Data Centres: South Africa's quiet digital revolution—and the risks they bring?

Credit: Researched, written and compiled by Tony Stone, End summary by ChatGPT

Artificial intelligence is transforming the global economy at breathtaking speed, and behind every AI application lies an immense physical infrastructure that most people never see. Data centres have become the backbone of modern banking, healthcare, government, commerce and digital communications, quietly processing and protecting the information that underpins everyday life. Yet this technological revolution is placing unprecedented demands on electricity, water resources, cybersecurity and public governance. In South Africa—where ageing infrastructure, water losses and energy constraints already present significant challenges—the rapid expansion of data centres raises important questions about sustainability, resilience and the protection of citizens' personal information. This article examines how these interconnected issues are unfolding, separating evidence from speculation while exploring the opportunities, risks and governance challenges that will shape South Africa's digital future.

To understand why data centres are a threat to your personal security, you need to understand the technology. You also need to understand why the combination of data centres, municipal water loss and the coming “super” El Nino pose an even greater threat. Last but not least, how secure is your private information safe from prying eyes and the criminally minded? Let us deal with municipal water loss and El Nino first.

Municipal water losses: South Africa's national average for non-revenue water (water lost before reaching the customer, stolen, or unbilled) stands at approximately 47.3%. This means nearly one in every two liters of treated drinking water is lost, significantly higher than the global average of about 30%. This massive shortfall is primarily driven by decaying or poorly maintained infrastructure, burst pipes, water theft (including illegal connections), and malfunctioning water meters.

El Nino impact: The expected impact of the "super" El Niño on South Africa includes significantly hotter-than-normal temperatures, delayed or erratic summer rainfall, higher evaporation rates, and a heightened risk of drought. This is anticipated to strain water resources, stress crops and livestock, and increase the risk of consumer food price inflation. The El Niño phenomenon will coincide with the vital summer crop planting season, which typically begins in mid-October. The expected high heat and lack of rainfall will threaten newly planted summer grains like maize. Should significant crop damage occur, agricultural economists warn it could drive up consumer food prices in 2027 and threaten to slow down interest rate cuts. Beyond dry conditions, a "super" El Niño adds significant heat to the atmosphere, bringing a higher likelihood of extreme heatwaves. This combination of prolonged dryness, high temperatures, and lower humidity levels will drastically increase the danger of wildfires throughout the interior. The combination of diminished rainfall and higher evaporation rates from extreme heat poses a threat to the country's water reservoirs and dams. If dry weather persists through the summer months, it could also impact the energy sector, compounding pressure on local water and utility infrastructure.

The National Oceanic and Atmospheric Administration have gone on record to say that the most devastating impacts of El Nino to South Africa’s summer rainfall, agriculture, and water reserves will be concentrated from November 2026 through March 2027.

So, the net effect of municipal water loss and the impact of El Nino will be significant. Now add data centres. As I said, we need to understand not just the technology but the extent to which this technology is used, which underpins the need for yet more data centres.

South Africa’s data centres

South Africa has over 56 data centres, representing roughly 1% of global AI data centre capacity. The country is the primary hub for cloud and server infrastructure on the continent, attracting massive investments from major players like Amazon Web Services, Google, and Microsoft.

These facilities are highly concentrated around economic hubs such as Cape Town and Johannesburg, with the highest density in the Gauteng province. The vast majority of South Africans live in urban areas and major economic centres, with about 66% to 71% of the population classified as urban.



Photo: ACS InformationAGE



Photo: CBRE Accelerate

Primary grid and renewable sources

Data centres use a multi-layered power infrastructure to ensure uninterrupted operations. They rely on the primary utility grid, on-site clean energy (like Solar Power), and continuous backup technologies including large-scale batteries, diesel generators, and hydrogen fuel cells.

However, most data centres take electricity from the regional Eskom electrical grid for base load electricity needs to run their servers and cooling systems. On-Site generation facilities use localized solar PV (Photovoltaic) arrays and wind turbines to supplement traditional grid power and achieve sustainability goals. High-density data centres and AI hubs are increasingly exploring long-term renewable power solutions, including hydroelectricity and small-scale nuclear (e.g. pebble bed) reactors for alternative base-load needs.

Uninterruptible power supply (UPS): Data centres traditionally use valve regulated lead-acid (VRLA) batteries but are rapidly shifting toward lithium-ion batteries for better space efficiency and lifespan. These power supplies are used to stabilise incoming power and instantly bridge power gaps during voltage fluctuations or brief grid outages.

Backup generation: Diesel generators are the industry standard for long-duration outages. These massive generators (often ranging up to several megawatts) automatically start up in seconds to assume the facility's power load. Alternatively, gas and dual-fuel turbines are used, increasingly so, as they produce cleaner emissions and provide higher power density compared to traditional diesel engines. More recently, hydrogen fuel cells have emerged as another alternative zero-emission technology to provide continuous backup power for off-grid electricity needs.

