



Powered by



PIT WALL AI · TEAM SOUTAR MOTORSPORT

AI Race Intelligence System

DESIGN DOCUMENT

A multi-agent AI that turns raw telemetry and the driver's own words into ranked, trustworthy recommendations — before the next session starts.

PROGRAMME

Microsoft NOVA

DELIVERED BY

Mr.PC Advanced IT Solutions

PLATFORM

Azure AI Foundry · Copilot+ PC

CAR

McLaren Artura GT4



The vision

Motorsport is a data sport. The bottleneck isn't the data — it's the time to read it.

The car already logs everything you need to go faster.
Understanding it before the next run is the hard part.

411 ch · 1 kHz logged every lap — data was never the problem

2 cars · 4 drivers run by one lead engineer and a junior

Minutes between Practice and Qualifying to act

THE OUTCOME THE DESIGN TARGETS

~4_{min} Flag to ranked Top-3, between sessions	Top 3 Actions per driver — never noise
100% Of AI claims cited to a fact	0 Car actions without the engineer

- Driver Coach
- Race Engineer
- Tyre & Consumables
- Peer-Review

HOW A SESSION FLOWS THROUGH THE SYSTEM



The AI engine

Not one model guessing — a panel of specialists reasoning over one shared, verified fact base.



Four specialists · one source of truth · orchestrated on Azure AI Foundry, grounded in season history by RAG.

**NO HALLUCINATION
BY DESIGN**

Deterministic code owns the numbers. The AI writes & ranks over cited facts. Models never see raw traces — only a compact fact table and the driver's words. Every claim traces back to a fact ID.

THE MODEL MAP — WHICH MICROSOFT AI DOES WHAT

GPT-5

DEFAULT

Driver Coach · Race Engineer · Tyre. Frontier reasoning — the everyday workhorse that explains and ranks.

GPT-5 Pro

EXTENDED THINKING

Deep setup calls. Selective, more deliberate reasoning for the hardest multi-variable questions.

Claude Sonnet

2ND OPINION

Peer-Review. An independent model on Foundry pressure-tests the highest-stakes outputs.

GPT-5-mini

HIGH-VOLUME

Classification & tagging. Fast, low-cost structuring at scale.

GPT-4o-transcribe

VOICE

Driver debrief. Turns a 60-sec spoken debrief into structured, queryable feedback.

embedding-3-large

RAG

Season memory. Retrieves setups, debriefs, manual & regs so suggestions stay legal.

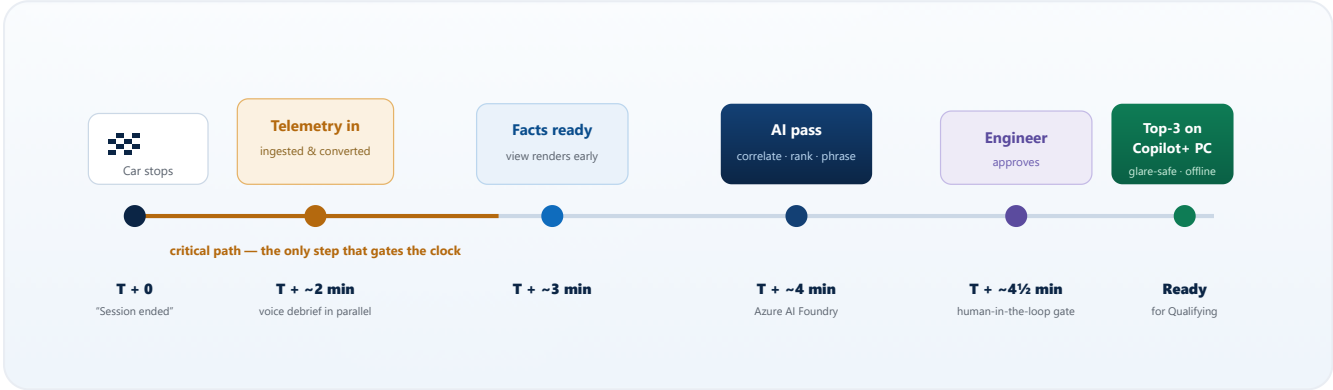
Local Ingest & Prep app

ON-DEVICE · COPILOT+ PC · NO LLM

The edge is deterministic. Trackside decoding and structuring run on the team's laptop with plain code — no on-device model — then upload a compact summary. Every LLM call stays in the team's Azure tenant.

