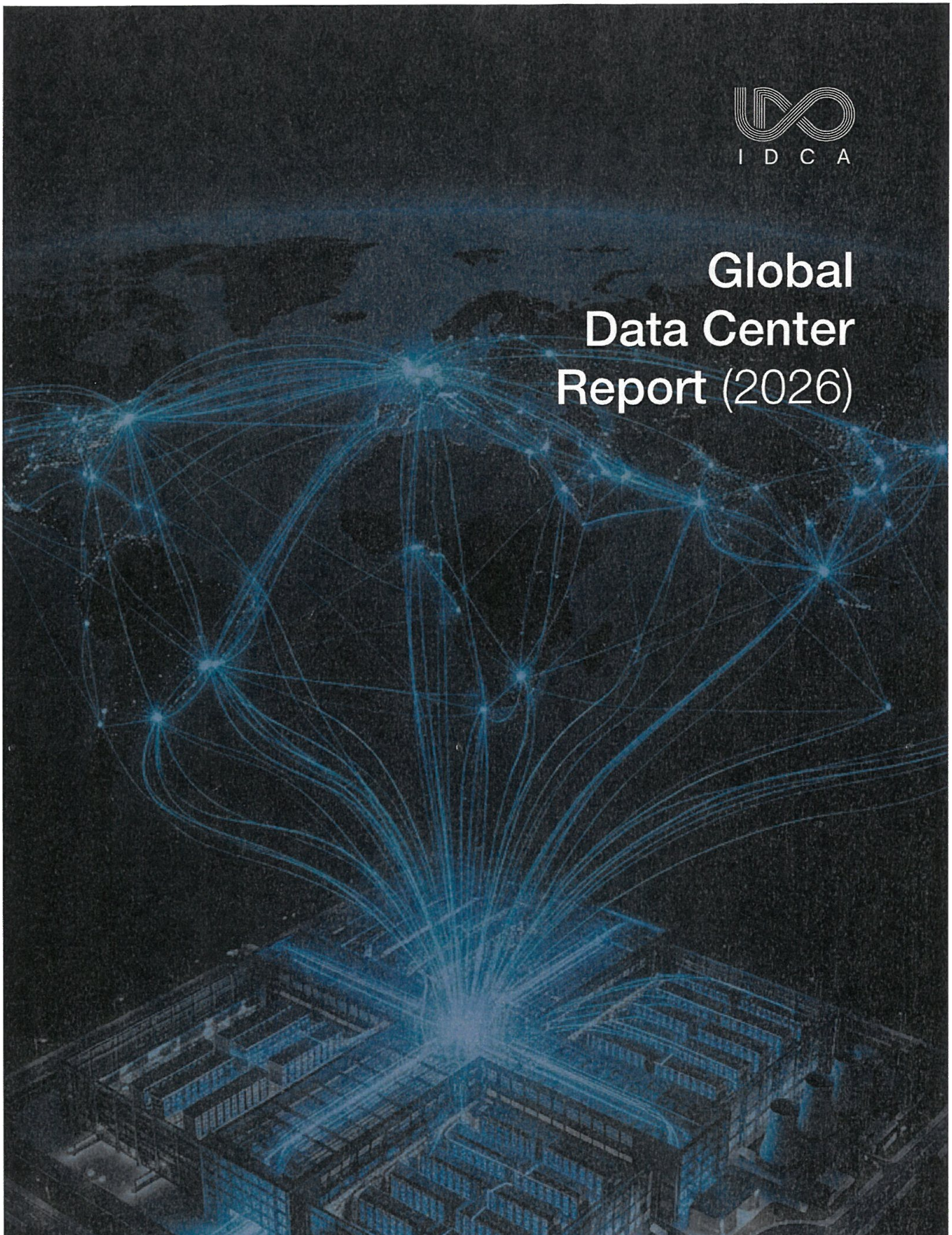
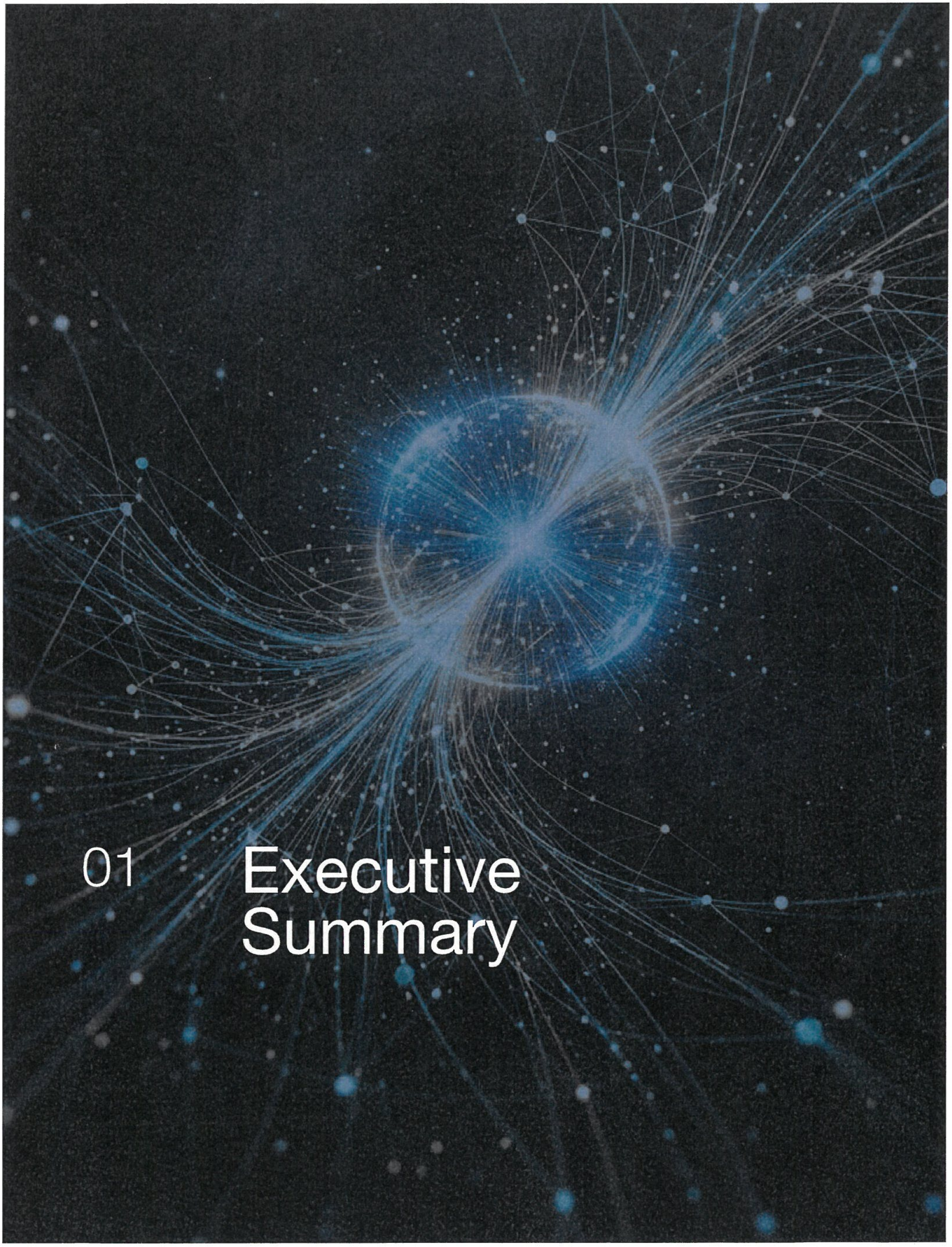
Global Data Center Report (2026)





01 Executive Summary

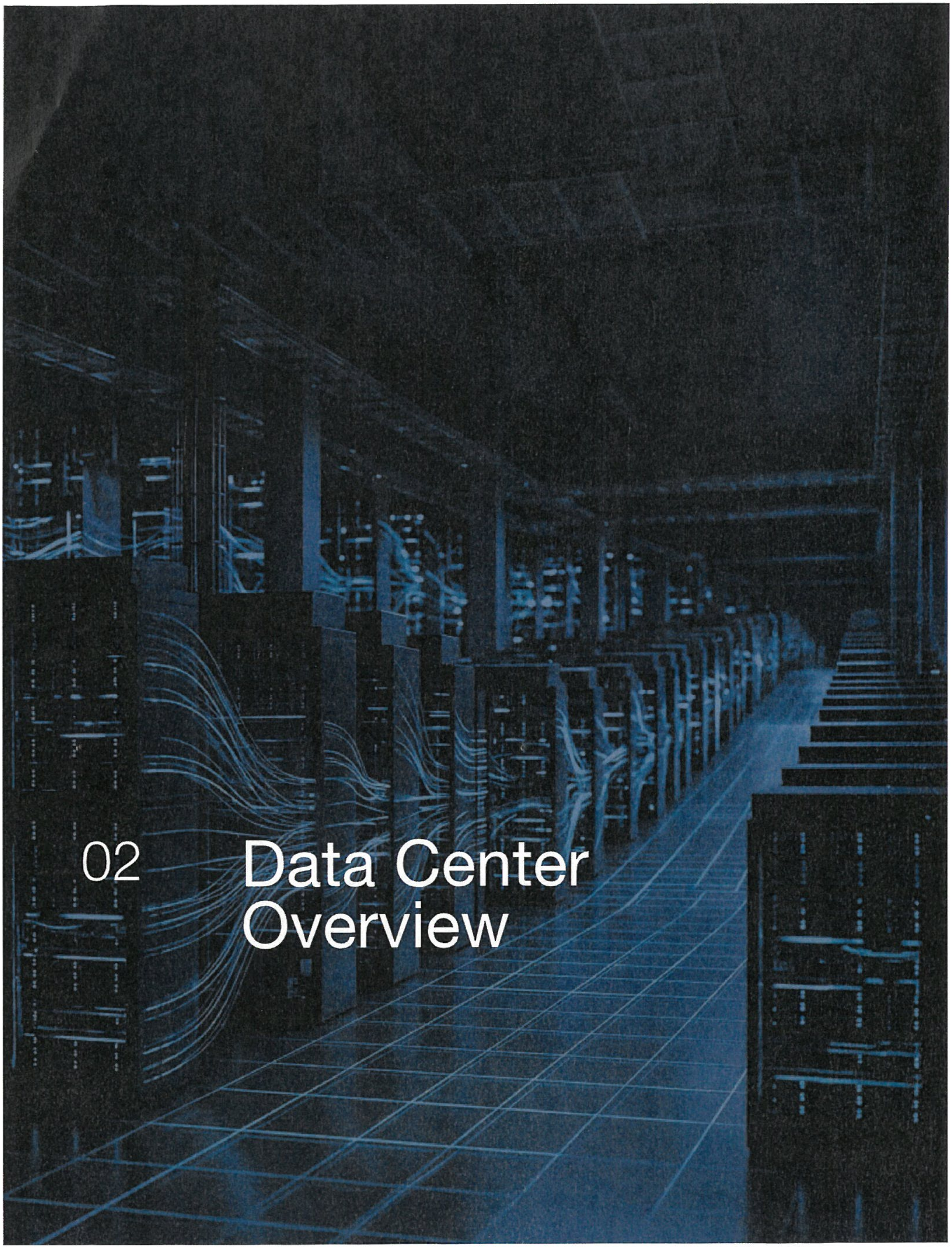
Executive Summary

This report accompanies the IDCA Digital Economy Report 2026, which was released in January 2026. The previous report focused on the global buildout of the Digital Economy, with an analysis of all aspects of digital infrastructure and its impact on societal prosperity.

This report, in contrast, focuses solely on data centers – their consumption, their location, their national footprints, their design, and the considerations that businesses, governments, other organizations, and investors should take into account when planning new data center facilities and hubs.