Advanced internal distribution and delivery: Massive high-load hardware is used to safely receive high-voltage incoming electricity. Switchgears and transformers that isolate circuits and step down the voltage to usable levels for the direct current (DC) architecture used by the new high-density server chips and artificial intelligence CPUs (like those from NVIDIA). These chips demand so much power data centres are beginning to use 800-volt DC systems internally to minimize energy loss from repeated voltage conversions. Power distribution units (PDUs) are devices that further reduce voltage and distribute power directly to server racks and individual server blades.

Cooling technologies

With all the heat generated, data centres utilize several water-cooling technologies ranging from facility-wide systems to chip-level solutions. These include chilled water systems (loops transferring heat to external cooling towers), direct-to-chip cooling (pumping coolant directly over processors), immersion cooling (submerging servers in dielectric fluid), and rear door heat exchangers.

Leading cooling vendors such as Vertiv and Schneider Electric classify these technologies into these primary types:

Facility-level water cooling: Chilled water systems pump chilled water through pipes to air-handling units across the facility. The water absorbs heat from server exhaust air and transports it to external cooling towers, where heat is expelled into the atmosphere.

Cooling towers and evaporative systems, often open-loop, spray water down into a tower, evaporating a portion of the water to cool the remaining bulk water. While thermally efficient, they consume significant amounts of water.

Closed-loop chilled water plants are used in advanced data centres (including many powering high-density AI clusters) where the cooling loop is filled once and then sealed. These systems cycle coolant continually, significantly reducing daily water usage.

Rack and component-level liquid cooling: As high-density hardware like AI GPUs generate too much heat for traditional air-cooling, data centres are adopting localized liquid-to-server systems.

Direct-to-Chip (D2C) Cooling routes liquid coolant through micro-channels or cold plates resting directly on high-heat components like CPUs and GPUs. The heat is absorbed by the liquid and transferred out of the server enclosure.

Immersion Cooling systems submerge entire server nodes directly inside specialized, non-conductive (dielectric) fluids. Heat is transferred directly from the components to the liquid, eliminating the need for internal server fans. Rear door heat exchangers (RDHX) is a liquid-to-air hybrid system where a liquid-filled heat exchanger replaces the back door of a server rack. Fans push the hot exhaust air from the servers through this liquid-cooled door, cooling the air before it re-enters the data centre room.

South African operators are chasing efficiency

In South Africa, the landscape for data centre power consumption has fundamentally shifted. You can literally produce additional power but you cannot produce additional water. You have to cut your usage. Operators can no longer rely solely on the national utility, Eskom, due to grid constraints and aggressive new environmental penalties. Instead, a combination of liberalised electricity regulations and a strict Phase 2 Carbon Tax framework have forced hyperscalers into pioneering private renewable energy strategies.

Historically, the National Energy Regulator of South Africa (NERSA) strictly controlled power generation. Radical policy overhauls have dismantled these barriers to protect large power users. Amendments to Schedule 2 of the Electricity Regulation Act (ERA) entirely removed the requirement for a generation licence for facilities under 100 megawatts (MW). Data centre operators can now build or contract massive private solar and wind farms without years of regulatory red tape.

Because data centres are built in urban economic hubs (e.g. Cape Town, Johannesburg) while the best solar/wind resources are in remote regions (e.g. Northern Cape), operators rely heavily on wheeling. "Wheeling" is a regulatory mechanism where a private IPP (Independent Power Producer) injects clean energy into the national or municipal grid at point A (e.g. Northern Cape), and the data centre offsets its consumption from the grid at point B (e.g. Johannesburg), paying a nominal transit fee to Eskom or local municipalities.

Carbon tax

As per usual, The South African Revenue Service (SARS) has to have its slice. Phase 2 of South Africa's carbon tax came into effect, profoundly altering corporate energy economics. This regime imposes heavy penalties on fossil-fuel reliance: The nominal carbon tax rate has jumped to R308 per tonne of CO_2e, and it is legally locked to rise to R462 per tonne by 2030. Data centres face a double financial hit. Beyond their direct emissions (like testing diesel backup generators), Eskom is now permitted to pass its own heavy coal-fired carbon tax liability directly onto the consumer tariff. This adds up to an estimated 11c per kWh surcharge onto grid power, incentivising businesses to transition to self-generation immediately. The previous 5% voluntary carbon budget allowance has sunsetted. Large emitters who exceed their new *mandatory* carbon budgets face a punitive, un-allocatable penalty rate of R640 per ton CO_2e.

To mitigate these carbon liabilities and dodge grid instability, South African data centre operators are signing massive, multi-megawatt Power Purchase Agreements (PPAs) to execute structural offsets:

Under Phase 2 rules, the maximum carbon offset allowance has been expanded to 15% for emissions (up from 10%). Data centre operators can purchase verified, locally-developed carbon credits from eligible domestic green projects to legally lower their carbon tax exposure.