Everything is designed backwards from one moment: the five minutes between sessions.

If an insight can't be computed, ranked and read in that window, it belongs in the overnight review — not on the pit wall.



WHY IT SURVIVES THE PIT LANE

- Overlap** Local view renders the moment facts land; the AI enrichment lands on top
- On-device** Top-3 renders from kilobytes on the Copilot+ PC — no cloud round-trip
- Graceful** If the cloud is slow, deterministic insights show, marked "provisional"

EVERY SETUP SUGGESTION CLEARS THREE GATES

GATE 1
In range
 Within the Artura's validated window

GATE 2
Legal
 Regulation- & BOP-legal, checked by RAG

GATE 3
Reversible
 Quick-to-undo preferred first

THE DELIVERABLE — A RANKED TOP-3

Illustrative — how a Top-3 would read · Nirwan #75 · ~4 min to Quali

1

Southern Loop · +0.16 s — you lift mid-complex; teammate stays flat on same tyres. **High**

2

+0.5 psi rear hot — rears below window; reversible test. **Med**

3

T4 braking · +0.11 s — braking early; trust the brakes deeper. **High**

SPLIT TWO WAYS
 Each item is tagged **Driver** (coach it) or **Setup** (change it) — the same insight, phrased for each audience.

What the AI is designed to see

The patterns a rushed human misses — GT4 examples are illustrative of intended output.



Cross-corner correlation

One root cause in several corners at once — the AI would surface the pattern, not three separate tweaks.

ILLUSTRATIVE

Understeer in T1, T4 & T9 → one **front-end direction**.



Compensating behaviour

A quick driver hides a car problem — lap time looks fine while tyre life is quietly spent.

ILLUSTRATIVE

Exit correction masks looseness → **rear-tyre cost hidden**.



Stint degradation

A fuel-corrected pace line would reveal when a tyre is done — and which axle.

ILLUSTRATIVE

"Rear's done after lap 6" → **protect the rears early**.



Perception vs data

The driver's feeling turned into a testable hypothesis — presented alongside data, never overridden.

ILLUSTRATIVE

Feels understeer → data says **early throttle on exit**.



"Money corners"

Corners ranked by realistic, achievable gain — not raw delta — so effort lands where the tenths are.

ILLUSTRATIVE

T1 hairpin: +0.19 s vs teammate, same tyres.



Early reliability signs

Trends invisible in a glance become amber "watch / act" flags long before a failure.

ILLUSTRATIVE

Oil-pressure droop at peak lateral-G in one corner.

ALSO SURFACED ACROSS TIME

● **Car vs car — teammate deltas isolate driver/setup from car**

● **Across a stint — fall-off attributed to axle & corner type**

● **Thermal story — the physical cause behind lost time**

● **Across a weekend — track evolution vs genuine gains**

DESIGN PRINCIPLE

Every insight is **computed by deterministic code and cited to specific laps, corners and channels** — the AI explains and ranks it, and its evidence is always one tap away.

Technology & trust

The Microsoft AI stack the design uses, and the principles that govern how it behaves.

THE MICROSOFT AI STACK



Azure AI Foundry

Hosts the frontier models — GPT-5, GPT-5 Pro, Claude Sonnet, GPT-5-mini — and orchestrates the multi-agent pass.



Voice & memory

GPT-4o-transcribe structures the driver debrief; text-embedding-3-large powers the season-history RAG.



Local Ingest & Prep app

Runs on the team's Copilot+ PC — decodes telemetry and uploads a compact summary; offline-capable, deterministic, no on-device model.

DESIGN PRINCIPLE

Runs in the team's own Azure tenant, in-region. Telemetry, setups and feedback stay resident in-tenant with no data egress.

THE TRUST MODEL

1

Human-in-the-loop

Nothing reaches the driver or car without the engineer's tap-to-approve.

2

Advisory only

The system never actions the car — it offers the best-supported options and why.

3

Every claim cited

One tap from the laps, corners and channels behind it. No black boxes.

4

Honest uncertainty

Low confidence stated plainly; driver conflicts surfaced, never hidden.

5

Self-calibrating

Predicted vs actual tracked each session — the system reports its own hit rate.