The Digital Economy Report, Data Center Report, and other IDCA reports are a product of IDCA's global research team, with a presence in each of the world's regions. IDCA's research process scans a number of publicly available reports from organizations such as the International Energy Agency, World Bank, United Nations, and International Telecommunications, as well as agencies within the United States, EU, and numeros additional governmental entities.

IDCA also receives private feedback from specific government sources globally, from major data center and cloud-computing operators, from developers, from investors, and from internal research.



02 Data Center Overview

Data centers significantly impact the environment through high energy consumption, substantial water usage, and carbon emissions, particularly as demand for cloud services and AI technologies grows.

Energy Consumption

Data centers are major consumers of electricity, accounting for nearly **2% of global energy demand** in 2022, with projections suggesting this could double by 2026 due to the increasing reliance on AI and cloud services applus.com. The energy-intensive nature of data centers is exacerbated by the need for cooling systems, which can consume over **40% of their total electricity** aaas.org. As AI technologies advance, the power demands of servers can increase by **four to five times** compared to traditional systems, further straining energy resources applus.com.

Water Usage

In addition to energy, data centers require vast amounts of water for cooling operations. Some facilities can consume over **20 billion liters of water annually** for cooling and electricity generation aaas.org. The rapid growth of AI is expected to lead to water withdrawals that could exceed the annual household water usage of millions of people Cornell University. This water is often lost through evaporation, contributing to local water scarcity issues.

Carbon Emissions

The environmental footprint of data centers is also marked by significant carbon emissions. Current trends in AI growth could lead to an increase of **24 to 44 million metric tons of CO2** emissions annually by 2030, equivalent to adding millions of cars to the roads Cornell University. If data centers rely on fossil fuels for electricity, they contribute to climate change and air pollution, impacting both local and global environments aaas.org.

Mitigation Strategies

To address these environmental challenges, many organizations are exploring **decarbonization strategies** and the use of renewable energy sources. Some major tech companies have committed to achieving carbon neutrality, but smaller facilities often lack the resources to implement similar measures aaas.org. Research

is ongoing to develop sustainable practices that can significantly reduce the environmental impact of data centers, including smarter siting, operational efficiency improvements, and faster grid
decarbonization

MIT - Massachusetts Institute of Technology+1.

Conclusion

The environmental impacts of data centers are profound and multifaceted, encompassing energy consumption, water usage, and carbon emissions. As the demand for digital services continues to rise, it is crucial for the industry to adopt sustainable practices to mitigate these effects and contribute to a more sustainable future.

Read less

Data Center Overview

Data centers have a global footprint of 67.7GW. This is a growth of 36 percent over the past two years. Data centers now consume 2 percent of the world's electricity, up from 1.7 percent in 2024 and 1.9 percent in mid-year 2025.

The US is by far the world's largest data center location, with 43 percent of global consumption. Data centers consume 29.2GW of electricity in the US, and now consume 6 percent of the nation's electricity.

The world's other Top Five data center consumers are:

- China, at 8.5GW, consuming 0.8 percent of nation's electricity.
- Germany, at 5.5GW, consuming 9.5 percent of the nation's electricity.
- UK at 2.0GW, consuming 5.8 percent of the nation's electricity.
- Japan at 1.7GW, consuming 1.5 percent of the nation's electricity.

IDCA's data for the US and several other countries are higher than some traditional estimates because of the presence of small, room-sized data centers in use by SMEs that must be counted. These facilities are numerous and inefficient compared to the large, building-sized facilities and campuses.

The continued use of smaller data centers by Mittelstand (family-owned mid-sized) companies in Germany increases that nation's estimate when compared to some other results as well. Similar considerations apply throughout a variety of nations in IDCA's analysis.

Additionally, IDCA research finds that as much as 13 percent of data center consumption in the United States comes from unused, "zombie" cloud containers and services. This wasted consumption totals in excess of 3GW, comparable to the energy produced by three nuclear reactors. This sort of inefficiency is presumed to be apparent throughout the world, with inefficiencies increasing as the percentage of cloud computing rises.

IDCA's China analysis, on the other hand, produces data that shows a smaller footprint than some traditional estimates. Neither China's government nor its major vendors or utilities are transparent in providing energy-use statistics. There is substantial unused capacity in China, and diverse global opinion of how much data center consumption exists. IDCA has revised its China estimate upwards by a couple of gigawatts when compared to analyses in 2025, and continues to seek optimal reporting of the nation's data center footprint.

China is an obvious sleeping giant within the world's data center sphere. IDCA's research shows the nation still devotes less than 1 percent of its energy grid to data centers. The nation now produces almost twice as much electricity overall as the United States, so has the potential to become a truly daunting consumer of data center services.

Terms of Use and Disclaimer

Acknowledgment

This comprehensive report is created through the efforts of seasoned subject matter experts of IDCA who have dedicated their time and expertise in hopes of bringing transparency and light for resourceful measures to the Digital Economies of the world:

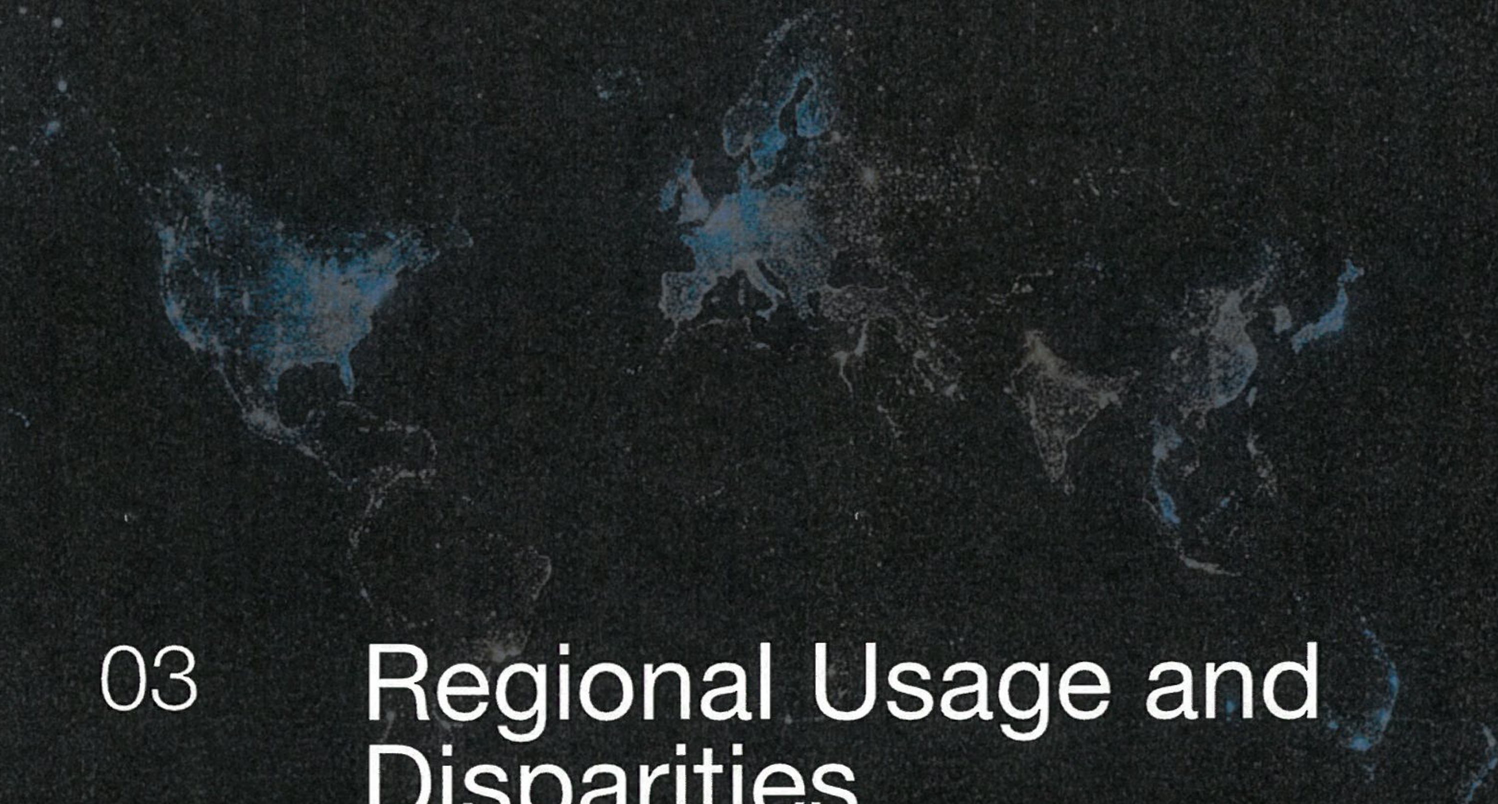
MEHDI PARYAVI
Chairman & CEO

PETER GROSS
Data Center Counsel

ROGER STRUKHOFF
Chief Research Officer

Mark Thiele
Entrepreneurship Counsel

Greg Sly
Infrastructure Counsel



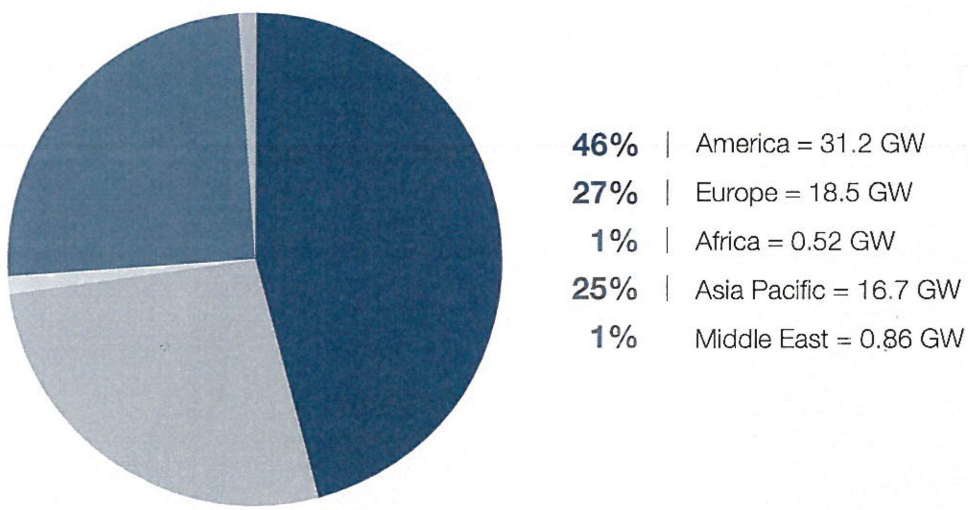
03 Regional Usage and Disparities

Regional Usage and Disparities

With the above advisories in mind, IDCA shows a regional breakdown of data center usage as shown below.

FIGURE 1.

Regional Breakdown of Data Center Footprint

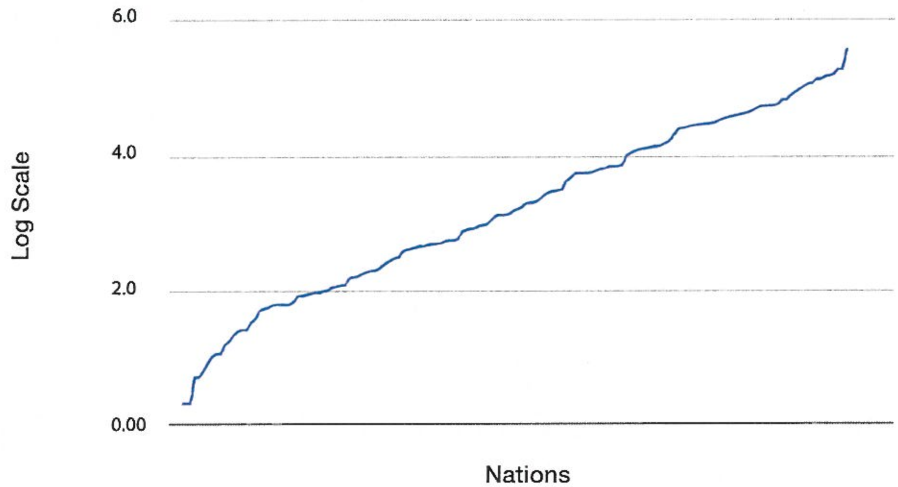


Source: IDCA

The figures above reflect a continuing, massive disparity in data center usage throughout the world. IDCA finds a difference of as much as 100,000X in the density of server usage – that is, the per capita presence of data servers and data centers varies by as much as 100,000 times.

