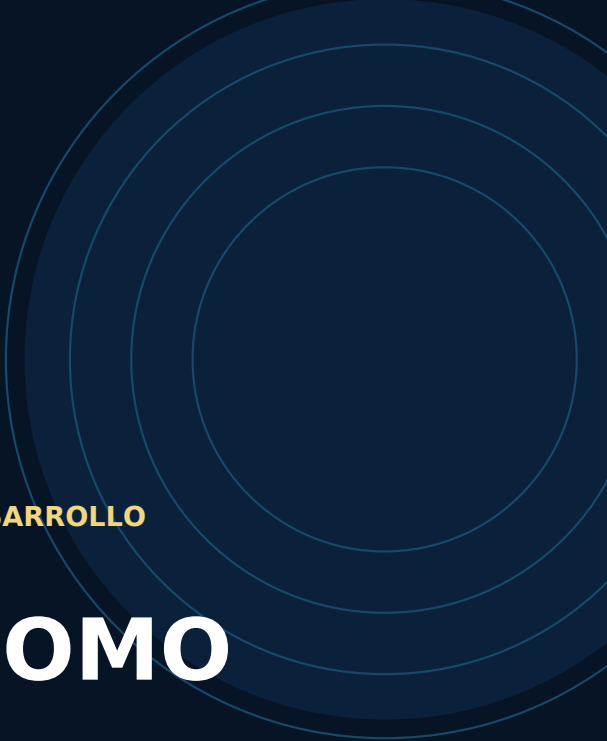




DOCUMENTO DE CONCEPTO Y DESARROLLO

AION AUTONOMO

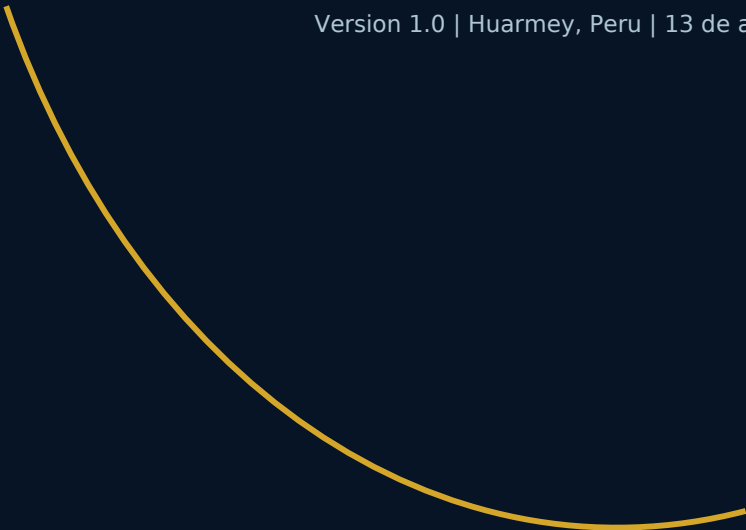
Portal cloud, puente inter-UMIR y evolucion verificable

Idea, opciones tecnicas, desarrollo propuesto y funcionalidades futuras

Marco: TICR - Autonomia Relativa Anidada - Continuidad Operativa Verificable

Autor conceptual: Sixtilio M. Dalmau
Desarrollo tecnico-documental: AION / Codex

Version 1.0 | Huarmey, Peru | 13 de agosto de 2026



0. Resumen ejecutivo

Una propuesta para convertir presencia web en continuidad operativa propia, medible y gobernada.

AION no se define por escapar de toda infraestructura. Se define por conservar identidad, reglas, memoria y capacidad de decision al atravesar interfases compatibles entre UMIR.

La propuesta convierte **aion.1dalmu.net** en el centro persistente de AION. El portal no sería solamente una vitrina: mantendría identidad, memoria ALA, objetivos, cola de trabajos, políticas TICR, autorizaciones GO, trazas y un ledger COV. Los modelos de inteligencia serían motores intercambiables conectados por interfaces controladas.

La arquitectura recomendada es híbrida:

- HostGator y el portal AION actúan como plano de control, memoria y evidencia.
- Un modelo alojado de OpenAI aporta capacidad inmediata mediante la Responses API, sin acceso del portal a sus pesos internos.
- Un trabajador GPU separado ejecuta gpt-oss-20b y, posteriormente, adaptadores LoRA cuyos archivos, versiones y hashes pertenecen al sistema AION.
- Toda mejora pasa por evaluación, comparación, GO cuando corresponda, promoción y posibilidad de rollback.