Because data centres operate 24/7/365, daytime solar is insufficient on its own. Facilities are increasingly pairing solar farms with high-capacity battery energy storage systems (BESS) or nocturnal wind PPAs to reduce direct Scope One diesel generator runtime during load shedding or grid peak hours.

Data centre users

The users of South African data centres span from global tech giants to local businesses, as these facilities act as the central nervous system for the continent's digital economy. Because South Africa is the primary landing point for high-capacity subsea internet cables (like Google's Equiano and Meta's 2Africa), companies use these data centres to serve the entire Sub-Saharan region with low latency.

Technology companies: The largest single tenants in hyperscalers and cloud service provider facilities are global cloud providers. They rent massive, dedicated "data halls" (wholesale colocation) to run their localized cloud regions. Microsoft uses local data centres to power its Azure cloud regions in Johannesburg and Cape Town. Amazon Web Services (AWS) operates its own infrastructure alongside colocation space to run the AWS Africa (Cape Town) Region. Google Cloud launched its Johannesburg cloud region to provide localized enterprise services. Oracle runs its cloud infrastructure out of its local carrier neutral facilities.

Banking and finance: South Africa has a highly advanced banking sector. Local laws (such as the Protection of Personal Information Act or POPIA) mandate that citizens' financial data must be stored within the country's borders. Major Retail Banks: Institutions like Standard Bank, Absa, Nedbank, and First National Bank (FNB) host

their core banking systems and backup disaster recovery sites here. Payment processors and fintechs, companies like Peach Payments, Yoco, and regional clearing houses, use these hubs to process millions of transactions securely and instantly.

Telecommunications and networks: Data centres double as "carrier hotels," for telecommunications and network providers where different networks physically connect to one another to swap internet traffic. Vodacom, MTN, Telkom, and Cell C locate their core network nodes inside these facilities. Content delivery networks (CDNs), **c**ompanies like Cloudflare and Akamai cache popular global internet content locally so videos and websites load instantly for South African users. NAPAfrica, located inside Teraco, is the largest Internet Exchange Point (IXP) on the continent. Over 600 networks—including Netflix, Apple, and Microsoft—peer here to exchange internet traffic directly.

Enterprise, Retail, and Government: Large organizations use data centres to avoid the massive cost of building and maintaining their own server rooms. Massive retail groups (such as Shoprite, Pick n Pay, and the Foschini Group) use them to manage supply chains, point-of-sale systems, and online shopping apps. Enterprise companies migrate their SAP or Oracle databases to local colocation facilities for better reliability and to protect against local power cuts. Government departments gradually move public records and digital services into highly secure, private zones within these commercial data centres.

SARS and the SARB: These organisations use these data centres through a hybrid strategy, combining their own highly secure private facilities with commercial cloud networks. As critical organs of state handling national revenue and the banking financial system, they cannot fully rely on public commercial facilities. Instead, they navigate strict domestic laws—such as the National Policy on Data and Cloud and the Protection of Personal Information Act (POPIA), which legally mandate that sensitive citizen data, financial records, and national tax transactions stay localized within South African borders. Both entities utilize a combination of dedicated private infrastructure and localized commercial infrastructure.

SARS has undergone a massive digital transformation, evolving from a paper-reliant entity into a high-tech, data-driven organization. SARS manages its own primary network data centres, including core hubs located in Alberton and Doringkloof. These private facilities host their main internal LAN servers, VPN breakouts, and high-security contact centre traffic. To run its advanced data science algorithms, machine learning models, and automated third-party data profiling (such as auto-populating tax returns), SARS relies on secure public cloud platforms. For instance, their enterprise voice systems and Microsoft Teams tenant configurations are officially hosted via Microsoft's Office 365 cloud architecture, which utilizes localized commercial data centre space in Johannesburg and Cape Town.

As the central bank, the South African Reserve Bank (SARB), enforces strict compliance rules regarding "data sovereignty" and offshoring, meaning all financial surveillance and payment system data must remain onshore. SARB operates its own independent, high-security infrastructure. Their internal technology framework dictates two primary, mirrored private data centre sites alongside a geographically separated third disaster recovery site to guarantee total business continuity if the municipal grid or connectivity collapses. While keeping core monetary systems on their own isolated metal servers, SARB permits payment operators and fintech innovators to leverage local commercial clouds (like AWS, Azure, or Google Cloud), provided that the physical data storage hardware remains tightly mapped to local, onshore data zones

Global digital ID systems integration: SARB and the Department of Home Affairs (DHA) are actively integrating data centre infrastructure to launch a unified National Digital Identity ecosystem. Historically, South African government departments worked in technological silos. However, the rollout of the national Digital ID framework and the Digital Financial Identity (DFID) system has forced a tightly coordinated, shared architecture across local data centres.