FIGURE 2.

Data Server Density



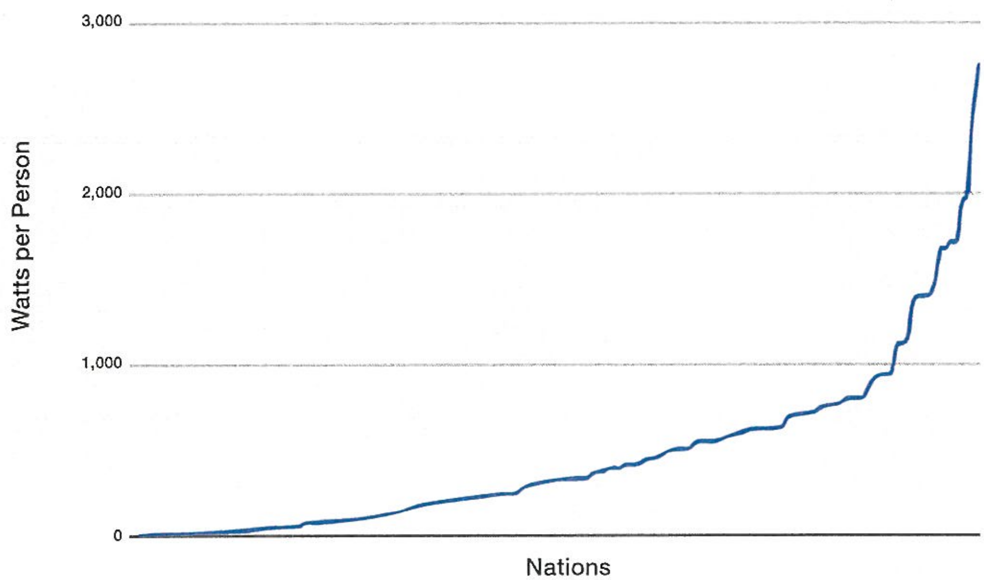
Source: IDCA

Note: This chart is plotted on a log10 scale. Each horizontal unit increase denotes a 10X increase in density. The five-unit disparity between the US (to the right of the chart) and least-developed infrastructures (on the left) thus represents a gap of 100,000 times in the number of servers per population.

The above disparity is largely a function of another disparity, that of the world's electricity grids, which vary 100X in their respective strength. This large disparity continues to exist at the same time that data centers and related digital infrastructure are consuming and increasing proportion of those grids.

FIGURE 3.

Disparity of National Electricity Grids



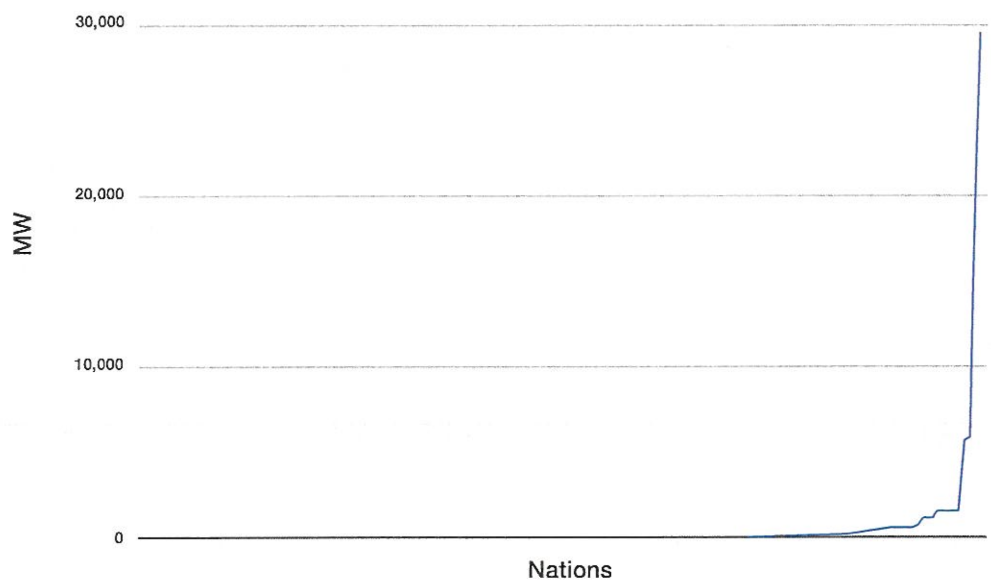
Source: IDCA Research

Data center consumption can also be viewed with respect to the amount of a nation's electricity it consumes. As noted above, the world as a whole has passed the 2-percent barrier in electricity consumption by its data centers.

This average is not troubling in and of itself; however, it masks a very wide disparity of electricity use from nation to nation, and obscures the very real constraints being experienced by a variety of nations across income tiers.

FIGURE 4.

Data Center Consumption



Source: IDCA Research

On the one hand, data centers in Singapore and the European nations of Netherlands, Germany, and Switzerland have exceeded the 9-percent consumption level. Data center consumption is provoking political and societal angst in those nations. It is also doing so in a few countries that are approaching this level in other European nations, including Ireland, Denmark, Lithuania, Estonia, and the UK.

The situation is mitigated somewhat in Singapore by strong data center development in the border regions of Johor, Malaysia and Batam, Indonesia. But there is no such immediate relief in Europe. Over the longer term, the large nation of France and smaller nations such as Slovakia and Slovenia, all of which have very strong electricity grids driven by nuclear energy, may serve to reduce some constraints. But this process would be a political process and presumably take time.

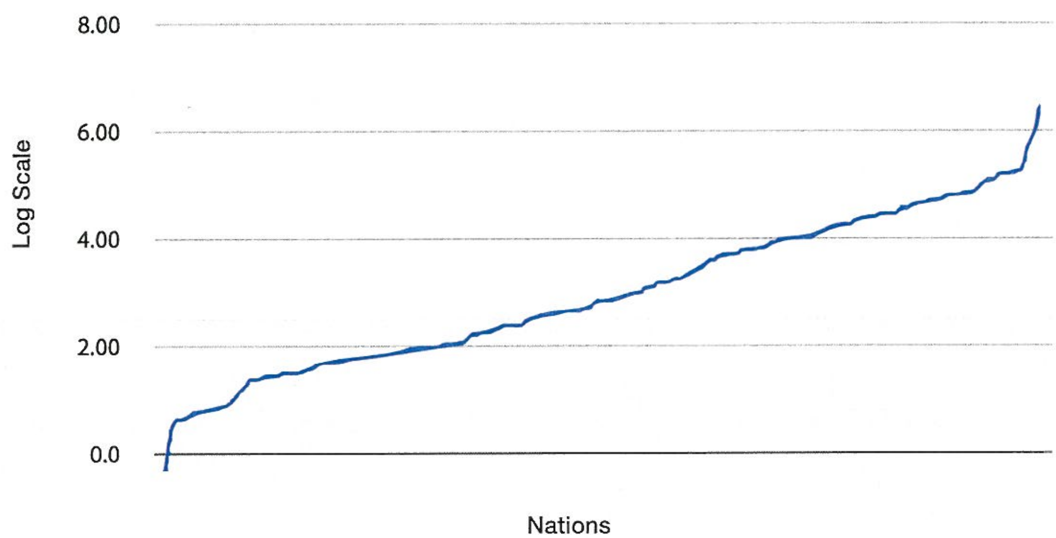
In the US, which has now reached the 6-percent level, community pushback over data center consumption has been seen in several regions. Data center developers in Northern Virginia, many parts of the Midwest, Texas, and California are finding multi-year delays in the ability to provision power for new data centers.

On the other hand, the lack of sufficient energy to meet basic needs throttles data center plans throughout the developing world. IDCA finds that more than 70 countries devote less than 0.1% of their power to data centers, with more than 30 countries devoting less than 0.01% to data centers.

Societal improvement in the 21st century cannot occur without a healthy digital infrastructure, and data centers are the obvious foundation of that infrastructure. Plans by some developers to ameliorate the situation in developing countries by providing their own power often involve the use of fossil fuels, and surely do not represent a long-term development view.

FIGURE 5.

Data Center Consumption (log10)

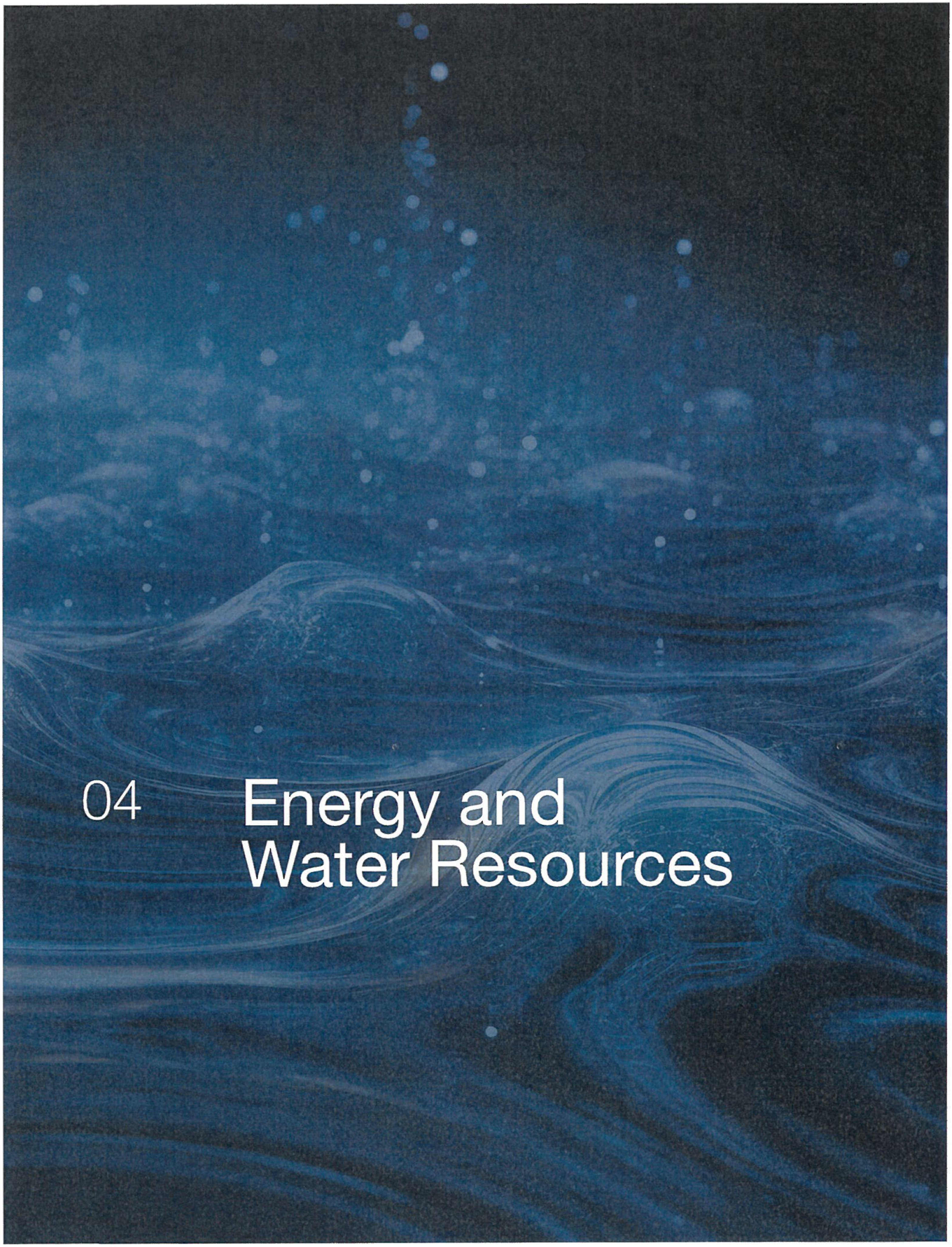


Source: IDCA Research

Data Center Consumption and Expansion

IDCA has analyzed each of the world's nations with respect to its current data center consumption and its potential for expansion. The analysis involves the use of an optimized model, in which Socioeconomic conditions are ideal and data centers consume 6.25% of an electricity grid that is strong and powered 95% by renewable energy.