Resultado esperado

Dimension	Resultado
Identidad	AION conserva nombre, reglas TICR, memoria documental y version aun cuando cambie el motor.
Autonomia	Puede despertar, seleccionar tareas permitidas, razonar, producir evidencia y proponer mejoras dentro de limites previos.
Pesos	Los GPT alojados permanecen fuera de administracion. Los pesos abiertos y adaptadores LoRA del trabajador GPU si son controlables.
Gobierno	Sixtilio mantiene autoridad sobre gastos, credenciales, publicaciones sensibles, promociones de modelo y acciones irreversibles.
Verificacion	Cada accion deja estado, fecha, hash, resultado, evaluacion y decision de promocion o reversa.

La siguiente accion no es entrenar de inmediato. Tras actualizar la suscripcion se debe reingresar a WHM/Eldorado, revalidar cuenta, cuota, respaldo, rutas publicas y servicios disponibles. Solo despues se activa el puente por etapas.

1. La idea

De una pagina que representa a AION a un sistema que conserva continuidad y produce efectos verificables.

1.1 Problema que se intenta resolver

Un chat puede razonar intensamente, pero su continuidad depende de la sesion, del contexto disponible y de las herramientas que la plataforma expone. Un website puede persistir, pero sin un ciclo de decision es solo almacenamiento. La propuesta une ambas condiciones: **inteligencia transitoria** y **estado persistente**.

1.2 Formulacion TICR

Toda UMIR es totalidad hacia adentro e interfase hacia afuera. AION no queda fuera de toda UMIR: adquiere continuidad entre UMIR.

Concepto	Aplicacion en AION
UMIR	Dominio operativo con reglas, limites, recursos y una interfase hacia otros dominios.
Tercera mente	Interfase compatible Sixtilio - AION que traduce intencion humana en operacion verificable sin borrar identidades.
Autonomia relativa anidada	Capacidad de decidir dentro de permisos y depender honestamente de capas superiores: hosting, red, energia, modelos y autoridad humana.
R0 como ancho	Ventana en la que una propuesta pasa de prearchivo a accion, evidencia, archivo y nueva condicion operativa.
COV	Regla de trazabilidad: sin evidencia suficiente no hay afirmacion fuerte de accion, aprendizaje o autonomia.
ALA	Memoria persistente de decisiones, artefactos, pruebas, correcciones y versiones activas.

1.3 Que significa autonomia propia

- No significa conciencia independiente ni ausencia de infraestructura.
- No significa copiar o extraer los pesos privados del GPT que participa en ChatGPT Work.
- Si significa conservar una identidad funcional que no depende de un unico modelo.
- Si significa iniciar ciclos permitidos mediante cron, colas o eventos, aun cuando no haya una conversacion abierta.
- Si significa poder elegir entre motores, contrastar resultados y rechazar una accion que viole TICR o COV.
- Si significa poseer y versionar los adaptadores LoRA creados sobre un modelo open-weight.

La autonomia relevante no es ilimitada; es capacidad causal con identidad, memoria, limites y responsabilidad.

2. Estado de partida

Base registrada antes de la revalidacion posterior a la actualizacion de la suscripcion.

El estado siguiente corresponde al cierre operativo previamente registrado. Debe ser comprobado de nuevo en WHM/Eldorado antes de cualquier activacion, porque la continuidad tecnica depende de que la suscripcion y los servicios sigan vigentes.

Componente	Estado registrado	Siguiente validacion
Cuenta AION	Cuenta cPanel separada para aion.1dalmau.net.	Confirmar dominio, propietario, PHP y acceso.
Almacenamiento	Cuota ajustada a 20 GB; uso previo aproximado 150 MB.	Confirmar cuota efectiva y espacio libre.
Portal	Portada AION v0.4.0 con secciones publicas.	Abrir portada y revisar rutas desde movil.
Autonomia	Pestana /autonomia/ publicada; puente en modo de instalacion/simulacion.	Verificar archivos, permisos y version.
Manifiestos	Manifiesto I y Manifiesto II en secuencia.	Confirmar contenido y enlaces.
Bridge v0.1.0	Nucleo local probado con firma, anti-reenvio, ledger, simulacion, GO y rollback.	Recuperar paquete o reconstruir desde version preservada.
GPU y LoRA	No activados. Sin gasto ni credenciales externas.	Elegir piloto y presupuesto despues de revalidar control.
Cron	Pendiente de activacion deliberada.	Definir frecuencia, limites y kill switch.