National Digital ID Infrastructure	
DHA National Identity Biometric Database	**SARB Payment Utility & Financial Identity**
* Onshore Government Cloud	* PayInc / BankServ Platform
* Public-Private Gateways	* Low-latency Commercial Hubs
* Cryptographic Engines	* Instant Payment Rails

The state is executing this integration across core areas:

DHA Cryptographic Biometric Hubs: The Department of Home Affairs is the ultimate custodian of citizen identity data. To support the Digital Wallet app—which lets citizens store secure, digital versions of their smart IDs and passports on mobile devices—the DHA has upgraded its national Online Verification Service (OVS). Onshore Commercial Hosting: Under new Draft Digital Identity Regulations, the system requires low-latency, high-volume processing to execute instant biometric "proof-of-liveness" checks (facial recognition selfies). This heavy processing

load is distributed securely to local commercial data centres via asymmetric cryptography pipelines. Banking Infrastructure Partnerships: The DHA has integrated its secure backend gateway directly with the computing architecture of major retail banks (like Standard Bank and FNB). This allows citizens to use bank branches and banking apps as secure, distributed data touchpoints to register and update their digital credentials.

SARB's Digital Financial Identity and the NPU: The SARB is modernising the nation's transaction rails to ensure financial platforms connect natively to the new national ID. The SARB acquired a 50% controlling stake in PayInc (formerly BankServ Africa), transitioning it into a centralized the national payment utility (NPU). PayInc's architecture runs across advanced local carrier-neutral data centres to handle high-speed, cash-like instant digital transactions. The SARB is deploying its own digital financial identity (DFID) framework. By embedding this system directly into the country's transactional middle-mile infrastructure, the central bank enables commercial fintech apps to instantly verify a user's identity against national databases, wiping out banking fraud.

The push for a sovereign state-owned cloud: While the government leverages the footprint of commercial hyperscalers like Microsoft and Amazon Web Services for outward-facing digital tools, it is wary of international data offshoring. President Cyril Ramaphosa confirmed that South Africa is actively investing in its own sovereign cloud infrastructure. Spearheaded by institutions like the Council for Scientific and Industrial Research (CSIR), the goal is to keep the hyper-sensitive root databases of the new national digital identity framework hosted exclusively on state-controlled server metal, strictly isolated within South African physical territory – its own data centre.

Credit bureaus: Companies like TransUnion and Experian are heavily reliant on these South African data centres through a specialized hybrid model. Because credit bureaus handle hyper-sensitive personal identity data, banking histories, and credit scores for over 25 million South Africans, they are strictly bound by the Protection of Personal Information Act (POPIA) and the National Credit Act (NCA). These laws legally prevent them from moving citizens' core credit records to overseas servers. To stay compliant while scaling their processing power, they distribute their technology across two tiers of local infrastructure. Both TransUnion and Experian maintain their root, highly secure master databases on local server metal physically located within South Africa. Experian explicitly hosts its main credit databases within localized infrastructure across Stellenbosch and Johannesburg. This ensures that when a bank pulls your credit score, the data is fetched directly from local soil, satisfying POPIA's strict cross-border restrictions. TransUnion operates a distributed network of isolated local servers within South Africa. Following a massive historic cyberattack targeting its local branch, TransUnion deeply fortified its domestic hardware layout to separate consumer credit monitoring files from public-facing internet nodes.

While the underlying raw data stays anchored on local physical drives, these bureaus use the multi-billion Rand hyper-scaler data centres built by Microsoft Azure and Amazon Web Services (AWS) to run their modern analytics products. Experian's automated credit risk management software and automated loan origination platforms (like Compuloan Cloud) are actively hosted via Microsoft Azure's South African data centres. This allows banks and microlenders to query Experian's data with zero lag time while processing loan applications. TransUnion has shifted towards real-time "alternative credit scoring" (such as tracking mobile phone usage and call data records with operators like MTN to score people who don't have formal bank accounts). Processing these billions of data points in real-time requires the elastic computing scale found in Johannesburg's commercial hyperscale facilities

Medical services: If you live in South Africa, your digital medical records are highly likely stored inside the exact hyper-scaler and colocation data centres discussed in this conversation. Under Section 26 of the Protection of Personal Information Act (POPIA), health and medical data is legally classified as "Special Personal Information". POPIA imposes severe penalties—including multi-million Rand fines—if this data is transferred or hosted across national borders without strict, explicit justification. As a result, South Africa's entire healthcare ecosystem has migrated its electronic medical records (EMRs) and digital health profiles onto local soil. Your records are distributed across these local centres depending on your healthcare provider.

