THE ORBA ECOSYSTEM

L’intelligence invisible et souveraine a désormais son écosystème.

Concept de Déploiement Souverain | Architecture Zero-Cloud

Priorité au Local (Zero-Cloud)

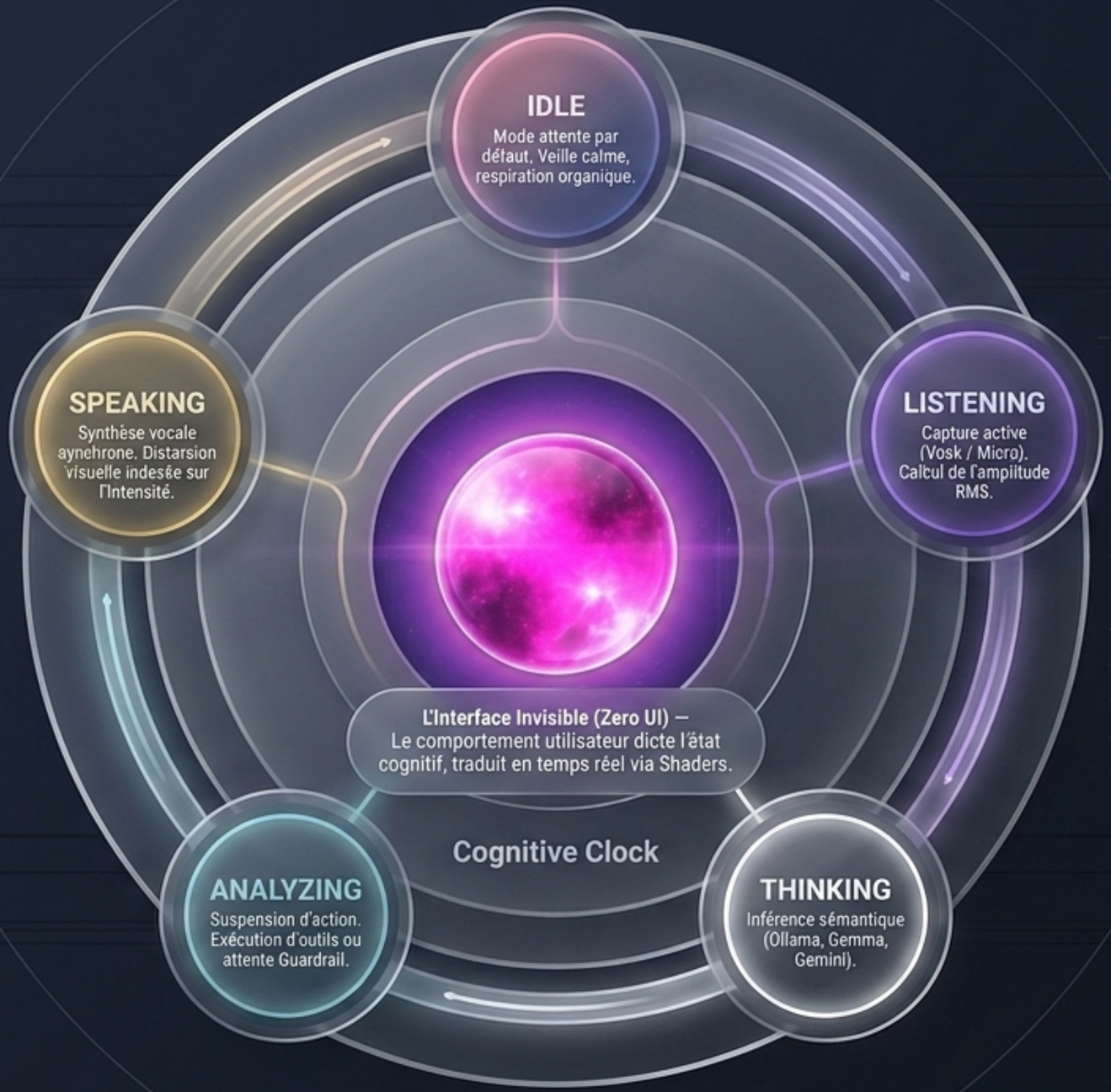
100% Offline-First. Inférence locale (Ollama / LiteRT) gardant les données en RAM volatile locale. Vosk et Piper traitent la parole à la source, sans exfiltration audio.

