Act of 1940) and "accredited investors" (as defined in Rule 501(a) under the Securities Act).



This document does not constitute an offer to sell, purchase, subscribe for or otherwise invest in units or shares of any fund managed by Redwheel. Any offering is made only pursuant to the relevant offering document and the relevant subscription application. Prospective investors should review the offering memorandum in its entirety, including the risk factors in the offering memorandum, before making a decision to invest.

CONTACT US

Please contact us if you have any questions or would like to discuss any of our strategies. invest@redwheel.com | www.redwheel.com

Redwheel London Verde 10 Bressenden Place London SW1E 5DH +4420 7227 6000

Redwheel Europe Fondsmæglerselskab A/S, Havnegade 39, 1058 København K, Denmark

Redwheel Miami 2640 South Bayshore Drive 80 Raffles Place Suite 201 Miami Florida 33133 +13056029501

Redwheel Singapore #22-23 UOB Plaza 2 Singapore 048624 +65 68129540