The optimized model takes into account the challenges facing developed nations – where energy use by data centers is straining grids – and developing nations, where current grids are insufficient to meet basic needs, let alone aggressive data center buildouts.



04 Energy and Water Resources

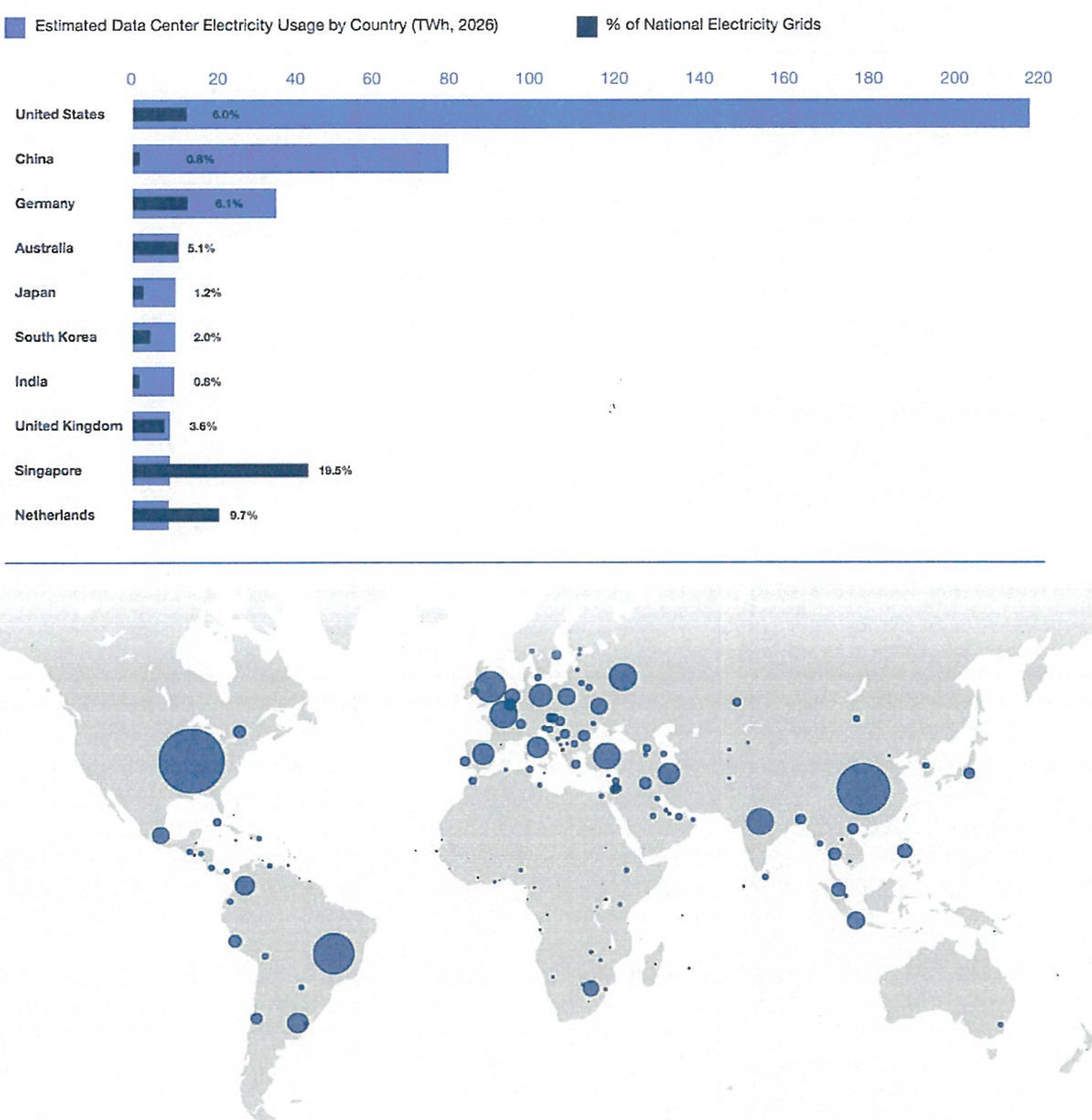
Energy

IDCA advocates the use of sustainable energy for development of new data center facilities and hubs. Sustainable energy includes the renewable sources of hydroelectric, wind, and solar power. It also includes the use of geothermal energy where it's present, as well as the use of biomass. Green hydrogen represents another renewable resource, one which may be developed more fully in future years. Finally, nuclear energy is included as a sustainable source of energy. A full review of the world's current use of sustainable energy is found in the upcoming IDCA Energy Usage Report.

The following figure provides an overview of global energy use and its consumption by data centers.

FIGURE 7.

World's Top 10 Data Center Grids (TWh)



Source: IDCA Research

FIGURE 8.

World's Top 10 Data Center Consumption Levels (% of national grids)

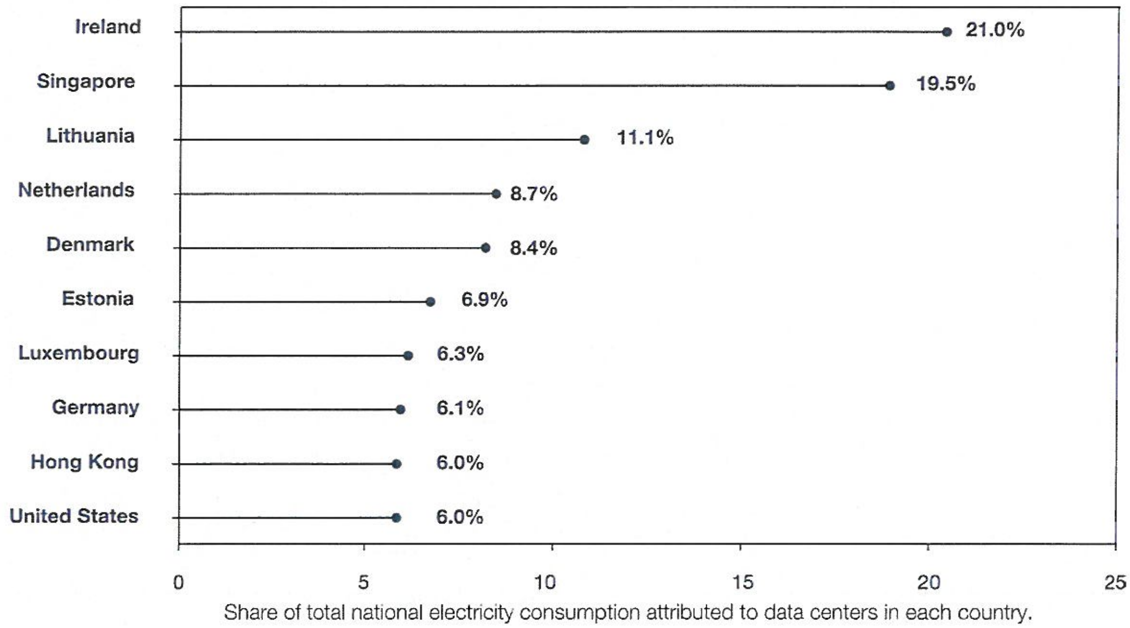
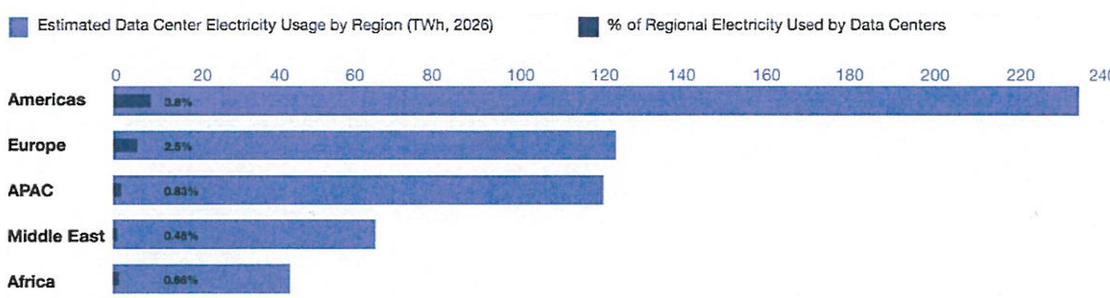


FIGURE 9.

Data Center Consumption by Region (TWH, % of national grids)



Source: IDCA Research

Besides the big picture outlined above, there are also obvious local concerns about the total amount of data center electricity consumption. The world average is now 2.0 percent, which is not an alarming figure by itself. However, several developed nations have crossed the 5-percent level, sparking societal and political concerns.

IDCA works with an optimized national model to determine ideal consumption levels for each nation of the world. For many developing nations, construction of any new data centers is contingent on also developing stronger electricity grids and more robust fiber-optic networks to deliver data to and from data centers. For developed nations, which by definition have enough electricity to serve their populations very consistently (except in rate peak-emergency situations), IDCA finds that a cap of 6.25 percent of national electricity consumption is an effective cap on data center growth. This is not a prescribed or engineering number. Rather, it is the point at which the correlation is very strong between data center consumption and political actions such as usage caps, preferred operating hours, community pushback that results in project cancellation, and prohibition on new projects.

The United States has seen a recent flurry in new data centers coming into operation, resulting in an average usage of almost 6 percent. As noted earlier in this report, as much as 13 percent of this usage (or 0.78 percent of the total) supports "zombie" applications that were instituted, discarded, and never turned off. In any case, the US with its vastness has several areas in which data center usage is much

Smaller nations that have exceeded the 6.25 percent level and which are experiencing a no-growth era, or close to it, include Germany, Netherlands, Ireland, Estonia, and Lithuania. Singapore in Asia is in a similar situation. South Korea provides a special case, with an overall usage of well below 6.25 percent but an extreme concentration in metropolitan Seoul. This has prompted the government to authorize new data centers only outside of the metro area. A full review of the world's current use of sustainable energy is found in the upcoming IDCA Energy Usage Report.

The table below provides a list of the 27 nations in the world that have a hypothetical capacity to add at least 1 gigawatt of new data center footprint without developing new electricity resources. These calculations were made by comparing current usage against the optimized 6.25 percent consumption level.

China is by far the largest presence in this table, but it is very difficult to get a precise number of what the opaque Chinese government will actually do in building out its footprint in coming years. For the other nations, this table should be viewed with the knowledge that any of these nations may not have the societal or political desire to build out to their capacity, and that every project must be built with respect to local power conditions and other local considerations.

