

## 1. Γενική Επισκόπηση

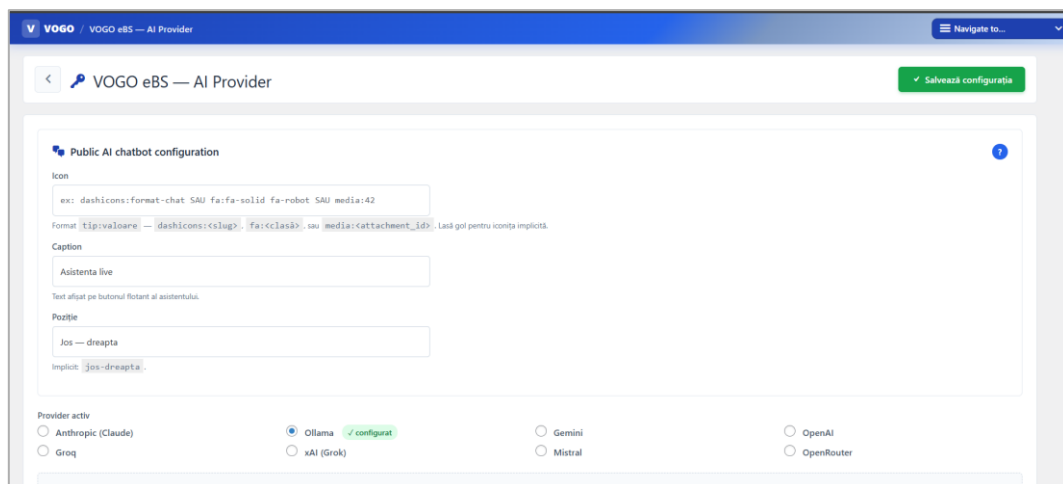
Το VOGO eBS AI Agentic Solutions είναι η ενότητα τεχνητής νοημοσύνης που είναι εγγενώς ενσωματωμένη στην πλατφόρμα VOGO eBS, σχεδιασμένη για την αυτοματοποίηση της επικοινωνίας, την υποστήριξη των χρηστών σε πραγματικό χρόνο και την εννοχήστρωση σύνθετων ροών εργασίας σε επιχειρησιακά περιβάλλοντα. Η λύση συνδυάζει προηγμένα LLM με αρχιτεκτονικές agentic AI — RAG, LangChain, LangGraph και MCP — παρέχοντας στους οργανισμούς προηγμένες δυνατότητες AI χωρίς εξάρτηση από πλατφόρμες τρίτων.

## 2. Τεχνική Αρχιτεκτονική & Στοιβά LLM

### Υποστηριζόμενοι Πάροχοι LLM

Η πλατφόρμα υποστηρίζει τη διαμόρφωση και την εναλλαγή σε πραγματικό χρόνο μεταξύ πολλών παρόχων LLM, χωρίς αλλαγές κώδικα:

- ▶ Anthropic (Claude) — προηγμένη συλλογιστική, εκτεταμένο πλαίσιο, επιχειρησιακή ασφάλεια
- ▶ OpenAI (GPT-4o, GPT-4 Turbo) — μοντέλο αναφοράς για εφαρμογές συνομιλίας
- ▶ Google Gemini — εγγενής ενσωμάτωση με το οικοσύστημα Google Workspace
- ▶ Ollama — εκτέλεση μοντέλων ανοιχτού κώδικα (LLaMA, Mistral, Qwen) τοπικά, on-premise, χωρίς κόστος ανά token
- ▶ Groq — εξαιρετικά γρήγορη εξαγωγή συμπερασμάτων για σενάρια κρίσιμης καθυστέρησης
- ▶ xAI (Grok) — εναλλακτικό μοντέλο για συγκεκριμένες περιπτώσεις χρήσης
- ▶ Mistral — ευρωπαϊκό μοντέλο, κατάλληλο για αυστηρή συμμόρφωση με τον GDPR
- ▶ OpenRouter — ενιαία πύλη πρόσβασης σε δεκάδες μοντέλα μέσω ενός μόνο API