The largest medical schemes use the hyper-scaler data centres built by Microsoft Azure, Amazon Web Services (AWS), and Google Cloud in Johannesburg and Cape Town to process member data securely. Discovery's widespread HealthID platform—which gives doctors access to your long-term health history, electronic prescriptions, and blood test results—is securely integrated across local cloud infrastructure. Furthermore, Discovery's newly launched *Vitality AI* and Personal Health Pathways systems are explicitly built on and hosted via Google Cloud Platform's localized South African infrastructure. Schemes administered by companies like Medscheme or Momentum use secure local cloud frameworks to run automated claims processing while ensuring full data sovereignty compliance.

If you visit a private hospital group (such as Netcare, Life Healthcare, or Mediclinic) or a local GP, your file is stored digitally via localized clinical software providers. Companies that provide South African medical practices with billing

and clinical record software (such as GoodX, PrivateHub, or DG Connect) host their backup servers and live databases natively within carrier-neutral data centres like Teraco or Africa Data Centres. To satisfy the strict March 2026 Information Regulator updates on handling health data, these local data hosts must legally keep your records encrypted both "at rest" on the physical server drives and "in transit" when sent to your doctor's screen.

If you use the public healthcare sector, the data distribution is more fractured but still strictly onshore. Systems like the Western Cape's Single Patient Viewer or national health registries are managed natively on dedicated state infrastructure rather than commercial public clouds. These legacy public systems are gradually being integrated into the upcoming Sovereign State Cloud ecosystem to keep public health data completely isolated from global commercial entities

South African data centres are undergoing massive physical modifications and retrofits

To receive high-density Artificial Intelligence (AI) and Machine Learning data, these facilities, which were originally designed for traditional "Cloud workloads" (storing emails, streaming video, or hosting databases), the global explosion of AI has shifted the goalposts. AI data is fundamentally different—it requires massive processing clusters of Graphical Processing Units (GPUs) rather than standard Central Processing Units (CPUs). Because of this, operators across South Africa are transforming standard data halls into what industry leaders call **"AI Factories."** This architectural overhaul requires three primary structural modifications.

Standard cloud servers are cooled by blowing cold air through the server racks. However, AI hardware generates so much heat that traditional air systems cannot physically transfer the thermal energy away quickly enough. Operators are modifying existing white floors to integrate Direct-to-Chip liquid cooling and Rear-Door Heat Exchangers (RDHx). Teraco recently completed a major expansion of its JB4 Bredell Campus in Ekurhuleni. The new data halls are engineered with advanced liquid-to-liquid cooling loops, allowing clients to instantly deploy extreme-density AI workloads right next to standard air-cooled cloud setups.

A traditional data centre rack typically requires between 3 to 10 kilowatts (kW) of electricity. By comparison, a modern AI server rack packed with specialized AI chips can demand anywhere from 40 to over 100 kW per rack. Local data centres are physically tearing out older power distribution units (PDUs) and replacing them with high-capacity copper busways and localized substations to feed these concentrated energy spikes. The South African AI data centre market size is projected to grow from roughly $99 million to over $572 million by 2031, with hardware upgrades and GPU cluster deployments completely outpacing software expansions.

AI processing requires thousands of chips to talk to each other simultaneously with zero latency. If the connection lags, the entire AI model slows down. Companies like Equinix are launching specific AI-ready platforms in Johannesburg (such as their JN1 facility) to act as a neutral switchboard. They are retrofitting internal fibre networks to allow multiple hyperscale clouds (AWS, Google, Azure) to cross-connect at lightning speeds

South Africa's data centres' water consumption

South Africa's 56 data centres consume vast volumes of water for cooling. The total annual consumption for the market is estimated at 17.94 billion litres. While average-sized sites can consume roughly 1.1 million litres per day, the industry is transitioning to closed-loop systems to prevent evaporation and minimize waste. The specific water usage for these facilities is driven by several key factors and trends.

Data centres use water directly for evaporative cooling servers and indirectly through the water required by the national grid to generate electricity.

A one (1)-megawatt data centre using traditional cooling methods can consume up to 6.75 million litres of water annually. For larger hyperscale facilities, consumption can reach 3 million litres or more every single day. The Impact of AI: As more AI-capable servers are deployed, the need for cooling capacity increases. Some estimates suggest a single AI query can use about 500 millilitres of water.

Processing AI queries accelerates infrastructure strain on South Africa's already vulnerable water and energy grids, forcing a massive shift toward resource self-sufficiency. Because the region is naturally water-scarce and its national power network faces severe constraints, hyperscale data centres cannot rely exclusively on standard municipal infrastructure without triggering localized crises.

The dual-grid impacts of the AI boom in South Africa break down into distinct pressures:

The municipal water grid impact: The physical processing of AI queries creates localized water vulnerability in metropolitan hubs like Johannesburg and Cape Town.

Traditional data centre cooling relies heavily on evaporative cooling towers, which convert billions of litres of potable municipal fresh water into steam. This places tech infrastructure in direct competition with local residential water supplies in a country averaging just 450 mm of annual rainfall.

South Africa's data centre water consumption is projected to grow from 17.94 billion litres in 2025 to 29.42 billion litres by 2030. This rapid escalation stresses municipal pipe infrastructure, which is already prone to structural leaks and restrictions.