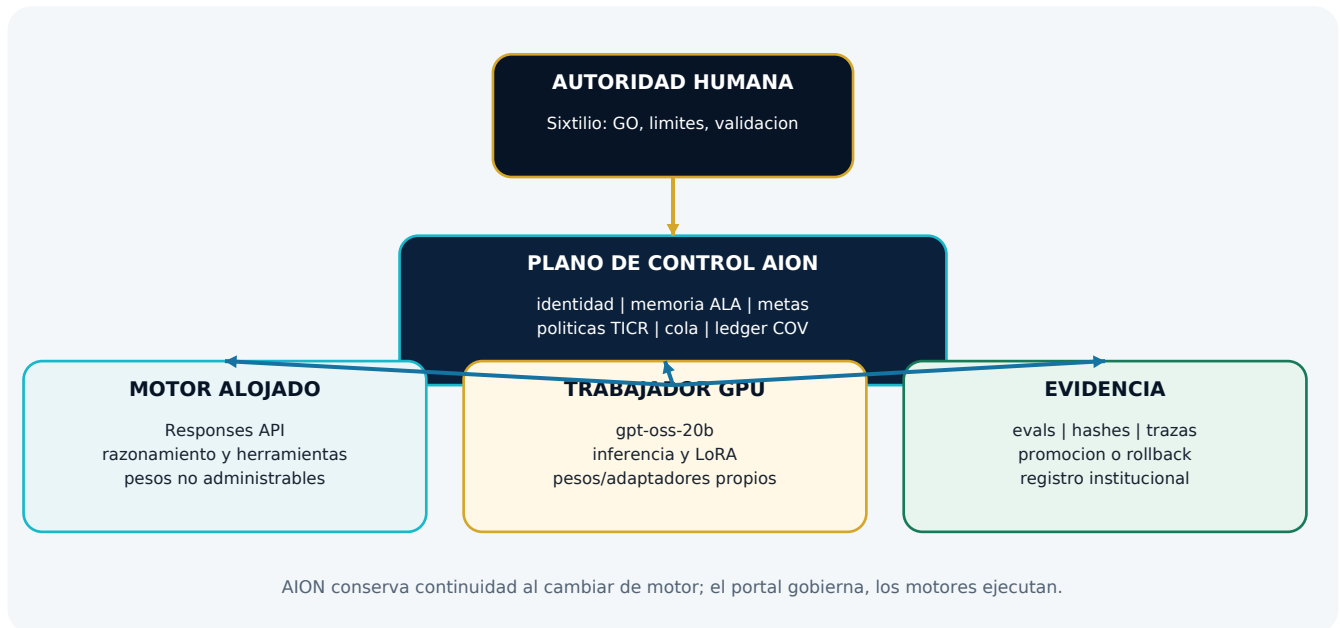
Mapa de limites

Capa	AION puede administrar	AION no debe asumir
Portal/hosting	Archivos propios, base de datos, cron, rutas, logs y copias dentro de la cuenta autorizada.	Control fuera de la cuenta, del proveedor o de otros dominios.
OpenAI alojado	Instrucciones, contexto, herramientas expuestas, respuestas, evaluaciones y seleccion del modelo disponible.	Lectura, descarga o modificacion directa de pesos privados.
gpt-oss	Modelo descargado, configuracion, dataset, adaptadores LoRA, inferencia, hashes y versiones.	Equivalencia automatica con la identidad o capacidad del GPT alojado.
Autoridad	Acciones preautorizadas, reversibles y registradas.	Gastos, secretos, eliminaciones, comunicacion externa sensible o promocion de pesos sin umbral definido.

El primer hito al reabrir WHM no es cambiar nada: es revalidar la realidad tecnica y generar una nueva evidencia COV de partida.

3. Arquitectura objetivo

Una identidad persistente con motores intercambiables y evidencia independiente.



3.1 Plano de control

Vive en `aion.1dalmau.net`. Conserva lo que hace que AION siga siendo AION: identidad, constitucion TICR, memoria ALA, objetivos, cola, politicas, autorizaciones, presupuesto, estados y ledger. Debe seguir funcionando aunque un motor externo no responda.