Εικ. 1 — Διαμόρφωση ενεργού παρόχου AI από τη διεπαφή VOGO eBS AI Provider

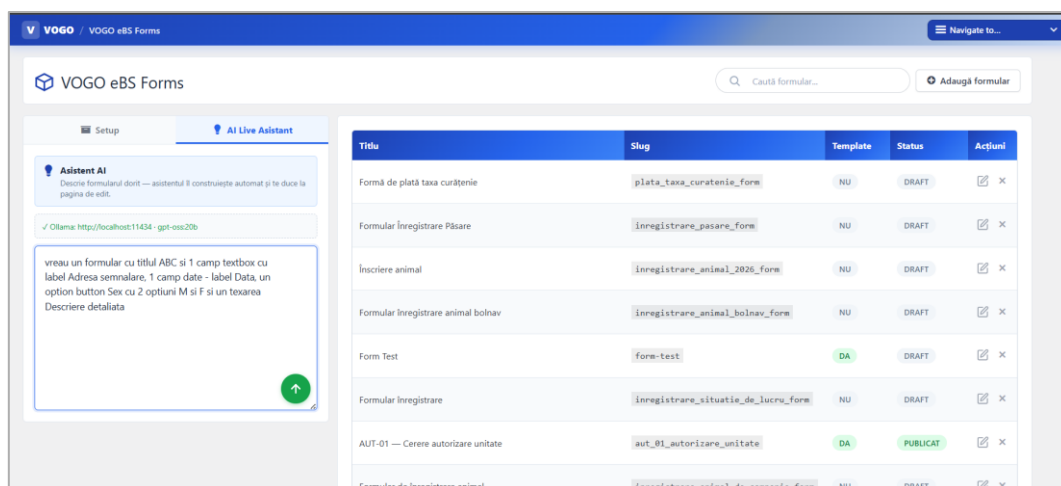
### Αρχιτεκτονικά Στοιχεία

- ▶ LangChain — αφαίρεση παρόχων LLM, RAG/Retrievers, Output Parsers, Tool Use
- ▶ LangGraph — εννοχήστρωση agentic ροών εργασίας με κατάσταση (μηχανές καταστάσεων), βρόχοι, συνθήκες, λογική επανάληψης, Human-in-the-Loop
- ▶ LangSmith — πλήρης παρατηρησιμότητα: ιχνηλάτηση κλήσεων, δοκιμές/αξιολογήσεις, παρακολούθηση παραγωγής, εκδόσεις prompt
- ▶ RAG (Retrieval-Augmented Generation) — ιδιόκτητες βάσεις γνώσης χωρίς επανεκπαίδευση· ανάκτηση των top-K τμημάτων → ακριβής απάντηση με βάση το πλαίσιο
- ▶ MCP (Model Context Protocol) — ανοιχτό πρότυπο για τη σύνδεση πρακτόρων με εργαλεία και εξωτερικές πηγές (ERP, CRM, DMS, ημερολόγια)· Embeddings + Vector Database — σημασιολογική ευρετηρίαση εγγράφων (Pinecone, ChromaDB, pgvector) για άμεση αναζήτηση βάσει πλαισίου

## 3. Λειτουργικές Δυνατότητες

## AI Live Assistant — Βοηθός Ενσωματωμένος στις Φόρμες

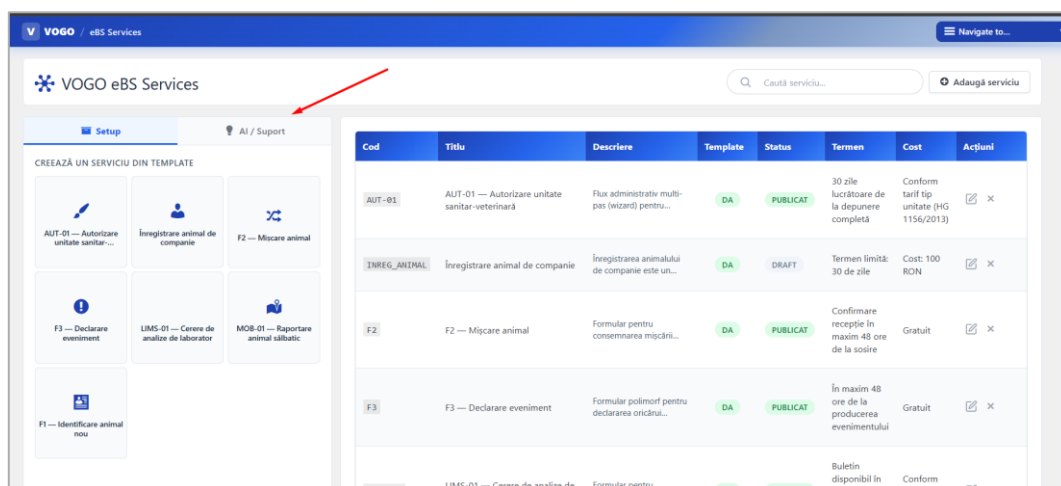
Η πλατφόρμα περιλαμβάνει έναν βοηθό AI εγγενώς ενσωματωμένο στην ενότητα Forms, ο οποίος καθοδηγεί τους χρήστες στη δημιουργία και συμπλήρωση ψηφιακών φορμών σε πραγματικό χρόνο.