Transparence Cognitive (ReAct Log)

Élimination de la boîte noire. L'agent affiche ouvertement ses pensées intermédiaires et les outils ciblés avant la formulation de la réponse finale.

Contrôle Absolu (Zero-Trust)

Architecture Human-in-the-loop. Aucune modification système critique (fichiers, scripts) sans une approbation explicite via Guardrails matériels ou distants.


IDLE
Mode attente par défaut. Veille calme, respiration organique.
LISTENING
Capture active (Vosk / Micro). Calcul de l'amplitude RMS.
THINKING
Inférence sémantique (Ollama, Gemma, Gemini).
ANALYZING
Suspension d'action. Exécution d'outils ou attente Guardrail.
SPEAKING
Synthèse vocale asynchrone. Distorsion visuelle indexée sur l'intensité.
Cognitive Clock
L'Interface Invisible (Zero UI) —
Le comportement utilisateur dicte l'état cognitif, traduit en temps réel via Shaders.

L’intelligence déployée sur trois niveaux matériels.

Mobile (Android Native)

Matériel/OS: NPU de poche, Android 13+ (Kotlin).

Moteur IA: LiteRT Gemma-2B-IT (100% Hors-ligne).

Voix: Android SpeechRecognizer / C++ NDK Piper (JNI).

UI & Accès: Shaders AGSL à 60 FPS, contrôle natif des capteurs système.

Desktop (Windows/Mac/Linux)

Matériel/OS: Tauri v2, Backend FastAPI Python.

Moteur IA: Agent ReAct Multi-tours (Ollama local privilégié).

Voix: Vosk (16000Hz) / Piper ONNX (Synchronisation WebSockets).

UI & Accès: OrbaSphere flottante sans bordures, Gateways WhatsApp/Telegram

Web (Showcase Platform)

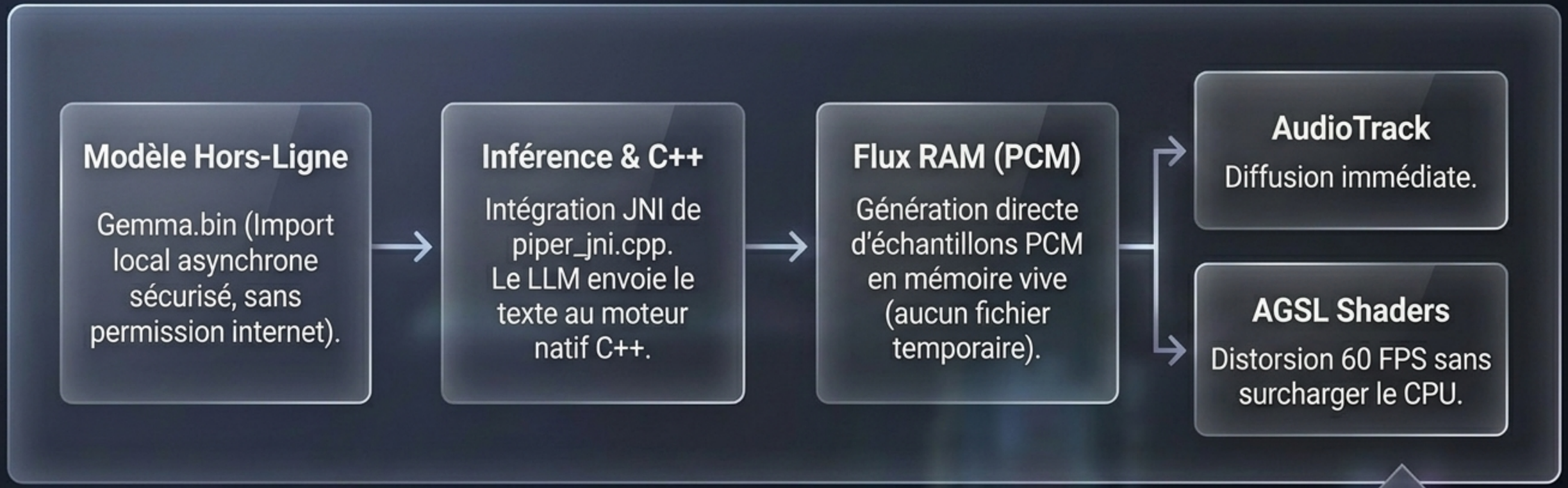
Matériel/OS: Navigateur Web (HTML5, JS, WebGL).

