



HARDWARE & SERVICE GUIDE

Private AI that runs on hardware **you control**

A ChatGPT-style workspace, document Q&A, Excel analysis and AI agents for your team, on a server in your office or, for lighter use, in Technosult's Teraco DB1 data centre in Durban. Your data and intellectual property stay in South Africa.

Your data stays put

On-site, nothing leaves your building. Hosted, it stays in South Africa under a POPIA operator agreement.

No per-token bills

The models run on your hardware. Use them as much as you like for a fixed monthly fee.

You own it

Open-source software, your hardware, your data. No lock-in to Technosult or any AI vendor.

What you get

01

Private Chat & Knowledge

A browser workspace on your intranet. Named logins and cited answers from your approved documents.

02

Data & Excel Agent

Upload a spreadsheet, ask questions, get cleaned workbooks, pivots and charts back.

03

Desktop Co-worker

An agent on staff PCs that sorts, summarises and converts files in approved folders.

04

Automations

Scheduled jobs into accounting, payroll, CRM and Microsoft 365, with your approval on every write.

05

Private Agents

Technosult AI Agents for email, calendar and WhatsApp, running on your local model.

06

Governance Pack

POPIA documentation, audit logs, access reviews and staff training.

Two ways to run it

On-Site Appliance

An AI server on your office network, with UPS, encrypted disks and firewall included in the quote.

Data never leaves your premises. TechnoSult configures, monitors and supports it over a secure VPN.

TechnoSult Hosted

A dedicated virtual server in TechnoSult’s rack at Teraco, Durban, with dual power feeds and multiple carriers. It runs small AI models on CPU and memory, with no GPU. Your data is held on its own isolated server, separate from other clients.

Suits light, entry-level use. No hardware to buy; South African data residency, fully managed by TechnoSult.

Hardware tiers

Complete, ready-to-run systems built from components sold by South African suppliers, with local warranty, and sized to the number of people using the AI at the same moment. Every tier runs US-made open-weight models by default, such as Google Gemma 4 and OpenAI gpt-oss; Qwen 3.5 is available on request.

Tier	Typical system	AI memory	Runs well	Indicative cost (incl. VAT)
Compact Up to 5 users	RTX 5090 AI workstation, built locally: Ryzen 9, 64 GB RAM, 2 TB NVMe	32 GB GPU	Google Gemma 4 26B and 31B at fast speeds	R90,000 – R120,000
Performance 5–15 users	Dual RTX 5090 server: Ryzen 9, 64–128 GB RAM, mirrored NVMe	64 GB GPU	Gemma 4 for many users at once, including long documents	R180,000 – R230,000
Scale 15–50 users	RTX PRO 6000 Blackwell server; the card is usually ordered in	96 GB GPU	OpenAI gpt-oss 120B and other 100B+ models, many users at once	From R380,000 quoted
Hosted Light, entry-level use	Virtual server in TechnoSult's Teraco DB1 rack, Durban; CPU only, no GPU	Server RAM	Small Gemma 4 models for drafting, summaries and document Q&A; slower replies	No hardware to buy

Users means concurrent users: people using the AI at the same moment, usually far fewer than your total staff.

Hardware prices are indicative and refreshed on every quote (valid 7 days), because global memory shortages move prices weekly. Every on-site quote includes UPS, encrypted disks, firewall set-up and encrypted off-site backup. Rent-to-own over 24–36 months is available.

Service pricing

Setup covers design, build, hardening, installation, knowledge-base loading and training. The monthly fee covers updates, monitoring, model refreshes, backups checks and support.

Tier	Setup & deployment	Monthly management	What's included
Local Starter Up to 5 users	R25,000 – R35,000	R3,500 – R4,500	Install and hardening, models and updates, one knowledge base, Data & Excel Agent, one training session, Governance Pack essentials, email support next business day
Local Professional 5–15 users	R45,000 – R65,000	R6,500 – R8,500	Everything in Starter, plus role-based knowledge bases, Desktop Co-worker for all users, two automations, WhatsApp support same business day, quarterly review
Local Enterprise 15–50 users	From R80,000	R9,500 – R14,000	Everything in Professional, plus custom integrations, Private Agent set-up, priority support within 4 business hours, monthly review, full Governance Pack
Technosult Hosted Light, entry-level use	Starter fee	Quoted	Local Starter services on a small-model virtual server at Teraco DB1, including hosting and monitoring. No hardware purchase.