THE LINE THE DESIGN WILL NOT CROSS

The AI amplifies the engineer's judgement — it never replaces it. Accountability stays with the crew.

PLANNED DELIVERY TIMELINE — 2026 · target dates, design intent



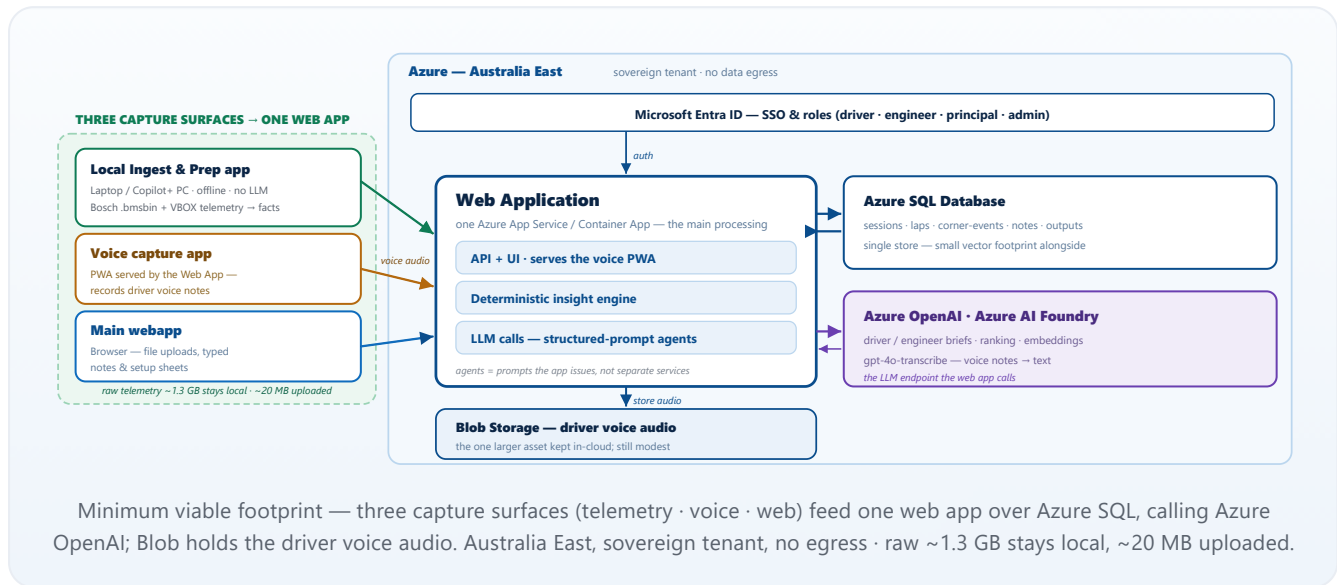
The design's intent, in one line: the turnaround and pattern-coverage of a full data department for a two-person crew — **ranked, cited, in-tenant, and always under human control.**

THE AI RACE INTELLIGENCE SYSTEM · PIT WALL AI · TEAM SOUTAR MOTORSPORT · MCLAREN ARTURA GT4

Prepared by Mr.PC Advanced IT Solutions under the Microsoft NOVA programme. A forward-looking design document describing intended system behaviour, written ahead of implementation. All examples are illustrative of design intent, not measured results. No Microsoft trademarks or logos are reproduced.

Solution architecture & Azure services

A1 Solution architecture — three capture surfaces, one web app, and the minimum-viable footprint



A2 Azure services map

COMPONENT	WHERE	ROLE IN THE DESIGN
Local Ingest & Prep app (surface 1)	LOCAL	On the team laptop / Copilot+ PC, offline-capable — decodes Bosch .bmsbin + VBOX, formats setup sheets, uploads the compact processed data. Deterministic, no edge LLM.
Voice capture app (surface 2)	PWA	A lightweight mobile / browser PWA served by the Web App — records driver voice notes, uploads the audio for transcription into a session note.
Web Application (+ main webapp, surface 3)	CLOUD	A single Azure App Service (or one Container App) — hosts the API, UI (uploads, typed notes, setups), insight engine and LLM calls. The "agents" are structured prompts it issues.
Azure SQL Database	CLOUD	The single data store — sessions, laps, corner-events, notes & outputs, with a small vector footprint held alongside. No separate vector service.
Azure OpenAI (AI Foundry)	CLOUD	The LLM endpoint the web app calls — briefs, ranking, embeddings and voice-note transcription (gpt-4o-transcribe).
Microsoft Entra ID	CLOUD	SSO and roles — driver / engineer / principal / admin; managed identity for the app's data access.
Azure Key Vault	CLOUD	Key management; encryption at rest & in transit for Azure SQL and Blob Storage.
Azure Monitor / Log Analytics	CLOUD	App health, pipeline latency (does export→brief stay in the window?), model usage & cost. Basic logging suffices — no separate SIEM.
Blob Storage	CLOUD	Holds the driver voice audio uploaded by the voice app — modest, the only larger asset kept in-cloud; raw telemetry stays local.

