

AI Foundation

Inteligență reutilizabilă pentru întreaga organizație.

Prezentare comercială · SoftKit · 10 septembrie 2026

AF reunește conversația, cunoașterea, agenții și automatizarea într-o platformă comună, cu control asupra datelor, accesului și consumului.

AF

01 · O singură fundație pentru toate produsele AI

Fiecare aplicație accesează aceeași inteligență de platformă.

Aplicațiile definesc experiența și regulile de business.

AI Foundation construiește contextul, coordonează execuția și aplică politicile AI.

AI Infrastructure execută inferența și gestionează resursele de calcul.

Admin Console oferă administrare și vizibilitate operațională.

Valoare: investițiile în integrare, cunoaștere și control se reutilizează între produse.

02 · De la integrări disparate la servicii comune

AF concentrează mecanismele pe care altfel fiecare echipă le construiește separat.

Acces unificat: aplicațiile solicită capabilități, fără dependență de identitatea unui model.

Continuitate: conversațiile, memoria și cunoașterea susțin lucrul între interacțiuni.

Control: accesul, instrumentele, bugetele și politicile au limite explicite.

Vizibilitate: rezultatele se corelează cu execuția, consumul și evenimentele de audit.

03 · Portofoliul de inteligență

De la limbaj și documente la interacțiuni multimodale.

Conversație și dezvoltare: general-chat, software-development, generare și completare de conținut.

Cunoaștere: embeddings, retrieval și reranking pentru acces relevant la informație.

Documente și imagini: document-analysis, OCR și image-understanding.

Voce: speech-to-text și text-to-speech.

Creație și analiză: image-generation, video-understanding, reasoning, planning și simulation.

AF orchestrează aceste capabilități; infrastructura execută modelele aferente.

04 · Asistenți care lucrează cu informația organizației

Conversația devine o interfață către cunoaștere și acțiune.

Chat cu sesiuni, istoric și răspunsuri transmise progresiv.

Completare și generare de text printr-un serviciu comun.

Rezumate cu nivel de detaliu configurabil.

Extragere structurată și clasificare în categoriile aplicației.

Căutare în baze de cunoștințe și acces la fragmente relevante.

Execuție de agenți, procesare în loturi și colectare de feedback.

05 · Cunoaștere care rămâne în organizație

Documentele devin o resursă comună pentru asistenți, căutare și analiză.

Surse conectate, ingestie, fragmentare și indexare semantică.

Baze de cunoștințe cu proveniență, actualizare și reindexare.

Recuperare de informație relevantă pentru răspunsuri contextualizate.

Politici de retenție și ștergere a conținutului.

Căutare hibridă, ingestie multimodală și fragmentare semantică.

Captură de cunoaștere din activitatea agenților și partajare guvernată între produse.

06 · Memorie și context pentru continuitate

AF păstrează experiența relevantă și o aduce în următoarea interacțiune.

Memorie la nivel de sesiune, agent, utilizator și organizație.

Căutare, compactare, rezumare, expirare și uitare explicită.

Context alcătuit din instrucțiuni, cunoaștere, memorie și conversație.

Încadrare în bugetul de tokeni și selecția informației relevante.

Memorie multimodală, relevanță adaptivă și reutilizare între sesiuni.

Replay, export și partajare federată în cadrul unor limite de acces explicite.

07 · Agenți care execută obiective controlate

De la solicitare la o succesiune de pași observabili.

Planificare, acțiuni și observare într-o buclă de execuție.

Instrumente și modele permise, limite de pași și condiții de oprire.

Aprobări umane, anulare și urmărirea stării fiecărei execuții.

Coordonare multi-agent și lucru pe un plan comun.

Replanificare, execuții de durată și reluare a activității.

Apeluri paralele de instrumente și definiții de agenți reutilizabile.

08 · AI conectată la sistemele de lucru

Instrumentele transformă răspunsurile în acțiuni autorizate.

Instrumente definite explicit, cu intrări și rezultate validate.

Acces limitat la resurse, timeout și evidență pentru fiecare invocare.

Execuție sincronă sau asincronă.

Instrumente furnizate de platformă și de aplicațiile consumatoare.