Start here: AI Readiness & POPIA Assessment

Workshop, data map, risk rating and a written recommendation. Credited in full against setup if you go ahead.

R15,000 – R25,000

Options & add-ons

Option	Price (excl. VAT)	What it does
Hybrid Gateway	R1,500/month + cloud AI usage at cost + 15%	Redacted, logged, policy-controlled access to frontier cloud models
Private Agents	Technosult AI Agent tier price	Email, calendar and WhatsApp agents on your local model; no OpenRouter account needed
Staff training	Per half-day session, quoted	Prompting, checking AI output, POPIA do's and don'ts
Hardware rental	Quoted over 24–36 months	Removes the upfront hardware cost

Built so the AI stays in its lane

"Rogue AI" usually means an agent acting beyond its brief, often because a document carried hidden instructions. These controls are switched on in every deployment.

Read-only by default

Agents only see the folders and systems you approve.

You approve every action

Nothing is written, sent or deleted without a person's OK.

Sandboxed code

Excel and data code runs in a sandbox with no network access.

No outbound internet

The AI server reaches nothing outside except an update allowlist.

Everything logged

Every prompt, tool call and change is recorded for audit.

One-click stop

Any agent or automation can be paused instantly.

Vetted add-ons only

No marketplace plugins; scoped, read-only API tokens.

Never exposed online

No public ports. Remote access only through a VPN.

Built for POPIA compliance

South Africa has no AI-specific law yet, so POPIA applies to every AI system that touches personal information. The Governance Pack gives your Information Officer the paperwork to show it.

- ✓ AI use policy for staff
- ✓ Operator agreement when TechnoSult hosts or manages your system
- ✓ Retention schedule for chats, files and indexes
- ✓ Staff training register and quarterly access review
- ✓ Personal-information impact assessment and data-flow map
- ✓ AI register: systems, models, data sources and owners
- ✓ Incident response runbook

Your organisation stays the responsible party under POPIA. TechnoSult provides the technical controls and documentation templates; it is not your legal adviser.

Typical use cases

Knowledge assistant

Policies, procedures and product information, answered with sources.

Excel & data processing

Reports, reconciliations and clean-ups from your spreadsheets.

Document Q&A

Contracts, statements and files for legal, finance and HR teams.

Quotes & proposals

First drafts in your house style, built from your templates.

Customer replies

Draft answers from your own knowledge base, for staff to approve.

Technical assistant

Private help with code, scripts and IT documentation.

An honest view of local AI

Today's open-weight models handle most daily office work well. For a few demanding jobs, frontier cloud models still lead.

Local AI handles well

Drafting emails, letters, quotes and policies
Summarising documents and meetings
Answering questions from your own documents
Excel and CSV analysis
Extracting data from invoices and PDFs

Cloud AI still leads at

Multi-hour autonomous projects
Complex software development
The hardest reasoning problems

Need both? The optional Hybrid Gateway sends redacted, non-personal tasks to a frontier model under your policy, with every request logged.

How we deliver

01

Discover

Week 1

Readiness & POPIA assessment, top tasks, success criteria.

02

Build

Week 2

System built, hardened and tested on your sample files.

03

Deploy

Week 3

Installed, connected and your staff trained.

04

Pilot

Weeks 3–5

Two or three use cases measured against written criteria.

05

Run

Monthly

Updates, monitoring, model refreshes and access reviews.

Ready to scope your system?

- 1 Book a discovery call.
- 2 Tell us how many people will use the AI at once, and for what.
- 3 Choose On-Site or Hosted.
- 4 Receive a fixed proposal: hardware, setup and monthly fee.

brett@technosult.co.za · www.technosult.co.za