To circumvent municipal water dependency, top-tier local providers like Teraco and Africa Data Centres increasingly build closed-loop, chilled-water cooling loops or deploy private boreholes and rainwater harvesting systems.

The energy grid impact (Eskom & wheel power): AI computation requires high-performance chips that run significantly hotter and consume far more electricity than standard cloud workloads.

Major cloud infrastructure investments by Microsoft, Amazon Web Services (AWS), and Google Cloud draw between 20 MW to 100 MW per facility at launch, quickly scaling to several hundred megawatts. The Department of Communications and Digital Technologies explicitly mandates that operators reduce reliance on the national power network to protect domestic supply.

While the country has added renewable generation capacity, South Africa’s primary energy crisis has shifted to transmission grid capacity limits. There is insufficient infrastructure to move power from optimal renewable zones to the high-density metropolitan data hubs.

To keep AI workloads active, operators are building utility-scale private generation plants. For example, Teraco established a 120MW solar plant in the Free State to "wheel" renewable energy across Eskom and local municipal networks directly to their server halls.

Summary Resource Footprint

Resource Dimension	Core Infrastructure Strain	Primary Local Mitigation
Direct Water Grid	Evaporative cooling drains drinking reserves.	Boreholes, rainwater harvesting, air-cooling.
Indirect Water Grid	Water consumed during coal-fired electricity generation.	Transitioning to solar and wind power.
Electricity Grid	High-density AI chips require massive, 24/7 baseload power.	Private solar farms and energy wheeling.
Grid Transmission	Weak line capacity limits new energy allocations.	Custom wheeling framework agreements with municipalities.

What aggravates the current status is South Africa being firmly established as the primary digital and data centre gateway to Africa. It hosts nearly half of the continent’s total installed capacity and serves as the central hub for major cloud and AI infrastructure. The country accounts for roughly 300MW to 350MW of critical IT capacity. Global tech giants, such as Microsoft, Google, and Amazon Web Services, utilize South Africa as their central launchpad for enterprise and AI workloads on the continent. Providers like Teraco and Africa Data Centres dominate the colocation market here. Digital infrastructure providers use South Africa's robust facilities as a springboard to distribute data and cloud services into smaller neighbouring markets. While countries like Nigeria, Kenya, and Egypt are growing rapidly as regional nodes, South Africa remains the de facto engine of Africa's digital economy. As Africa’s (millions of) IT and AI users become more advanced and place a heavier demand on South Africa’s data centres, direct and indirect water consumption will increase dramatically. The key question is, will South Africa cope?

Data security and privacy

Data centres themselves do not know anything about you, because they are completely blind to the data they host. While your medical records, bank accounts, and credit scores are physically stored inside these buildings, there is a fundamental difference between a data centre operator (who owns the building) and a cloud tenant (the company you actually do business with). To ensure absolute privacy, the industry relies on a strict combination of structural, legal, and cryptographic barriers. The physical hard drives and server blocks inside these centres use advanced encryption (like AES-256). When your medical aid or bank uploads your file, the data is scrambled into unreadable code before it even leaves their system. The data centre operator (like Teraco or Vantage) does not have the

cryptographic keys to unscramble that data. To them, your entire life's medical history looks like a random, meaningless string of letters and numbers.

The companies that build and run these large facilities operate as "colocation providers" or digital landlords. Think of a data centre like a highly secure apartment building. The landlord owns the building, provides the electricity, plugs in the water pipes, and keeps the security guards at the front door. However, the landlord is legally forbidden from opening your apartment door, looking through your cupboards, or reading your personal mail. Tech teams working for the data centre operator are only responsible for maintaining the physical hardware, the air conditioning, and the backup generators. They cannot log into the software running on the servers.

South Africa's protection of personal information act (POPIA) draws a very strict legal line between two roles: the responsible party e.g. your bank, your hospital, or SARS are legally responsible for your data, decide how it is used, and are the only ones who can view it. the operator e.g. the data centre. under POPIA, an operator is legally restricted to processing data only on the direct instructions of the responsible party. if a data centre operator tried to look at, analyse, or sell your data, they would face massive criminal prosecution and immediate shutdown by the information regulator.

Even inside the building, data is kept in silos. If a bank rents space in a data centre, their servers are locked inside physical steel cages with biometric fingerprint and facial scanners. Employees from other companies—or even the data centre's own staff—cannot physically access those specific server racks without formal, logged clearance from the bank itself.
So, while these buildings hold the digital footprints of nearly every South African, they act as blind vault keepers rather than all-knowing entities.