Moteur IA: Gemini 2.5 Flash Native Audio.

Voix: Traitement audio natif direct en ligne.

UI & Accès: Canvas 3D organique simulé, vision spatiale multimodale.

Le NPU de poche : Pipeline Audio Direct et Latence Zéro (Android)



Protection OOM Killer — Implémentation de onTrimMemory() pour décharger OrbaBrain dynamiquement lorsque l'application passe en arrière-plan, préservant la fluidité du système d'exploitation.

Station de Travail Autonome : L’Agent ReAct de Bureau (Tauri + FastAPI)

Widget Border-less Always-on-Top

Interface transparente WebGL. L'animation visuelle est synchronisée dynamiquement via WebSockets calculant la valeur RMS de la synthèse locale.

Boucle Cognitive Multi-tours

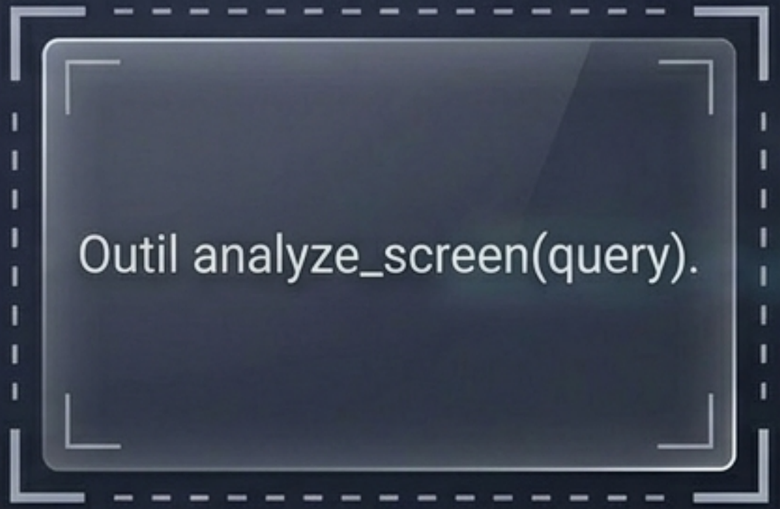
Le fichier main.py orchestre une boucle (jusqu'à 3 tours) permettant à l'agent de chaîner les requêtes locales (Ollama) et de réfléchir avant de répondre.

Gateways Multi-Canaux

Pilotage et validation à distance via intégration Twilio WhatsApp native pour une supervision déportée.

Proactivité et Agentivité : L'IA qui observe et agit

Vision Multimodale



Outil analyze_screen(query).

Capture en direct avec Pillow → Analyse sémantique par Gemini 1.5 → Extraction de réponses visuelles.

Protégé par Guardrail.

Planificateur Local

Boucle de surveillance asynchrone sur scheduled_tasks.json.

Exécution autonome en arrière-plan avec réveil de l'utilisateur via Notifications Push Natives Windows (PowerShell Toast).

Architecture Zero-Trust : Classification des Outils (tools.py)

SAFE

Exécution : Immédiate et transparente.

```
read_file
list_directory
open_app
list_scheduled_tasks
```

CRITICAL

Exécution : Bloquée par défaut. Bascule de l'OrbaSphere à l'état ANALYZING (Cyan).

```
delete_file
write_file
execute_system_command
schedule_task
analyze_screen
```

Aucune action destructrice ou modification système ne peut être finalisée sans l'insertion d'un humain dans la boucle décisionnelle.

Le Flow Guardrail : Validation à double canal

Requête Utilisateur : Supprime ce fichier.

ANALYZING (Interception par l'Orchestrateur)

Action suspendue. Timeout dynamique (ORBA_APPROVAL_TIMEOUT).

Validation Locale

Modale UI Tauri au centre du bureau.