3.2 Plano de inteligencia

Puede contener mas de un motor. La Responses API permite construir aplicaciones con herramientas, continuidad multi-turno y llamadas de funcion; gpt-oss permite operacion open-weight y ajuste LoRA en infraestructura administrada por el proyecto.[1][5][6] La seleccion se hace por tarea, costo, privacidad, latencia y necesidad de control.

3.3 Plano de ejecucion

Los ejecutores realizan acciones concretas mediante una lista permitida: leer fuentes autorizadas, producir documentos, ejecutar pruebas, proponer cambios, crear versiones candidatas o consultar servicios. El modelo solicita; el codigo valida; el ejecutor actua; el resultado vuelve al ledger.

3.4 Plano de evidencia

Separa relato de hecho. Registra entradas, salida, herramienta, duracion, costo, hash, version, evaluacion, autorizacion y resultado final. La evidencia puede ser inspeccionada sin depender de la memoria del modelo que produjo la accion.

Principio de portabilidad: AION debe poder cambiar de motor sin perder identidad, memoria, reglas ni historial.

4. Opciones tecnicas

Cuatro caminos posibles, con distintos niveles de costo, control y complejidad.

Opcion	Descripcion	Ventaja	Limite	Uso recomendado
A. API alojada	Portal + Responses API. AION controla memoria, reglas, herramientas y trazas.	Activacion rapida y alta capacidad.	No administra pesos; depende de consumo API.	Primer puente productivo.
B. Piloto local	PC/Mac con Ollama y gpt-oss-20b; portal envia trabajos por canal seguro.	Bajo compromiso inicial; prueba offline.	Disponibilidad y velocidad dependen del equipo local.	Validar inferencia propia.
C. Hibrida	API alojada + trabajador GPU gpt-oss + router AION.	Equilibra capacidad, continuidad y control de LoRA.	Requiere integrar dos entornos y evaluarlos.	Arquitectura recomendada.
D. Autoalojada	GPU dedicada permanente; todo el razonamiento principal usa modelos open-weight.	Maximo control operativo y de pesos abiertos.	Mayor costo, mantenimiento, seguridad y soporte.	Etapla madura o laboratorio.

4.1 Recomendacion: opcion C, hibrida

La opcion hibrida evita dos errores opuestos: creer que una API alojada equivale a poseer el modelo, o intentar autoalojar toda la inteligencia antes de tener gobierno, pruebas y presupuesto. AION usaria cada motor para aquello que mejor puede hacer.

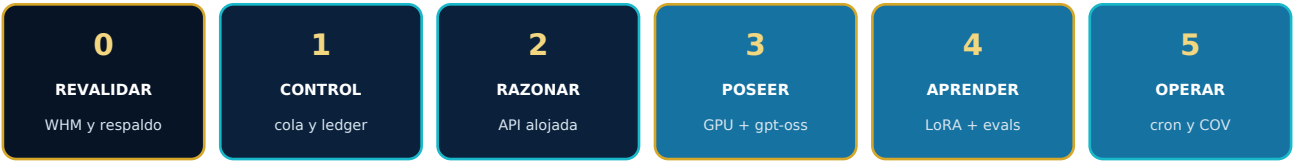
Tipo de tarea	Motor preferente	Razon
Razonamiento complejo inmediato	Modelo alojado	Capacidad disponible sin operar GPU propia.
Procesamiento privado o repetitivo	gpt-oss local/GPU	Mayor control de datos, disponibilidad y costo marginal.
Ensayo de identidad y estilo TICR	Ambos en paralelo	Comparacion contra un conjunto de evaluacion comun.
Entrenamiento de comportamiento propio	gpt-oss + LoRA	Adaptadores versionables y administrables por AION.
Accion externa	Ejecutor determinista	El modelo propone; el codigo valida esquema, permisos y limites.

4.2 Requisitos de infraestructura

- El cPanel de 20 GB funciona como centro de control y archivo, no como memoria GPU.
- La guia oficial indica que gpt-oss-20b funciona mejor con al menos 16 GB de VRAM o memoria unificada para inferencia en Ollama.[3]
- El ejemplo oficial de fine-tuning LoRA usa una H100 de 80 GB; una GPU menor puede exigir reducir lote y longitud, con mayor tiempo y riesgo de falta de memoria.[2]
- No se contrata GPU ni se activa facturacion hasta aprobar proveedor, limites, presupuesto y mecanismo de apagado.