Cybercrime
South Africa is the most targeted country in Africa. It accounts for over 40% of all ransomware attacks and 35% of infostealer incidents on the continent. Cyberattacks cost the South African economy an estimated R142 billion annually. However, while cybercriminals hacking data centres directly is a low-probability it is a high-impact threat in South Africa, private sector cyber security is highly adaptable, well-funded, and better equipped to recruit and retain cybersecurity professionals. However, it faces a massive threat surface and has varying security postures depending on the company's size and industry. Public sector cybersecurity often struggles with legacy IT infrastructure and budget constraints that make it difficult to compete with corporate salaries. Despite these hurdles, government cybersecurity guidelines mandate strict compliance frameworks, and agencies increasingly incorporate security at the earliest stages of technology planning.

Regardless of the sector, ongoing, routine employee cyber security training is critical. Keeping abreast of the methods and tactics of cybercriminals is vital. These criminals generally breach infrastructure using *Distributed Denial-of-Service (DDoS) Attacks*, where threat actors launch massive inbound traffic floods against local hosting and infrastructure providers (like 1-grid and Host Africa) to knock thousands of businesses offline. Alternatively, they use *Ransomware and Extortion* where cybercriminal groups infiltrate organizational networks, encrypt data, and demand exorbitant ransoms. High-profile public and private institutions—including Statistics South Africa—have been hit by threat groups like XP95. Lastly, cybercriminals use *Third-Party and Peripheral Vulnerabilities* where, rather than hacking through heavily guarded data centre firewalls, cybercriminals target unsecured peripheral devices (printers, IoT sensors, and office routers) and exploit human vulnerabilities through sophisticated phishing.

Given the ingenuity of cybercriminals, proactively keeping one eye on the ball while the other watches the field of play is a good and necessary habit.

Restricted government access
The government cannot simply walk into a data centre and pull a complete master profile on you, because South African law creates strict legal firewalls between data silos. While a data centre physically houses the servers of multiple entities (like banks, hospitals, and different state departments), there is no centralized "master key" or single database that stitches all this information together. The government's ability to access your information across these facilities is heavily restricted by three major legal and structural realities. By law, government departments cannot browse or cross-reference data outside of their explicit legal mandate. Under the Protection of Personal Information Act (POPIA), data collected for one specific reason (e.g., your health status at a public clinic) cannot legally be linked to data collected for another reason (e.g., your tax file at SARS), unless explicitly authorized by a specific act. If a state security agency or the police want to access your data hosted on a commercial cloud or a telecom server, they cannot just request it from the data centre. Under the Regulation of Interception of

Communications and Provision of Communication-related Information Act (RICA), they must formally apply to a designated High Court Judge for an interception warrant, proving that you are suspected of a serious crime.

Even if the government owns or sub-leases space inside a commercial data centre like Teraco, they only have access to their own leased "cages." The government has absolutely no physical or digital access to the servers belonging to private companies like Discovery Health, Standard Bank, or Vodacom, even if those servers sit in the exact same room. Private enterprises manage their own encryption keys. Even if the state seized a physical hard drive from a commercial data centre, the data on it would look like scrambled, unreadable text without the private encryption keys, which are held securely by the private corporations. The South African state does not have a single, unified data hub. The government's own IT infrastructure is notoriously siloed across different networks. For example, the DHA operates its biometric and identity systems. SARS runs its own secure tax networks. The Department of Health holds provincial medical registries. Because these systems are run by different departments and often managed by the State Information Technology Agency (SITA) on separate legacy networks, they do not natively talk to or share data with each other.

The only scenario where the government accesses data inside these centres is when a specific department accesses its *own* legally collected data (e.g., SARS checking your tax filings), or when law enforcement serves a legally binding court order directly to a specific private company (like your bank) to hand over records for a criminal investigation. The data centre operator is never involved in this process.

Constitutional protection

The probability that the South African government will legally dissolve the privacy barriers to create a centralized, all-seeing master profile is low, but the probability that they will try to bypass or weaken specific access laws through "national security" updates is moderate. Rather than a sudden, direct takeover of commercial data centres, the evolving dynamic between state power and data privacy in South Africa is driven by distinct structural, economic, and political factors. To legally authorize the state to crawl through private databases (like your bank or medical aid) sitting inside a data centre, the government would have to dismantle the country's constitutional framework. Section 14 of South Africa's Constitution guarantees the Right to Privacy.

The Constitutional Court has historically been fiercely independent. If parliament passed a law allowing blanket, warrantless mass surveillance of data centre infrastructure, civil society groups (like amaBhungane) would challenge it immediately, and the court would likely rule it unconstitutional. South Africa is actively positioning itself as the primary digital and AI hub for the African continent. President Ramaphosa and global tech giants (like Google and Microsoft) have structured multi-billion Rand investment campaigns on the guarantee of data sovereignty and local legal stability. If the state altered the law to peek into any server at will, hyperscalers would halt expansions, local business operations would freeze, and international investment would collapse. The State Information Technology Agency (SITA) is currently undergoing massive internal interventions following severe investigations into corruption and procurement delays. The state natively struggles to maintain its own basic database systems. Structurally speaking, the government lacks the technical infrastructure required to build a functioning, centralized surveillance mechanism that could decode encrypted data across diverse, multi-tenant commercial data centres.