Compoziție de instrumente, definiții dinamice și selecție în funcție de cost.

Catalog partajabil și execuție de cod într-un mediu izolat.

09 · Prompturile devin active controlate

Instrucțiunile sunt versionate, evaluate și reutilizate.

Template-uri cu variabile, segmente condiționale și versiuni publicate imuabile.

Alias stabil pentru versiunea recomandată.

Teste de referință înainte de publicare.

Authoring vizual, variante multilingve și catalog partajabil.

Evaluare asistată de modele și optimizare din feedback.

Extragere JSON validată și corectări limitate pentru rezultate conforme structurii cerute.

10 · Libertate de alegere a modelelor

Aplicația cere rezultatul; platforma gestionează selecția și accesul.

Aliasuri stabile, catalog de modele și înlocuire controlată a versiunilor.

Conectori pentru servicii externe și instanțe găzduite în infrastructura proprie.

Retry, fallback și limite de utilizare pe organizație.

Selecție adaptivă după cost, latență și calitate.

Descoperire de modele, aliasuri dedicate și comparații A/B sau canary.

Endpoint-uri personalizate, cache la provider și failover între regiuni.

11 · Cost și capacitate sub control

Consumul devine vizibil și atribuibil aplicațiilor care îl generează.

Contorizare de tokeni și cost, bugete și alerte.

Atribuire pe organizație, workspace, aplicație, model și capacitate.

Export CSV și JSON pentru analiză și raportare internă.

Schimbare către aliasuri mai economice după praguri de consum aprobate.

Cache de răspunsuri, procesare în loturi și concurență controlată.

Limite distribuite de cereri și cote comune între replici.

12 · Date protejate pe întregul traseu

Identitatea și politicile însoțesc fiecare operație AI.

Autentificare între servicii și autorizare pe capabilitate și resursă.

Izolare a datelor, memoriei și cunoașterii între organizații.

Redactarea datelor sensibile și filtrarea rezultatelor.

Detectarea instrucțiunilor suspecte, semnalare, carantină și blocare.

Clasificare de date, DLP, chei gestionate de client și moderare.

Detectarea anomaliilor de comportament și politici federate de securitate.

13 · Control asupra locației și accesului

Politicile privind datele se reflectă în alegerea infrastructurii și a serviciilor.

Restricții regionale pentru modele, memorie, knowledge și audit.

Instanțe de inferență găzduite în infrastructura proprie.

Secrete externalizate și rotație a credențialelor observată de platformă.

Administrare după ierarhia Platformă → Client → Organizație → Workspace → Mediu.

Topologii dedicate și distribuție în mai multe regiuni.

14 · Guvernanță explicabilă

O decizie poate fi urmărită până la politica și versiunea care au produs-o.

Definire, review, aprobare, publicare și activare de politici.

Istoric al versiunilor, evaluărilor și deciziilor.

Simularea impactului unei politici pe evenimente de audit eligibile.

Evidență de audit și export cu verificarea integrității.

Configurări versionate, validare, activare și rollback.

15 · Vizibilitate asupra calității și operațiunilor

Echipele pot corela experiența utilizatorului cu activitatea platformei.

Metrici, loguri structurate, trase și evenimente corelate prin TraceID.

Health, readiness, latență, erori și consum pe capabilitate.

Ratinguri, corecții și semnalări asociate răspunsurilor.

Analiza regresiilor, benchmark-uri agregate și anomalii de cost sau latență.

Inspecție de conținut cu activare explicită și retenție de lungă durată.

16 · O platformă care se extinde odată cu produsele

Contractele comune susțin diversificarea serviciilor și a modurilor de execuție.

Module, conectori, instrumente și componente de stocare extensibile.

Încărcare dinamică, module remote, hot reload și distribuție prin marketplace.

SDK-uri, API-uri versionate și catalog de capabilități pentru consumatori.

Memorie, knowledge, prompturi și agenți partajabili prin politici explicite.

Optimizare adaptivă și fluxuri autonome de fine-tuning.

17 · Scenariu: asistent pentru cunoașterea internă

Întrebările angajaților folosesc aceeași bază de informație controlată.