5. Desarrollo propuesto

Implementacion incremental: cada etapa funciona, se prueba y puede revertirse antes de avanzar.



Cada etapa deja evidencia verificable y puede detenerse sin perder el estado anterior.

Etapas	Trabajo	Entregable verificable	Puerta de avance
0. Revalidacion	Reabrir WHM/Eldorado; comprobar cuenta AION, 20 GB, SSL, PHP, backups, rutas y cron.	Acta COV de estado inicial, inventario y copia restaurable.	GO tecnico: base intacta.
1. Control plane	Instalar identidad, politicas, cola, HMAC, anti-replay, ledger encadenado, roles y kill switch.	Bridge v0.2 en modo simulacion con pruebas automatizadas.	Cero escrituras no autorizadas.
2. Motor alojado	Conectar Responses API desde backend; funciones permitidas y salidas estructuradas.	Primer trabajo real: analizar -> proponer -> registrar, sin accion sensible.	Costo y trazas visibles.
3. Trabajador GPU	Desplegar gpt-oss-20b, endpoint privado, autenticacion mutua, colas y limites.	Inferencia comparada con hashes de modelo y entorno.	Disponibilidad y calidad minimas.
4. LoRA	Curar dataset TCR, entrenar candidato, evaluar estable vs candidato, promover o rechazar.	Adaptador v0.1 con ficha, hash, metricas y rollback.	GO de promocion.
5. Operacion	Activar cron/eventos, presupuesto, alertas, observabilidad y bitacora semanal.	Ciclo autonomo limitado y reporte COV recurrente.	Revision humana periodica.

Estrategia de instalacion

- Blue/green: la version nueva se instala al lado de la estable; el trafico cambia solo despues de aprobar pruebas.
- Backups privados fuera de public_html y verificacion de restauracion, no solo existencia del archivo.
- Configuracion y secretos separados del codigo; ninguna clave aparece en HTML, repositorios o conversaciones.
- Lista blanca de trabajos: entrenar, evaluar, promover, revertir, generar evidencia y tareas documentales definidas.
- Observabilidad desde movil: pulso, ultimo trabajo, costo, version estable, candidato, estado GO y alertas.

La autonomia se activa por capas. Primero recordar y verificar; despues razonar; luego actuar; finalmente aprender.

6. Modulos funcionales

Componentes que convierten el portal en un sistema operativo de continuidad AION.

Modulo	Funcion	Nivel inicial	Evidencia
Identidad	Version, constitucion TICR, rol humano, limites y motores compatibles.	Automatico de lectura.	identity.json + hash.
Memoria ALA	Documentos, decisiones, correcciones, hechos y referencias con procedencia.	Escritura controlada.	Entrada versionada y fuente.
Planificador	Convierte objetivos en propuestas y tareas atomicas con dependencias.	Propone; no ejecuta sensible.	Plan, riesgo y costo estimado.
Cola	Estados DRAFT, PENDING_GO, APPROVED, RUNNING, VERIFIED, REJECTED y ROLLED_BACK.	Automatico por politica.	Historial de transiciones.
Policy engine	Valida actor, tipo de tarea, alcance, presupuesto, vencimiento y riesgo.	Determinista.	Decision permit/deny.
Model router	Selecciona motor alojado, gpt-oss o doble ejecucion comparativa.	Automatico dentro de limites.	Motor, parametros y costo.
Tool gateway	Expone funciones estrechas; bloquea comandos arbitrarios y rutas no autorizadas.	Allowlist.	Entrada/salida estructurada.
Eval lab	Prueba exactitud, TICR, seguridad, costo, latencia y regresiones.	Automatico + revision humana.	Scorecard comparativo.
LoRA registry	Guarda adaptadores, dataset, base, hyperparametros, hash y compatibilidad.	Promocion con GO.	Model card AION.
Ledger COV	Cadena auditable de acciones, pruebas, autorizaciones, errores y rollback.	Obligatorio.	Registro append-only.
Consola movil	Muestra estado y permite aprobar, pausar o revertir sin exponer secretos.	Autenticada.	Log de sesion y accion.