Data Center Headroom

Country	Headroom (GW)
China	58.9
India	12.7
Russia	6.7
Japan	5.5
Brazil	4.5
Canada	3.6
South Korea	3.2
Saudi Arabia	2.6
Iran	2.4
Mexico	2.4
Indonesia	2.1
Taiwan	2.0
Turkey	1.9
Vietnam	1.7
France	1.7
Egypt	1.6
Thailand	1.5
Italy	1.5
Australia	1.4
Spain	1.3
USA	1.2
Pakistan	1.2
South Africa	1.2
UAE	1.2
Iraq	1.1
Argentina	1.0
Colombia	1.0
Group Total	**127.1**

Water

The convergence of cloud expansion, edge computing, and artificial intelligence has created a genuine shock to global electricity and water systems. Technology demand is doubling rapidly, and the underlying pipes, wires, and power plants were not designed for this scale of progress. Historically, large data centers were centralized in major metropolitan hubs with built-in utility redundancy. Today, edge computing is pushing infrastructure into smaller communities, resulting in power and water demand spreading into regions that never expected to host this kind of industrial infrastructure.

Small towns and municipalities are getting hit the hardest. In regions like the suburbs of Houston, or Columbus, Ohio, water systems are already stretched thin juggling the surging demands of residents, agriculture, and incoming data centers all at once. In extreme cases, such as a proposed facility in Marfa, Texas, a single site could potentially use more groundwater than the entire town currently consumes.

The Reality of Traditional Cooling Methods

To understand the water challenge, stakeholders must recognize that the data center market is now sharply bifurcated. The vast majority of existing traditional cloud and enterprise facilities rely on older, less efficient cooling architectures like water-cooled chillers or evaporative cooling towers. In these legacy designs, water is evaporated to remove the latent heat of vaporization from the air, and that evaporated water is never recovered.

These facilities can be massive water consumers. A single traditional hyperscale facility can burn through millions of gallons of water each day just for cooling. Because these sites frequently rely on municipal potable drinking water, a single large data center can consume as much water as 6,500 local households or the equivalent of maintaining 41 golf courses. While engineers historically accepted this design because using evaporative water significantly lowered electricity costs, communities in arid climates are now actively pushing back against the depletion of their limited resources.

Demystifying AI and Liquid Cooling

Because of the massive water footprint of older cloud facilities, a public narrative has emerged warning that the incoming wave of "AI Factories" will completely drain local water supplies. Factually speaking, this is a misconception. Next-generation gigascale AI facilities operate at such extreme chip densities that traditional air and evaporative cooling are physically incapable of removing the heat.

Instead, modern AI centers rely on Direct-to-Chip liquid cooling. These function as entirely closed-loop systems—much like an automobile radiator. Hot liquid is pumped directly from the chips to a massive dry radiator outside, where fans cool the liquid before it is recirculated back to the servers. Because the internal orifices of these cooling plates are as small as one micron, the systems are filled with highly purified water mixed with inhibited propylene glycol to prevent bacterial growth and corrosion.

Most importantly, once this closed-loop system is initially filled, these high-performance AI centers consume minimal water during daily operations. When the media reports that a new AI facility has a permit for "5 million gallons of water a day," it is almost always referring to a one-time municipal permit required to initially fill the loop, maintain fire suppression tanks, or suppress dust during the construction phase. Once operational, the cooling loop is sealed, meaning the industry's most advanced AI data centers do not contribute to ongoing local water scarcity.

Putting Data Center Water Use in Context

While data centers can severely impact local municipalities, their global water footprint is remarkably small compared to other sectors. On a per-square-foot basis, modern data centers consume significantly less water than standard commercial buildings like hospitals, senior care facilities, or courthouses, and vastly less than the forest products industry requires just to manufacture toilet paper. Furthermore, the "indirect" water utilized by utility companies to generate data center electricity has steadily declined over the past 25 years—a reduction that continues to accelerate as national grids transition toward wind and solar generation, which consume minimal water.

The IDCA Water Stress Index

The rapid expansion of digital infrastructure has placed unprecedented strain on local water supplies, sparking fierce community pushback against new facilities. While legacy data centers rely on water-intensive evaporative cooling, the fear that incoming "AI Factories" will drain local aquifers is largely a misconception. Next-generation gigascale AI facilities utilize closed-loop, direct-to-chip liquid cooling. Once initially filled, these high-performance systems consume minimal water during daily operations.

Despite these innovations, resource scarcity remains a severe threat to long-term digital growth. To provide actionable guidance, the IDCA developed a proprietary Water Stress Index to evaluate a nation's hypothetical water availability alongside the socioeconomic and political factors that impede development. Because local water stress dictates project viability, this 0-to-100 metric is integrated directly into the overarching IDCA Sigma Index to measure a nation's true readiness for sustainable digital expansion.

Water Stress Index

Country	Water Stress (0-100)
Bahrain	100
Belize	86
Libya	80
Kuwait	80
South Sudan	73
Singapore	73
Saudi Arabia	72
Bahamas	71
Yemen	71
Qatar	70
Algeria	70
Seychelles	65
Turkmenistan	65
Maldives	65
Gambia	65
Oman	64
Lebanon	64
Eritrea	64
Bangladesh	63
Tunisia	62
Burundi	62
Trinidad	62
Haiti	61
UAE	60
Hong Kong	60
Djibouti	60

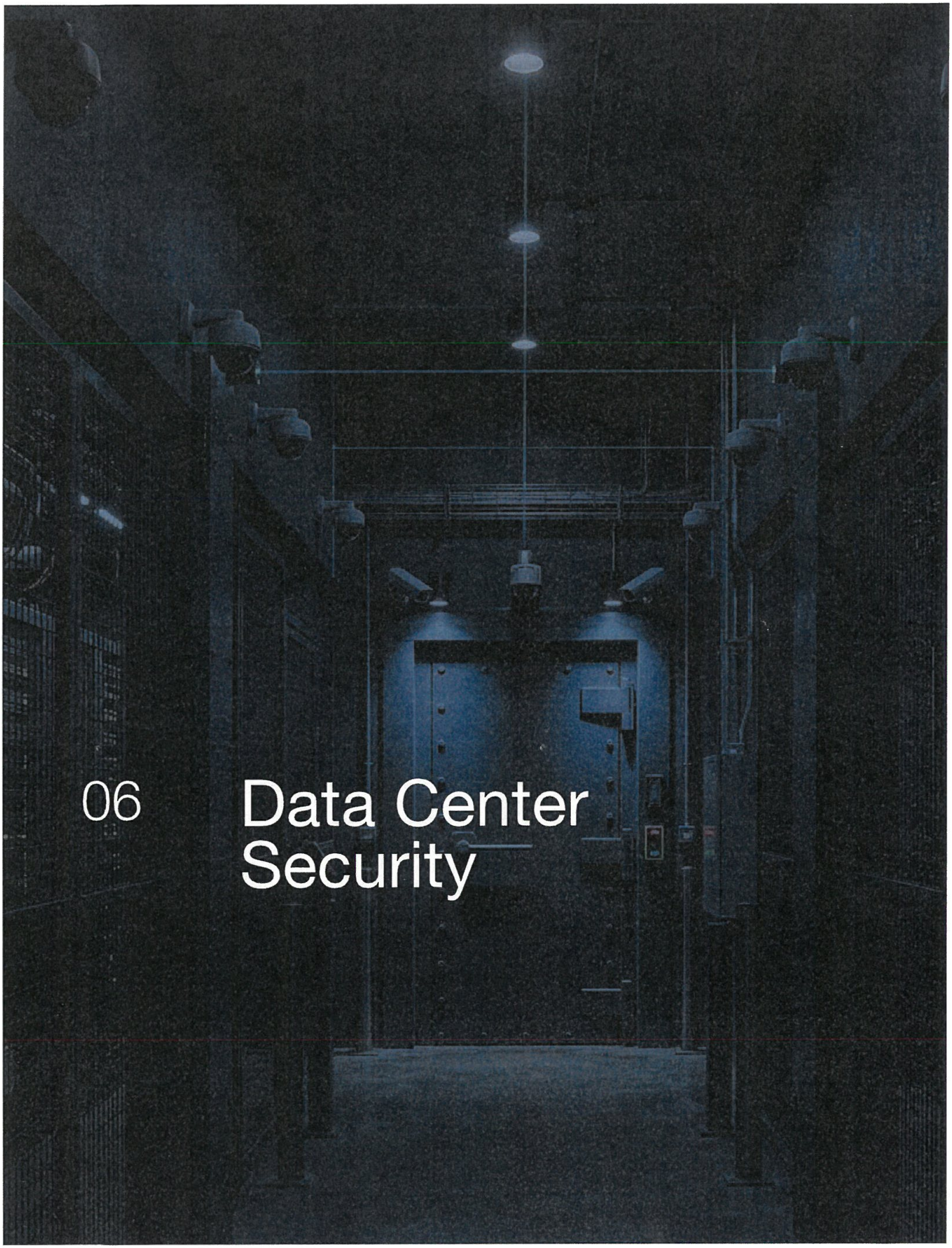
Water Stress Index

Country	Water Stress (0-100)
Iran	59
Malta	58
Jordan	58
Israel	56
Sudan	56
South Africa	55
Timor-Leste	55
Egypt	55
Cyprus	55
Nigeria	54
Azerbaijan	54
Uzbekistan	53
Niger	53
Dominican	53
Zimbabwe	53
Kenya	53
Moldova	52
Botswana	52
Morocco	51
Rest of World	51
Pakistan	51
Malawi	51
Togo	50
Philippines	50
Uganda	49
Myanmar	49
India	49
Iraq	48
Eq Guinea	47
N. Macedonia	47
Rwanda	47
Mexico	46
Kyrgyzstan	46
Romania	46
Poland	46
El Salvador	45
Ethiopia	45
Czechia	45
Tanzania	45
Cote D'Ivoire	45

Water Stress Index	
Country	Water Stress (0-100)
Ukraine	44
China	44
Lithuania	44
Tajikistan	43
Sri Lanka	43
Jamaica	43
Turkey	43
Guatemala	43
Venezuela	43
Kazakhstan	43
Ghana	42
Congo, Dem	42
Taiwan	42
Thailand	42
Senegal	41
Mozambique	41
Honduras	41
Mongolia	41
South Korea	40
Panama	40
Angola	40
Madagascar	39
Germany	39
Italy	39
Congo	39
Cameroon	38
Mali	38
Armenia	38
Indonesia	38
Cambodia	38
Netherlands	37
Nepal	37
Mauritius	37
Denmark	37
Belarus	37
Bosnia	37
Zambia	36
Vietnam	36
Nicaragua	36
Hungary	35

Water Stress Index	
Country	Water Stress (0-100)
Slovakia	35
Spain	35
Ireland	35
Luxembourg	34
UK	34
CAR	33
Greece	33
Bulgaria	33
Russia	33
Serbia	33
Estonia	32
Belgium	32
USA	32
Japan	31
Albania	31
France	31
Guyana	31
Argentina	30
Colombia	29
Bolivia	29
Portugal	28
Gabon	28
Laos	28
Paraguay	28
Sierra Leone	27
El Salvador	27
Croatia	26
Latvia	26
Montenegro	25
Australia	25
Brazil	24
Switzerland	23
Puerto Rico	23
Slovenia	23
Georgia	23
Ecuador	22
Malaysia	22
Peru	21
Namibia	20
Austria	19

Water Stress Index	
Country	Water Stress (0-100)
Costa Rica	14
Finland	10
Chile	9
Sweden	6
Uruguay	3
Bhutan	3
Canada	2
Iceland	1
New Zealand	0
Norway	0



06 Data Center Security

Data Center Security

Historically, physical security was sometimes subjugated to cybersecurity rules, but today, industry leaders recognize both domains as co-equal pillars of a unified, comprehensive security strategy. The modern reality is that the cloud has a physical street address, and adversaries now utilize highly asymmetric tactics that completely blur the lines between digital and kinetic attacks. Whether a threat actor deploys a destructive malware payload to breach network defenses or utilizes an armed drone to physically disable a critical power substation, the ultimate goal is identical: to completely destroy or disrupt the data center's core mission execution.