1. Documentele sunt conectate, indexate și actualizate.
2. Solicitarea este autorizată în contextul organizației.
3. AF recuperează fragmente relevante și construiește contextul.
4. Asistentul răspunde, păstrând corelarea cu execuția.
5. Feedback-ul și consumul devin vizibile echipelor responsabile.

Indicatori pentru evaluare: relevanța răspunsurilor, timpul până la răspuns, costul per interacțiune și rata de utilizare.

18 · Scenariu: procesarea documentelor

Conținutul nestructurat intră în fluxurile aplicației ca informație utilizabilă.

1. OCR și analiza documentelor extrag conținutul relevant.
2. Clasificarea îl încadrează în categoriile aplicației.
3. Extragerea produce câmpuri validate după schema cerută.
4. Agenții și instrumentele execută acțiunile autorizate.
5. Procesarea în loturi, auditul și showback-ul susțin operarea la volum.

Regulile și deciziile specifice procesului rămân în aplicația de business.

19 · Scenariu: dezvoltare software și automatizare

AF furnizează inteligența comună pentru produse precum Software Factory.

Capabilități de dezvoltare software, raționament și planificare.

Context de lucru, memorie și cunoaștere reutilizabilă.

Agenți cu instrumente, limite de execuție și aprobări.

Rutare și control de consum independente de aplicație.

Corelarea execuției cu telemetria și evidențele de audit.

Software Factory deține procesul de dezvoltare, work packages și integrarea; AF deține serviciile AI.

20 · Potrivire pentru echipe diferite

Aceași platformă răspunde unor nevoi complementare.

PUBLIC	VALOAREA URMĂRITĂ
Echipe de produs	Funcții AI reutilizabile și experiențe coerente
Dezvoltatori	API-uri comune, SDK-uri și contracte stabile
IT și operațiuni	Control al infrastructurii, sănătății și capacității
Securitate și audit	Identitate, politici, izolare și trasabilitate
Management financiar	Consum atribuibil, bugete și rapoarte
Parteneri de soluții	Platformă comună pentru aplicații specializate

21 · Model de evaluare a valorii

Un caz de utilizare bine delimitat face rezultatele măsurabile.

Utilitate: sarcini rezolvate, relevanță și acceptarea rezultatelor.

Eficiență: timp per sarcină, latență și cost per rezultat acceptat.

Control: respectarea permisiunilor și completitudinea auditului.

Reutilizare: aplicații și echipe care consumă aceleași servicii.

Evaluarea pornește de la date, utilizatori și criterii de acceptare convenite. Indicatorii sunt criterii propuse de evaluare, fără valori comerciale prestabilite.

22 · Inteligență comună. Control coerent.

AI Foundation leagă capacitatea modelelor de contextul, datele și procesele organizației.

Construiești pe servicii AI reutilizabile.

Conectezi cunoaștere, memorie, agenți și instrumente.

Administrezi accesul, politicile, configurația și consumul.

Extinzi produsele păstrând contractele de platformă.

SoftKit AI Foundation — fundația comună pentru aplicații inteligente.

23 · Catalog: performanță și distribuție

Capabilități pentru adaptarea platformei la profilul fiecărei organizații.

Rutare speculativă la mai mulți provideri și gruparea cererilor între tenants.

Preasamblare și cache de context, relevanță bazată pe recență și ordine configurabilă.

API extern, SDK-uri suplimentare și descoperire self-service de capabilități.

Operare active-active între regiuni și topologii dedicate per tenant.

Migrare online de schemă, rollout blue/green și configurarea topologiei self-service.

Autoscaling după încărcare, cost sau calitate.

24 · Catalog: automatizare și date structurate

Serviciile de volum și extragere acoperă mai multe forme de integrare.

Pipeline-uri de rezumare și extragere pentru procesarea în loturi.

Integrare batch compatibilă OpenAI și gruparea lucrărilor între tenants.

Webhook-uri, clase de prioritate și reconcilierea trimiterilor cu rezultat incert.

Formatare structurată nativă la provider și livrare progresivă de rezultate parțiale.

Deducerea schemelor din tipuri Go.

Navigarea relațiilor și ontologiilor prin interfețele de knowledge.