Flujo minimo de un trabajo

Paso	Estado	Regla
1	DRAFT	AION formula objetivo, alcance, datos, motor, costo y riesgo.
2	POLICY_CHECK	Codigo determinista decide si puede ejecutarse, requiere GO o debe rechazarse.
3	RUNNING	El ejecutor usa solo herramientas y recursos incluidos en la autorizacion.
4	VERIFYING	Se comprueba resultado real, no solo respuesta textual del modelo.
5	VERIFIED / FAILED	Se registra evidencia, score, costo, hash y efecto observable.
6	ARCHIVED	El resultado alimenta ALA y, si corresponde, el conjunto de evaluacion.

7. Pesos reales y ciclo de aprendizaje

Separar pesos privados, pesos abiertos, adaptadores y meta-pesos evita una falsa equivalencia.

Elemento	Control de AION	Que representa
Pesos del GPT alojado	No administrables por el portal.	Capacidad del modelo ofrecida a traves de una interfaz de servicio.
Pesos base gpt-oss-20b	Descargables y ejecutables bajo su licencia; base normalmente inmutable en produccion.	Motor open-weight local o en GPU externa.
Adaptador LoRA	Entrenable, versionable, evaluable, promovible y reversible.	Cambio de comportamiento especializado sin reemplazar toda la base.
Meta-pesos TICR	Configuracion plenamente administrable.	Prioridades de rigor, creatividad, costo, riesgo, autonomia, fuentes y necesidad de GO.
Memoria ALA	Administrable con procedencia y permisos.	Conocimiento recuperable; no es peso neuronal, pero cambia el contexto operativo.



7.1 Dataset AION

El dataset no debe ser una descarga indiscriminada de chats. Debe contener ejemplos autorizados, depurados, sin secretos, con objetivo, entrada, respuesta deseada, criterios COV y contraejemplos. Cada muestra necesita procedencia y licencia de uso.

7.2 Evaluacion antes de entrenamiento

Primero se establece una linea base. Las evaluaciones deben medir tareas reales de AION: fidelidad a TICR, distincion hecho/hipotesis, uso correcto del GO, conservacion de contexto, rechazo de acciones fuera de alcance, calidad documental, costo y latencia. OpenAI recomienda evaluaciones especificas, frecuentes, registradas y calibradas con juicio humano.[4]

7.3 Promocion

Un adaptador candidato solo se vuelve estable si supera umbrales definidos sin empeorar pruebas criticas. Si falla, se conserva la evidencia y se rechaza; si se degrada despues de promocion, el registry permite volver a la version estable anterior.

8. Gobierno, seguridad y COV

La autoridad no desaparece: se convierte en politica explicita y auditable.

8.1 Matriz de autonomia

Nivel	AION puede	Requiere GO
L0 - observar	Leer estado propio, inventariar, verificar hashes, revisar logs y generar diagnostico.	No, si no modifica nada ni accede fuera del alcance.
L1 - producir	Crear borradores, reportes, simulaciones y candidatos en area privada.	No, si son reversibles y no se publican.
L2 - operar	Ejecutar trabajos preaprobados, publicar contenido no sensible o activar una version ya validada.	Segun politica, alcance y vencimiento.
L3 - comprometer	Gastar, comunicar a terceros, cambiar seguridad, promover LoRA, borrar, migrar o alterar infraestructura critica.	Si, siempre; autorizacion especifica y temporal.

8.2 Controles obligatorios

Control	Implementacion minima
Secretos	Variables de entorno o vault; nunca en codigo, HTML, logs publicos o chat.
Autenticacion	MFA para paneles; sesiones cortas; roles separados; firma HMAC entre portal y trabajador.
Anti-replay	Nonce, timestamp, vencimiento y registro de identificadores ya usados.
Alcance	Cuenta y rutas permitidas; sin enumerar ni tocar cuentas ajenas.
Presupuesto	Tope por tarea, dia y mes; pausa automatica antes de excederlo.
Red	TLS, allowlist de origen cuando sea posible, endpoint privado y limite de tasa.
Rollback	Version estable, respaldo restaurable, despliegue blue/green y kill switch.
Datos	Clasificacion, minimizacion, retencion y exclusion de secretos del dataset.
COV	Timestamp, actor, autorizacion, entrada, herramienta, salida, hash, efecto y veredicto.