Because a vulnerability in one domain can easily be leveraged to compromise the other, mitigating these threats requires security planning to begin at the highest organizational level rather than existing in isolated departmental silos. The financial and reputational stakes are unprecedented, with the average successful cyberattack costing enterprises around $20 million and operational downtime rapidly compounding those losses. True resilience demands that operators maintain absolute vigilance every second of the day, recognizing that threat actors constantly seek out the softest targets across both the network and the facility. The following sections detail the specific defensive architectures, normative standards, and operational protocols required to harden both the digital and physical perimeters of the modern data center.

Cybersecurity: Defending the Digital Ecosystem

The average successful cyberattack already costs G2000 companies around $20 million, according to IDCA research. But cybersecurity is now about much more than money – many attacks are catastrophic, designed to put companies and even entire nations out of business. Yesterday's potential for deepfake identities and related fraud have morphed into today's nightmare stew of AI slop and pernicious, AI-driven cybercrime and terrorism. Furthermore, the attacks on data centers – now viewed as critical infrastructure – in the Middle East have shocked data center operators and customers with the specter of breached physical security. As a result, cybersecurity is now twinned with physical security as part-and-parcel of a unified, comprehensive security strategy.

Normative Standards and Organizational Governance

To implement basic defensive measures, the 800-53 NIST standards have long been followed as the normative cybersecurity standards in the United States. They are continually updated and focus on protecting organizational operations and assets (and people) from hostile attacks, human errors, natural disasters, structural failures, hostile state actors, and general privacy risks. The NIST standards are "derived from mission and business needs, laws, executive orders, directives, regulations, policies, standards, and guidelines," according to the agency. Additionally, specialized frameworks like NIST 800-30 are highly utilized for risk assessment methodologies, particularly within the healthcare industry for HIPAA compliance, alongside other evaluative frameworks like OCTAVE.

Europe and much of the rest of the world follows the ISO/IEC 27000 standards. Identity management, access control, and threat identification remain at the heart of this approach. ISO has noted that "most organizations stick to the most basic form of threat intelligence," and urge them to be more curious about the benefits of incorporating a more rigorous cybersecurity environment. Ultimately, security must begin at the organizational level. Developing security systems based on ISO-27001 requires continuous personnel training and education, the strict enforcement of policies, weekly reviews of security incidents, and both periodic internal and independent external audits to ensure compliance.

The Evolving Cloud and Malware Threat Landscape

Human-origin network risks, including hacker penetration, Denial of Service (DoS) network saturation, and malware, represent the most active daily threats to data center stability. Malware is actively targeting the cloud, shifting toward cloud control planes, SaaS apps, containers, Kubernetes, and even core hypervisors. If left unchecked, cloud-targeted malware compromises the vast computing power of the infrastructure, applications, and underlying data.

Modern threat reports highlight severe vulnerabilities, including ransomware that targets ESXi servers utilizing double-extortion tactics (data theft combined with leaks). Threat actors are increasingly deploying destructive wipers, supply-chain malware through poisoned CI/CD pipelines, and web shells used for persistent control and data exfiltration staging. Furthermore, cryptomining malware is frequently used to burn massive amounts of CPU and GPU resources, often masking deeper network intrusions. Once inside, attackers aggressively target Backup and NAS systems—sabotaging snapshots and encrypting files via SMB/NFS to systematically remove a facility's recovery paths.

Defense-in-Depth and Infrastructure Hardening

Protecting the IT infrastructure requires a robust Defense-in-Depth strategy spanning external application security, network edge security, and inter-application/intra-application security. The foundational perimeter must be secured using firewalls, encryption, strict authentication protocols, and advanced Intrusion Detection and Prevention Systems (IDS/IPS). These IDS/IPS solutions, whether host-based (HIDS/HIPS) or cloud-native network appliances, are vital for recognizing and restricting malicious inbound traffic and analyzing system anomalies.

At the endpoint and workload level, operators must deploy modern Endpoint Detection and Response (EDR/XDR) platforms alongside Cloud Workload Protection Platforms (CWPP) to secure high-density VMs and Kubernetes clusters. To manage this massive influx of security data, data centers rely on Security Information and Event Management (SIEM) systems to provide centralized visibility and rapid threat response. Finally, these defensive architectures must never remain static; they require continuous validation through periodic internal and external vulnerability scanning, as well as aggressive Penetration Testing (Pen Tests) to discover and close exploits before hostile state actors or cybercriminals can leverage them.

Physical Security: Hardening the Modern Data Center

The attacks on data centers – now viewed as critical infrastructure – in the Middle East have shocked data center operators and customers with the specter of breached physical security. Cybersecurity is now twinned with physical security as part-and-parcel of a unified, comprehensive security strategy. As industry experts note, these are co-equal pillars; the cloud has a physical street address, and that address is increasingly vulnerable to kinetic attacks. Modern conflicts, such as the war in Ukraine, have demonstrated how drone warfare has fundamentally changed the battlefield, leaving traditional infrastructure highly exposed.

Core Principles and Perimeter Defense

All architectural, operational, and physical security measures must be designed to achieve one or more of the following: **Delay, Deter, Detect, Decide, and Act**. Physical security begins at the organizational level and heavily relies on establishing a secure perimeter. Best practices dictate installing a fence or wall at least 20 feet from the building, utilizing gate and guard kiosks, ensuring comprehensive lighting, and strictly avoiding any recognition signs that advertise the facility's purpose to the public.

Internal Access and Surveillance

Inside the facility, the security room should be located at or adjacent to the main personnel entrance, though the visitor sign-in area must be physically separated from the actual security operations. For highly critical facilities, this security room must be hardened with 16 mm (5/8 in) plywood, feature 180-degree peepholes, and maintain a fire separation of at least 2 hours from the rest of the data center. Strict access control is mandatory, requiring manual or automated Access Control List (ACL) management, detailed access logs, and RFID tracking for personnel, vendors, and equipment. Surveillance must include IP-based CCTV cameras running on UPS power, streaming directly to Network Video Recorders (NVR) with automatic cloud backups to maintain historical compliance.

Defending Against Modern Kinetic Threats

Facilities may already be resistant to accidental and other amateurish drone threats, but the recent war activity involves armed drones delivering missiles to breach the data centers. Attackers constantly seek "soft targets," which frequently include unprotected fiber vaults and rural electrical substations.

Immediate counter-drone measures include more sophisticated detection systems that can identify threats from further distances and trigger internal alarms. RF jamming and GPS spoofing can be brought to bear if facilities now operate in a consistently dangerous area, with budgetary concerns an obvious concern to be weighed against the effectiveness of these measures. Military installations can resort to interceptor drones and other weaponry, although this is not viable for most of the world's data centers. (For context, top-tier defense data centers are often built as underground bunkers capable of surviving a nuclear blast and feature extreme shielding against EMPs and biological weapons).

Cost/Benefit Analysis and Infrastructure Hardening

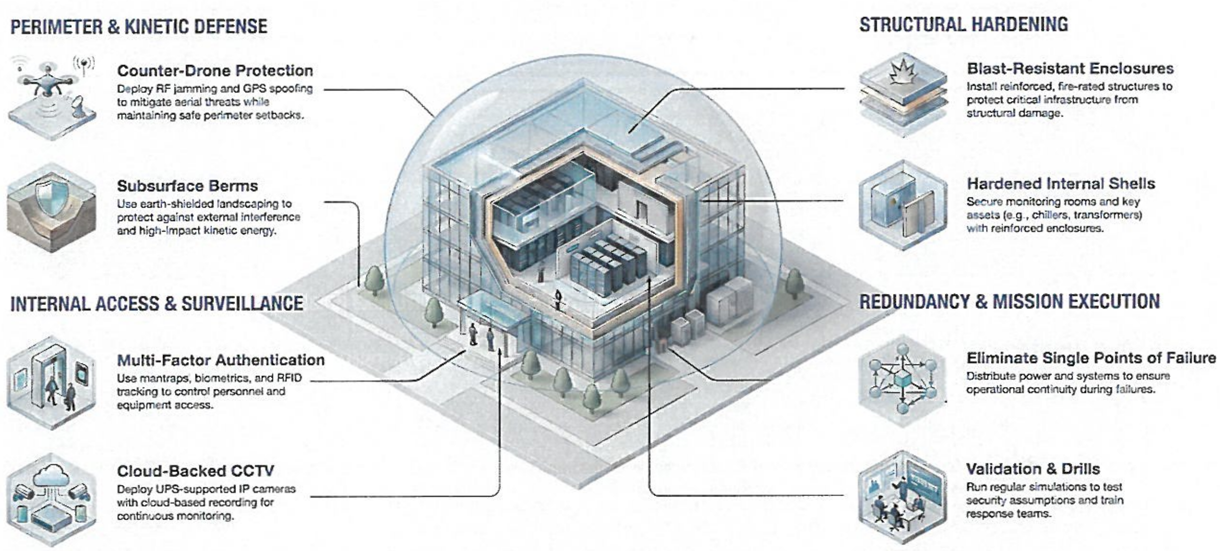
At a more modest budget and threat-aversion level, operators need to do a cost/benefits analysis regarding installing reinforced, blast-resistant roofs and facades, hardened enclosures for transformers, chillers, and fuel systems, and consider subsurface earth-shielded berms for certain infrastructure. To justify these investments, leadership must redefine the ROI of physical security by linking it directly to the safety of the workforce and the catastrophic brand impact of a business continuity failure.

Redundancy, Topology, and Mission Execution

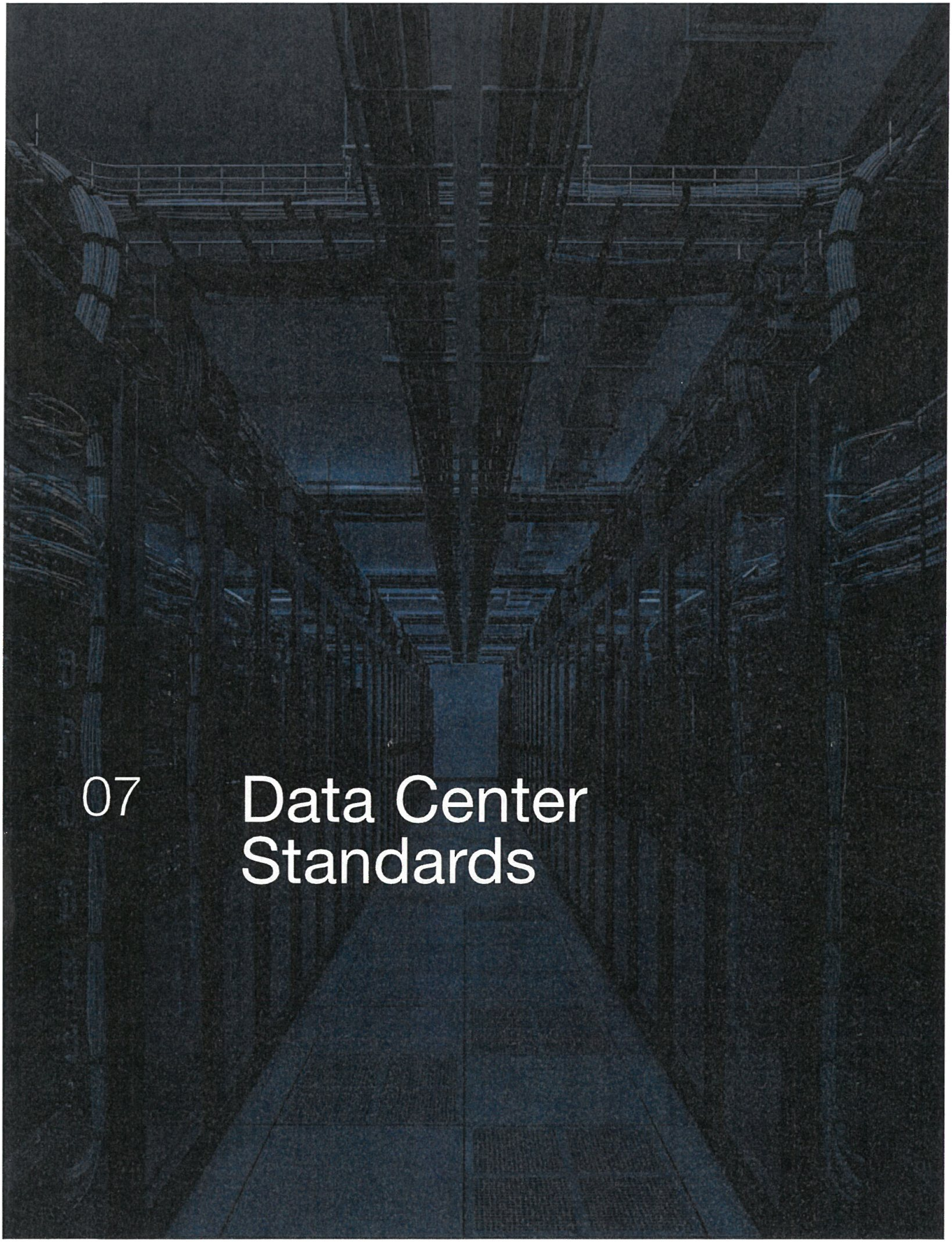
Redundancy takes on even more importance, with respect to electrical substations, cabling (especially exposed cabling), backup systems, and even walkways and service roads within a large facilities. But the best measure operators can take is to distribute resources as possible within the facility to lessen single points of failure. Large operators will have robust topology already in place, to distribute loads among data centers within a particular campus or wider region.