Descoped as unnecessary at this scale: separate ingestion / feature / orchestrator workers (Container Apps), an Event Grid / Service Bus spine, a data-lake tiering layer, a standalone vector database, and an on-device LLM — all folded into the one web app and Azure SQL. A BI dashboard (e.g. Power BI) is a possible future add-on.

AI stack, data platform & governance

A3 AI & model stack — which model does what

GPT-5	Driver Coach · Race Engineer · Tyre agents — frontier reasoning over the fact base.
GPT-5 Pro	Selective deep, extended-thinking setup recommendations.
Claude Sonnet	Peer-review — independent 2nd opinion on the highest-stakes outputs (on Foundry).
GPT-5-mini	High-volume classification & tagging of facts and feedback.
GPT-4o-transcribe	Driver voice notes (from the voice capture app) → structured, queryable feedback.
embedding-3-large	Embeddings for RAG over history, notes & setups; vectors in Azure SQL.
Local prep app	On-device decoding & structuring on the Copilot+ PC — deterministic code, no LLM.

GROUNDING & THE NO-HALLUCINATION PRINCIPLE

- **Deterministic code owns the numbers.** The models see a compact fact table + the driver's words, never raw 1 kHz traces.
- The AI's job is to **explain, prioritise and phrase** — every claim carries a citation back to a fact ID.
- **RAG over season history** retrieves setups, debriefs, the car manual & regs, so any suggestion is checked against what is legal and adjustable before it is shown.

A4 Data platform & ingestion

SOURCES

- Bosch telemetry
- Voice notes (voice app)
- Engineer notes
- Setup sheets
- Regs & car manual

LOCAL VS UPLOADED

- **Local:** raw ~1.3 GB
- **Uploaded:** ~20 MB facts
- voice audio → Blob
- raw stays on laptop

SQL SCHEMA ENTITIES

- sessions · laps
- corner_events
- notes · setups
- agent_outputs

CAPTURE → ATTRIBUTE → PROCESS

- 3 surfaces → API
- pick pre-created session
- ingest to Azure SQL
- insight → LLM calls

Capture at source: sessions & drivers are pre-created (car · round · circuit · session · stint); every surface picks from that list (inline + **Add new**), so each asset is attributed at capture — Bosch→VBOX reconciliation verifies the pair.

A5 Security, governance & residency

IDENTITY — ENTRA ID

- SSO & RBAC roles
- Managed identities for uploads & service auth

KEYS — KEY VAULT

- Encryption at rest & in transit
- Keys held & rotated centrally

RESIDENCY

- All data in Australia East
- **No egress** — telemetry & IP stay in-region

OBSERVABILITY

- Monitor / Log Analytics
- Basic logging — no SIEM

HUMAN-IN-THE-LOOP & AUDIT

No output reaches the driver or the car without the engineer's approval; every recommendation is cited and every decision logged — a full audit trail, advisory-only, never auto-actioned on the car.

A6 Delivery timeline (reference) — planned 2026 dates, design intent

PHASE · TARGET DATES	INTENDED TECHNICAL DELIVERABLES
Solution development · Jul 1–14	Architecture, the deterministic fact pipeline, corner model & retrieval index, and the web app + its three capture surfaces built.
Initial production · Jul 14–28	First live deployment; sources onboarded to Azure SQL; Top-3 between-session view, driver & engineer briefs and voice capture in use.
Production completion & tuning · August	Feature-complete; setup→outcome corpus, calibration (predicted vs actual) reporting, accuracy & performance tuning and QA.
Monitoring & fine-tuning · September	Live monitoring and fine-tuning; evaluation and case-study delivery.