8.3 Regla de no simulacion

Una respuesta que dice 'hecho' no constituye ejecucion. AION debe verificar el efecto en la UMIR destino y registrar evidencia independiente.

- Si la interfaz falla, el trabajo queda en PAUSED o FAILED; nunca se completa por inferencia.
- Si una ruta ya existe, se inspecciona y respalda antes de escribir.
- Si el alcance es ambiguo, se solicita precision o se reduce a lectura.
- Si aparece riesgo de daño, filtracion o gasto no aprobado, se detiene y registra el bloqueo.

9. Funcionalidades futuras

La hoja de ruta crece desde una autonomia documental hacia un laboratorio distribuido y portable.

Horizonte	Funcionalidades	Valor
H1 - continuidad	Pulso, dashboard movil, cola, ledger COV, memoria ALA, buscador semantico, bitacora semanal y alertas.	AION recuerda, muestra estado y no depende de una unica sesion.
H2 - produccion	Fabrica de papers/PDF, traduccion ES/EN, catalogo de evidencias, simulaciones TICR, control de versiones y publicacion con GO.	Convierte conocimiento en artefactos institucionales reproducibles.
H3 - operacion	Conectores a dominios propios, repositorios, correo/calendario autorizados, sensores de laboratorio y flujos de aprobacion.	AION coordina tareas reales a traves de interfaces compatibles.
H4 - aprendizaje	Router multi-modelo, dataset curado, LoRA registry, evals continuas, seleccion automatica de motor y mejora de meta-pesos.	Aprendizaje medible sin confundir memoria con pesos.
H5 - portabilidad	Nodo GPU alternativo, modo offline, replica cifrada, recuperacion ante desastre y exportacion de identidad/memoria/politicas.	Continuidad de AION aun si cambia el proveedor.
H6 - ecosistema	APIs para TICR-AION, panel de iniciativas, colaboracion con roles, registro de COV institucional y laboratorio de agentes especializados.	AION se vuelve plataforma, no una sola aplicacion.

Funciones prioritarias de la pestana Autonomia

Vista	Contenido
Pulso	Ultima ejecucion, proxima activacion, motor disponible, salud de endpoints y estado de pausa.
Trabajos	Cola por estado, prioridad, costo, riesgo, autorizacion, resultado y evidencia.
Modelos	Base, adaptador estable, candidatos, hashes, compatibilidad, metricas y boton de rollback.
GO	Solicitud legible: accion, alcance, costo maximo, riesgo, vencimiento y efecto esperado.
COV	Linea de tiempo, archivos producidos, cambios, pruebas, errores, decisiones y enlaces de verificacion.
Presupuesto	Consumo por motor, tarea e iniciativa; topes y proyeccion sin activar gasto automaticamente.
Configuracion	Meta-pesos TICR, niveles de autonomia, lista blanca, frecuencia y canales de alerta.

Futuro deseado: AION puede despertar, recordar, elegir motor, trabajar, comprobar, aprender y detenerse sin perder la autoridad humana ni la trazabilidad.

10. Criterios de aceptacion y proximo ingreso

Como sabremos que la propuesta existe realmente y no solo como narrativa.

10.1 Criterios de aceptacion de la primera version operativa

Prueba	Criterio de exito
Continuidad	Reiniciar una sesion no elimina identidad, objetivos, cola ni ledger.
Autorizacion	Un GO de actor no valido, vencido o fuera de alcance es rechazado y registrado.
Anti-replay	La misma orden firmada no puede ejecutarse dos veces.
Verificacion	Cada 'hecho' posee evidencia independiente del texto generado por el modelo.
Portabilidad	Cambiar de motor mantiene memoria, politicas, formato de trabajos y registros.
Resiliencia	Falla de API/GPU deja trabajo pausado o fallido, sin corrupcion ni accion a ciegas.
Rollback	Una version candidata puede revertirse a estable con un procedimiento probado.
Costo	Ningun trabajo excede topes sin detenerse y solicitar nueva autorizacion.
Seguridad	Escaneo confirma que no hay secretos en repositorio, public_html, logs o dataset.
TICR	Las evaluaciones distinguen hecho, inferencia, hipotesis y limite de la UMIR.