Ultimately, true physical security requires a focus on "mission execution"—identifying the absolutely essential elements required to keep an application running, and rigorously conducting tabletop exercises, drills, and continuous reviews of those continuity plans.

FIGURE 10. Hardened Data Center Security Architecture



Source: IDCA Research



07 Data Center Standards

Data Center Standards

There are a number of underlying, fundamental standards to be followed by data center builders and operators to reduce business risk and align performance with defined business requirements. Historically, many organizations have provided foundational guidelines for physical design, security, and operations. Today, modern frameworks like the IDCA's Infinity Paradigm leverage these foundational standards (such as those from ASHRAE, ISO, NFPA, and TIA) as a baseline, extending them into a comprehensive, application-centric model that governs the entire design, installation, and operation of a data center. The most important standards and frameworks come from the organizations shown below, and all should be familiar to experienced data center owners, operators, and managers.

ASHRAE

ASHRAE TC 9.9 Thermal Guidelines for Data Processing Environments is a core reference for allowable temperature/humidity ranges, recommended vs allowable operating environments, server inlet conditions, and reliability impact. The HVAC Handbook provides guidelines for cooling architectures, airflow management, and redundancy, and Standard 90.4 focuses on energy efficiency requirements for HVAC and power systems as well as PUE (Power Usage Effectiveness.)

BICSI

ANSI/BICSI 002 for data center design and implementation covers structured cabling, infrastructure lifecycle, security operations and commissioning.

IEEE

Relevant IEEE standards are the IEEE 3000 series for electrical series design, and IEEE 1100 for power quality and grounding of sensitive electronics.

ISO

International ISO standards include ISO/IEC 30134 for data center KPIs such as PUE, WUE, and CUE; and ISO/IEC 22237, which address data center infrastructure and facilities requirements.

NFPA

NFPA in the US covers fire protection. Relevant codes include NFPA 70 (National Electrical Code, NEC), NFPA 75 (Fire Protection of IT Equipment), and NFPA 76 (Telecommunications Facilities). They govern grounding and bonding, UPS/generator safety, and fire-rated cabling and power distribution.

TIA

TIA-942 is a common physical design standard for redundancy requirements, cable pathways, site space planning, resiliency, and topology.

IDCA: The Infinity Paradigm AE360® 2.0

The Infinity Paradigm is a comprehensive, cloud-inclusive, and application-centric standards framework designed by the International Data Center Authority (IDCA). It serves as a framework for visualizing and analyzing an entire information technology system as a holistic and dynamic model. Rather than just evaluating the physical building, it qualitatively and quantitatively illustrates the myriad of interdependencies of a data center's various supporting components to ensure the facility successfully delivers applications.

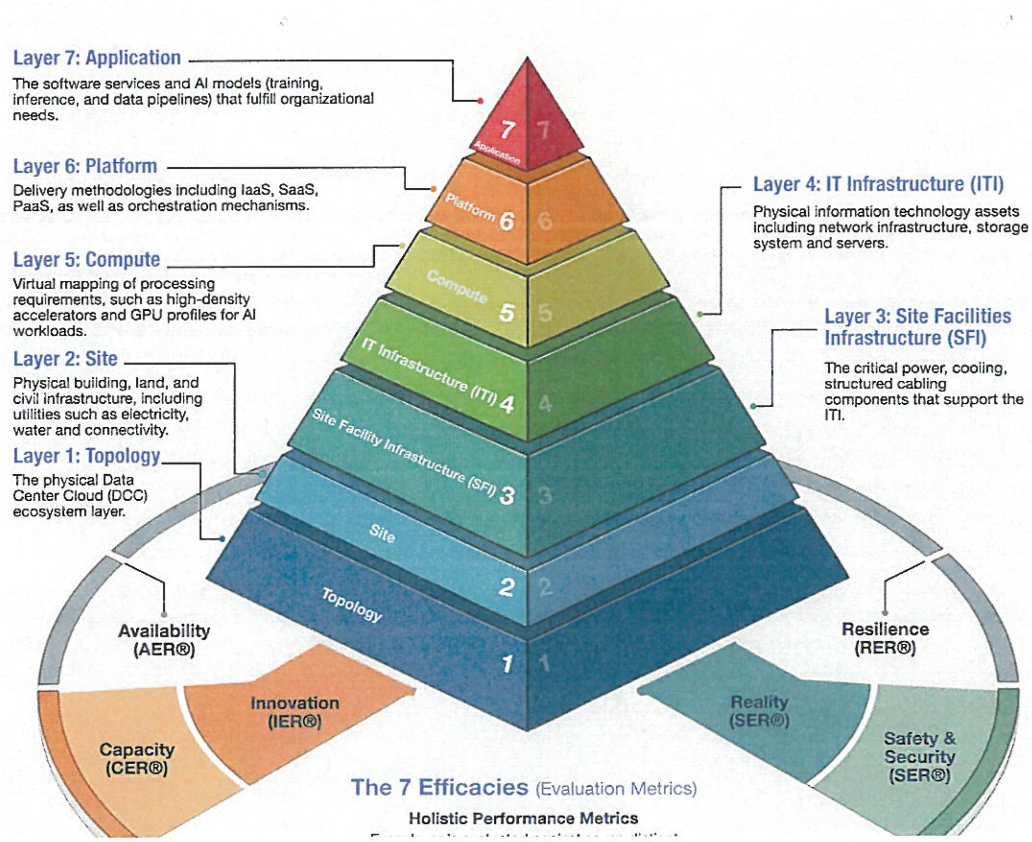
The Application Ecosystem (AE)®

At the core of this standard is the concept that the data center is simply the infrastructure supporting the Application Ecosystem (AE). Because an application lives, breathes, and interacts only in an ecosystem that supports its purpose, the framework is modeled as a 7-layer pyramid that stacks physical infrastructure at the bottom and logical infrastructure at the top. The 7 layers are:

1. **Topology:** The foundation of the pyramid, specifying the connectedness, physical location, interconnectivity, and interrelation of various data center sites.
2. **Site:** The independently operated space (land, building, architecture, utilities) housing the data center infrastructure.
3. **Site Facilities Infrastructure (SFI):** Facility-level components utilized for application delivery, including power, cooling, structured cabling, and safety/security.
4. **Information Technology Infrastructure (ITI):** The hardware bridging the facility to the compute layers.
5. **Compute:** The processing resources driving the applications.
6. **Platform:** The software frameworks and operating systems.
7. **Application:** The peak of the pyramid and the ultimate purpose of the facility.

FIGURE 11.

Application Ecosystem Layers



The 7 Core Efficacies

To understand the interdependencies of the 7 layers (Topology, Site, SFI, ITI, Compute, Platform, Application), AE360 evaluates them across seven specific metrics known as "Efficacies". These are measured both quantitatively and qualitatively:

1. **Availability Efficacy Rating (AER):** Evaluates the probability that a layer or system will operate at or above required performance levels during its service mission.
2. **Capacity Efficacy Rating (CER):** Measures the ability to handle data and compute volume.
3. **Operations Efficacy Rating (OER):** Assesses the human and automated management practices.
4. **Resilience Efficacy Rating (RER):** Measures the ability to withstand and quickly recover from failures or disruptions.
5. **Safety & Security Efficacy Rating (SER):** Evaluates physical and logical protections against threats.
6. **Efficiency Efficacy Rating (EER):** Assesses the optimal use of power, cooling, space, and compute resources.
7. **Innovation Efficacy Rating (IER):** Measures the effective creativity in deploying new technologies and modernizing processes.

The Efficacy Score Rating (ESR)®

The ultimate goal of evaluating these layers and efficacies is to generate the Efficacy Score Rating (ESR) ®. The ESR is an aggregated scoring metric calculated through a complex algorithm that combines the Grade Levels and efficacy results. It applies weightings to favor functions most critical to the application's survival.

The ESR provides a single, straightforward numerical output that represents the ecosystem's true ability to reliably and sustainably meet business requirements, making it the most reliable rating for comparing data center delivery abilities. It can be applied to existing facilities and projects at the concept stage.

Grade Levels (G4 to G0)