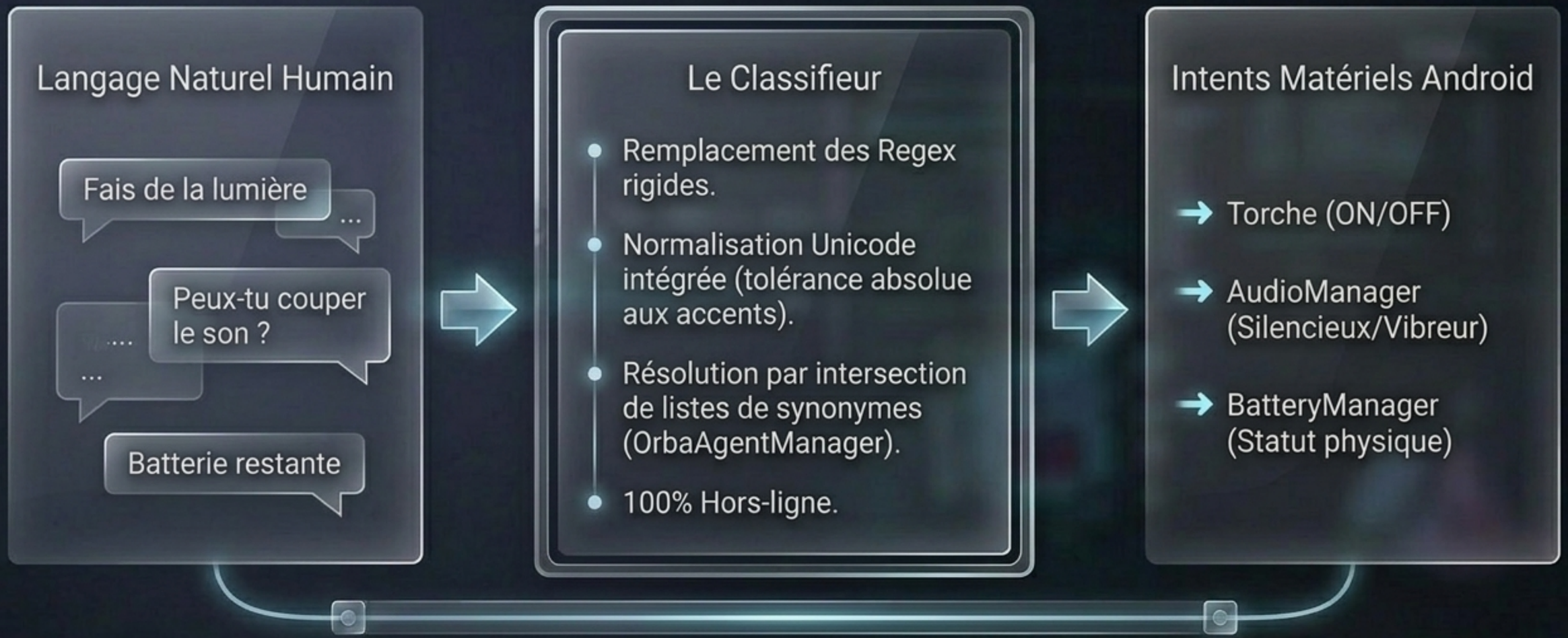
Validation Distante (WhatsApp)

Alerte envoyée via Twilio sur smartphone.

L'utilisateur répond textuellement : OUI [ID] ou NON [ID].

Exécution de l'action système ➔ Rapport de succès renvoyé.

Le Routeur Sémantique Mobile : Piloter le matériel sans le Cloud



Synthèse de l’Audit 2026 : Preuve de Robustesse

Boucle de Feedback Vocale Infinie

[100% RÉSOLU]

Problème : Le micro STT (Vosk) écoutait la propre voix du TTS (Piper).
Solution : Ignorance logicielle stricte des transcriptions pendant les états SPEAKING, THINKING ou ANALYZING.

Surcharge GPU (Canvas WebGL)

[100% RÉSOLU]

Problème : Consommation d'énergie en arrière-plan sur configs modestes.
Solution : Suspension de requestAnimationFrame dès que l'onglet/fenêtre passe en veille.

Rigidité des Timeouts de Sécurité

[100% RÉSOLU]

Problème : Fenêtre d’approbation humaine (Guardrail) codée en dur.
Solution : Rendu dynamique et personnalisable via variable .env (ORBA_APPROVAL_TIMEOUT).

Déploiement Modulaire : Reprenez le contrôle de votre intelligence

Desktop

git clone Orba_OS ➔ python main.py ➔ npm run tauri dev

Mobile

Import Android Studio ➔ Compile CMake/NDK ➔ Lancement du ModelDownloader local

Showcase

Hébergé sur Netlify (Simulation WebGL & Gemini 2.5 Flash Native Audio)

"The future of interaction is invisible."

— Alex Koncept