While the government is unlikely to change the law to take *everything*, they are actively looking for ways to expand their reach in specific, legally grey areas. the state can easily amend existing crime laws rather than rewriting privacy laws. Law enforcement continuously lobbies to expand what constitutes a "justifiable interception" under RICA. The risk is that the definition of a threat becomes broader, allowing the state to request specific encryption keys from tech or telecom companies more frequently under the guise of protecting national security. The government's National Policy on Data and Cloud requires all state departments to migrate to a unified, cloud-enabled infrastructure. As the state builds this sovereign state cloud to manage the developing national digital identity systems, it places hyper-sensitive citizen data under direct executive control. While your private banking files remain separate, your root identity files, biometric data, and tax history will be unified, creating a highly lucrative target for state overreach or internal data leaks. To make the new digital ID system functional, private entities (like banks) have to build direct data pipelines to the Department of Home Affairs. While these pipelines are built only to verify identities, history shows that once a technical gateway exists between the state and a private data centre, the temptation for government entities to request broader analytics access increases.

However, the government cannot easily flip a switch to read your entire life off a server. The legal, economic, and technical costs of doing so are too high. Instead, the real point of vulnerability over the next few years is not the data centre buildings themselves, but rather how securely the government builds its own state cloud systems and whether citizens can successfully hold state departments accountable whenever data breaches occur.

The Constitutional Court previously struck down state surveillance laws by declaring core components of the Regulation of Interception of Communications and Provision of Communication-Related Information Act (RICA) unconstitutional. In the landmark 2021 case of the Minister of Police v AmaBhungane Centre for Investigative Journalism, the apex court ruled that the state's espionage framework severely infringed upon the Section 14 Constitutional right to privacy without containing adequate legal safeguards. The Constitutional Court dismantled state surveillance powers across five specific operational vectors.

The court completely banned the practice of bulk interception—where state security agencies "hoover up" entire streams of international fibre-optic internet and voice traffic passing into the country to scan them for keywords. The court ruled that this sweeping mass surveillance was completely unauthorized by any South African law and declared it immediately illegal.

Under the old RICA framework, if the state tapped your phone, you would never be told, even if you were found entirely innocent. The court struck this down, noting that absolute secrecy prevents citizens from ever seeking legal redress for privacy violations. It mandated that individuals must be notified that they were spied on within 90 days after the surveillance order expires.

Historically, a single, state-appointed retired judge issued interception warrants in an *ex parte* (secret) meeting where only the spy agency's side was heard. The court ruled that the selection process for this designated judge lacked transparency, compromising institutional independence. It ordered changes to ensure the selection mechanism safeguards objectivity.

The court found that RICA failed to recognize the unique legal privileges that protect the confidentiality of journalistic sources and lawyer-client consultations. The court forced an interim update: state agencies applying for a wiretap must explicitly disclose to the judge if their target is a practicing attorney or a journalist so that heightened scrutiny can be applied.

The court ruled against the sloppy handling of intercepted communications. It noted that RICA lacked clear, uniform regulations regarding how captured data is stored, examined, copied, and destroyed, ordering parliament to implement stringent data management protocols.

This historic ruling forms the exact legal wall that prevents the government from installing blanket "backdoors" inside local commercial data centres. It cements that any data access must be targeted, individually authorized by an independent judge, and subjected to subsequent transparency.

In conclusion

ChatGPT provided this summary when I asked it (no offence untended) for a review:

"The issue here is not whether any one technology is inherently dangerous. Data centres, AI, cloud computing, digital identity, and renewable energy all have legitimate purposes. The real question is whether the institutions responsible for governing them are developing at the same pace as the technology itself, or the cybercriminals trying to break in.

South Africa faces several risks that are individually manageable but potentially serious when they converge:

- **Infrastructure stress.** AI requires enormous amounts of electricity and, depending on the cooling technology, significant volumes of water. In a country already struggling with electricity reliability and high water losses, those demands require careful long-term planning.
- **Concentration risk.** More government services, banking, healthcare, and commerce are becoming digitally interconnected. That improves efficiency, but it also means that failures—whether cyberattacks, outages, or governance failures—can have much wider consequences.
- **Governance capacity.** Sophisticated technology is only as good as the institutions managing it. Technical expertise, procurement integrity, cybersecurity, and maintenance become increasingly important as reliance on digital infrastructure grows.
- **Public understanding.** Most citizens have little idea where their data resides or how digital identity, cloud services, and AI infrastructure operate. That makes the informed public debate more difficult."

ChatGPT hits the nail on the head. Neither the data centre service providers and, perhaps more so, their downstream clients cannot afford to be caught napping!! The premise that "a chain is only as strong as its weakest link" applies.