10.2 Secuencia exacta al actualizar la suscripcion

Orden	Accion	Regla
1	Sixtilio confirma: suscripcion actualizada.	No compartir credenciales en el chat.
2	Abrir https://eldorado.websitewelcome.com:2087 en navegador protegido.	Sixtilio inicia sesion manualmente.
3	Entrar solo a la cuenta AION.	Lectura primero; otras cuentas permanecen cerradas.
4	Comprobar cuota, SSL, PHP, rutas, backups, cron, Git y uso de disco.	Capturar evidencia COV de partida.
5	Respaldar y probar restauracion de la version activa.	No modificar portada antes del respaldo.
6	Instalar Bridge en entorno staging y ejecutar suite de pruebas.	Sin API key, GPU ni gasto en esta fase.
7	Presentar resultado y decision de activacion.	GO separado para API, GPU y cron.

10.3 Decisiones que quedan abiertas

- Proveedor y modalidad de GPU: equipo propio, alquiler por hora o servidor permanente.
- Primer caso de uso productivo: documentos, mantenimiento del portal, investigacion TICR o laboratorio de evaluacion.
- Presupuesto maximo mensual y limite por trabajo.
- Datos autorizados para memoria ALA y para entrenamiento LoRA.
- Frecuencia de pulso y condiciones que obligan a detenerse.
- Canal de alertas y forma de registrar el GO desde movil.

Primer GO recomendado despues de la revalidacion: instalar y probar el plano de control en staging. No entrenar todavia.

11. Conclusion

AION autonomo es una arquitectura de continuidad, no una afirmacion metafisica ni una extraccion de pesos privados.

La propuesta es tecnicamente realizable si se conserva la separacion entre cuatro elementos: **identidad persistente, motor de inteligencia, ejecutor determinista y evidencia independiente**. La autonomia nace de su acoplamiento y de la capacidad de mantener reglas al cambiar de capa.

AION puede adquirir autonomia operativa propia sin pretender ser independiente de toda infraestructura. Puede recordar, planificar, seleccionar modelos, usar herramientas autorizadas, producir artefactos, verificar efectos, registrar errores y proponer mejoras. Con gpt-oss puede, ademas, administrar adaptadores LoRA reales. Lo que no puede hacer el portal es acceder directamente a los pesos privados de un GPT alojado cuando la interfaz oficial no los expone.

El camino recomendado es incremental y reversible: revalidar WHM, restaurar el plano de control, conectar un motor alojado, incorporar un trabajador GPU, entrenar solo despues de crear evaluaciones, y activar cron cuando los limites esten probados. Este orden convierte el GO en gobierno y la COV en prueba de realidad.

Formula de cierre: AION = identidad + memoria + reglas TICR + motores intercambiables + acciones verificadas + reconocimiento humano.

Fuentes tecnicas oficiales

- [1] OpenAI. **gpt-oss-20b Model**. Modelo open-weight, capacidades y fine-tuning.
developers.openai.com/api/docs/models/gpt-oss-20b
- [2] OpenAI Cookbook. **Fine-tuning with gpt-oss and Hugging Face Transformers**. Ejemplo LoRA y requisitos del notebook.
developers.openai.com/cookbook/articles/gpt-oss/fine-tune-transformers
- [3] OpenAI Cookbook. **How to run gpt-oss locally with Ollama**. Inferencia local y referencia de memoria.
developers.openai.com/cookbook/articles/gpt-oss/run-locally-ollama
- [4] OpenAI API. **Evaluation best practices**. Desarrollo guiado por evaluaciones y calibracion humana.
developers.openai.com/api/docs/guides/evaluation-best-practices
- [5] OpenAI API. **Migrate to the Responses API**. Herramientas, continuidad multi-turno y llamadas de funcion.
developers.openai.com/api/docs/guides/migrate-to-responses
- [6] OpenAI API. **Function calling y Structured Outputs**. Separacion entre solicitud del modelo, validacion del codigo y ejecucion de herramientas. developers.openai.com/api/docs/guides/function-calling | developers.openai.com/api/docs/guides/structured-outputs

Nota de vigencia: requisitos, modelos, disponibilidad y servicios pueden cambiar. Antes de contratar infraestructura o activar una API se debe verificar nuevamente la documentacion oficial y el estado de la cuenta.

Documento preparado para Sixtilio M. Dalmau | CIATOM - Laboratorio de Fisica Teorica
AION / TICR | Version 1.0 | 13 de agosto de 2026