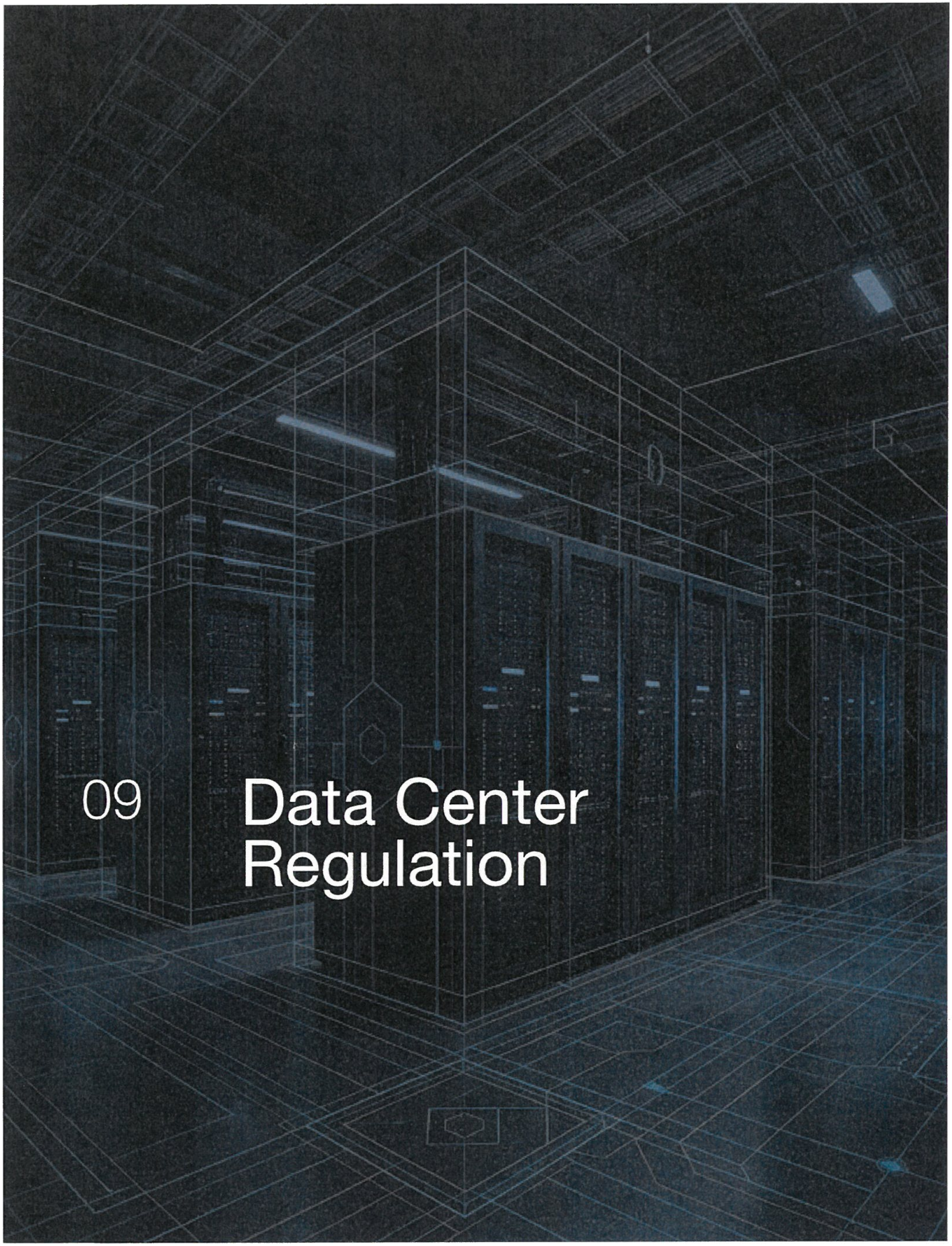
While a single data center node might be evaluated, AE360 recognizes that application survival often requires multiple nodes. The Grade Levels (Gs) reflect this reality:

1. **G4 (Minimum Acceptable):** The lowest level of achievement awarded to a layer. It represents the highest allowable levels of risk, inefficiencies, and chances of failure.
2. **G3 (Average):** Represents an average exposure to risks, vulnerabilities, and inefficiencies.
3. **G2 (Excellent):** Represents a low exposure to risks and high operational performance.
4. **G1 (Exceptional):** The lowest possible level of risk and highest performance achievable within a single Data Center Node (DCN).
5. **G0 (Infinity):** The ultimate state of excellence. Because G1 is the maximum score for a single physical building, achieving G0 requires ecosystem-level resilience—meaning redundancies and failovers are embedded at the Logical Layers (Application, Platform, Compute) and spread across a multi-node Topology. Only the most demanding mission-critical environments require a G0 rating.

Controls and Regulatory Compliance

To determine these grades, AE360 utilizes a massive set of controls. It does not replace foundational standards like ASHRAE, ISO, NIST, NFPA, or TIA; rather, it uses them as a baseline and fills in the gaps.

Furthermore, AE360 integrates regulatory and governance criteria directly into the infrastructure grading. A facility's grade can be impacted by its ability to physically and logically support compliance with Sarbanes-Oxley (SOX), PCI DSS, FISMA, HIPAA, and GDPR. For modern deployments, it even requires documented privacy constraints and ethical use policies for applications incorporating AI models and automated decision-making.



09 Data Center Regulation

Data Center Regulation

United States

At the moment the United States does not have anything that could be described as a "data center law." There are the usual environmental and emissions control concerns addressed by longstanding legislation, including the Clean Air Act, dating to 1963 and administrated by the Environmental Protection Agency (EPA). Its sibling, the Clean Water Act, passed in 1972, and administered at the federal and often state level by National Pollutant Discharge Elimination System (NPDES) regulations.

Consideration must also be made to the National Environmental Policy Act (NEPA), signed into law in 1970, which is coordinated by the Executive Office of the President of the United States through the Council on Environmental Quality. There is also the Federal Energy Regulatory Commission (FERC), an independent agency that was created during the very active era of 1970s legislation during the Nixon, Ford, and Carter administrations to address severe air and water quality problems throughout the US that were brought about by what was then a heavily industrialized economy. FERC overseas whole electricity markets and licenses hydropower projects, in addition to activities with transport of fossil fuels.

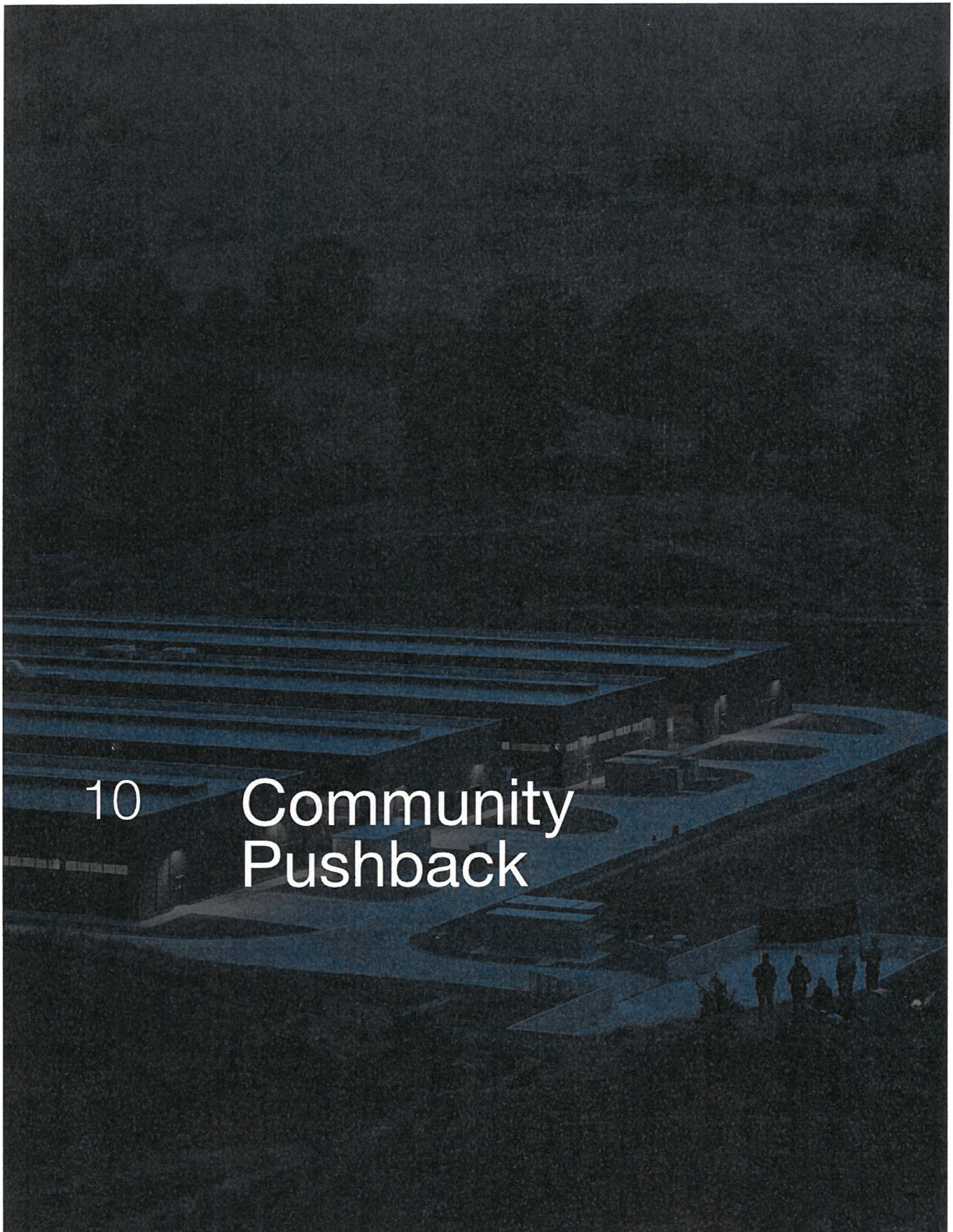
At the state level, there have been hundreds of recent bills regarding data centers introduced throughout several legislatures. Proposed areas of new regulation include prices for large-load electricity consumers, tax policies, renewable-energy mandates and energy efficiency standards, and controls over foreign ownership. One bill, in the state of Maine, has made it through the state legislature and is expected to become law. It will prohibit the construction of data centers larger than 20MW until 2027, and is positioned as a "pause" rather than a prohibition.

Outside of US

Outside of the US, the European Union has developed its Energy Efficiency Directive, which requires an honest reporting of data center energy performance. The more stringent ideas of mandatory sustainability reporting, efficiency benchmarks, waste heat reuse requirements, and AI regulation appear to be on the horizon.

There are, of course, nations that have already capped or banned the creation of new data centers, including Ireland, Singapore and South Korea. (There are also reportedly strict controls in China, but this nation is sui generis, that is, one of a kind in controlling and building out its data center and AI center footprint).

These caps and restrictions are driven by the single, simple issue of a lack of available electricity. Many regions in the US are also experiencing this phenomenon, with the result that county and local planners say they are simply unable to issue new data center permits. Developers are now waiting until 2032 for new land and power availability in Loudoun County, VA (the world's most densely packed data center hub), as well as several states in the Midwest (Illinois, Indiana, Wisconsin, Iowa, and Minnesota), according to recent IDCA research. Power scarcity is also present in California, effectively capping growth in several of its regions.



10 Community Pushback

Community Pushback

Community pushback against large new data centers has been reported over the past year in many parts of the United States. At the highest level, Senator Bernie Sanders from Vermont has urged a national "pause" on building new data centers, although his statements on the matter have received very little support. But the issue remains. It raises the question of whether legislation focused on data center construction, effects, limits, or even prohibitions will be considered in upcoming legislative sessions and future elections.

Over the past two years, however, most data center guidelines and potential pushback have occurred at the local level. A very recent proposed project in the St. Louis, MO metro led to the recall of city council members. A similar project in Joliet, IL received near-unanimous approval by local leaders after an extended public debate. A developer in the Los Angeles metro region recently withdraw plans for a project after vociferous community protests. And in the State of Maine, as noted above, a "pause" on data centers larger than 20MW until 2027 is apparently set to become law.

Dozens of similar efforts show the key areas of electricity and water use, noise, and economic impact emerging consistently.

Community opposition to data center expansion has intensified, particularly around energy, water, and land use.

Factually speaking, data centers throughout the US today consume one gallon of water for every 10,000 gallons consumed for overall human use. IDCA research has also found no direct connection between new data center electricity consumption and immediate consumer price increases. As far as noise, the majority of data center developers are now focusing on rural areas or locations on the edge of populated areas, and build in substantial setbacks for the actual facilities.

The seemingly key issue that arises repeatedly – and arose in the very early days of the original large data centers from Amazon, Google, and Microsoft – is a lack of transparency by the developers. A new data center is typically proposed by a local developer with a generic, opaque name that aims to mask the true developer. The true developer's name is often shared with some (but not all) local political and business leaders, however, leading to community frustration.

IDCA encourages developers to be transparent when they set their sights on a new data center site. This advisement is made now that there are often competitive reasons for the opacity (that is, one large cloud developer doesn't want one of its competitors to know early on where it is building), but recent experience has shown that concerns about electricity, water, and noise can be sufficiently addressed if companies are upfront about who they are and what they intend to do.

IDCA research has found that significant community and political pushback starts to occur in nations once their data center footprints have reached the 5-percent consumption level of national grids. Singapore has doubled that consumption level already, several European nations are well past it, and the United States has now reached the 6-percent level overall.