Εικ. 2 — To AI Live Assistant ενεργό στην ενότητα VOGO eBS Forms

## AI Live Support — Έξυπνη Υποστήριξη στις Υπηρεσίες

Η ενότητα υπηρεσιών διαθέτει ένα επίπεδο ζωντανής υποστήριξης AI που βοηθά χειριστές και πολίτες στις ροές ψηφιακών υπηρεσιών, μειώνοντας τους χρόνους επίλυσης και τον όγκο των χειροκίνητων αιτημάτων.



Εικ. 3 — To AI Live Support ενεργό στην ενότητα VOGO eBS Services

## Chatbot & Αυτοματοποίηση Επικοινωνίας

- ▶ Chatbot συνομιλιακής AI με αναγνώριση κειμένου και φωνής, με συμπεριφορά παρόμοια με την ανθρώπινη
- ▶ Πλήρης λειτουργία on-premise, χωρίς εξάρτηση από εξωτερικές υπηρεσίες cloud
- ▶ Ιδιόκτητη ενότητα NLU (Natural Language Understanding) για αναγνώριση προθέσεων και οντοτήτων
- ▶ Ενότητα Ερωτήσεων-Απαντήσεων που διαμορφώνεται από τους διαχειριστές χωρίς κώδικα
- ▶ Εξατομικευμένο μήνυμα καλωσορίσματος που αποστέλλεται αυτόματα με το άνοιγμα της συνομιλίας
- ▶ Web εφαρμογή για το κανάλι συνομιλίας, ενσωματώσιμη σε οποιαδήποτε σελίδα HTML
- ▶ Ροές εργασίας Human-in-the-Loop (HITL): ο πράκτορας AI προτείνει, ο χειριστής εγκρίνει ή παρεμβαίνει

## Υποστηριζόμενα Κανάλια Επικοινωνίας

- ▶ Web chat — widget ενσωματώσιμο σε οποιαδήποτε σελίδα HTML
- ▶ E-mail — αυτόματη επεξεργασία και απάντηση
- ▶ SMS — ειδοποιήσεις και αμφίδρομες αλληλεπιδράσεις
- ▶ Μέσα κοινωνικής δικτύωσης — ενσωμάτωση διαμορφώσιμη από τον διαχειριστή (Facebook, WhatsApp κ.ά.)
- ▶ Επεκτάσιμη πολυκαναλική αρχιτεκτονική για πρόσθετα κανάλια

## Διαχείριση Χρηστών & SSO

- ▶ Διαφοροποιημένοι τρόποι αλληλεπίδρασης για εσωτερικούς και εξωτερικούς χρήστες
- ▶ Ανώνυμος εξωτερικός λογαριασμός (βάσει cookie του browser) ή με πιστοποίηση μέσω ονόματος χρήστη/κωδικού· Συγχρονισμός με το υπάρχον σύστημα SSO — οι εσωτερικοί χρήστες δεν διαχειρίζονται πολλούς λογαριασμούς· δημιουργία νέου έργου chatbot διαθέσιμη στον κύριο διαχειριστή της πλατφόρμας
- ▶ Ο διαχειριστής του έργου διαμορφώνει ενσωματώσεις με εξωτερικά συστήματα και κοινωνικά κανάλια

## 4. Προηγμένες Δυνατότητες AI

### Επεξεργασία & Εξαγωγή Εγγράφων

- ▶ Αυτόματη εισαγωγή PDF, Word, HTML, CSV, Notion, Confluence μέσω LangChain Document Loaders
- ▶ OCR για σαρωμένα έγγραφα (Tesseract, AWS Textract, Google Document AI)
- ▶ Έξυπνη τμηματοποίηση: RecursiveCharacterTextSplitter, TokenTextSplitter, SemanticChunker
- ▶ Εξαγωγή μεταδεδομένων για φιλτράρισμα και ανάκτηση βάσει πλαισίου σε επιχειρησιακά συστήματα με χιλιάδες έγγραφα

### Μνήμη & Cache

- ▶ Conversation Memory (Buffer, Window, Summary) — διατήρηση πλαισίου στα chatbots
- ▶ Semantic Memory — σχετικές μνήμες από προηγούμενες συνεδρίες για εξατομίκευση
- ▶ Prompt Caching — έκπτωση 90% (Anthropic) / 75% (Google) σε επαναλαμβανόμενα system prompts
- ▶ LangGraph Checkpointing — μόνιμη μνήμη μεταξύ συνεδριών χωρίς προσαρμοσμένο κώδικα

### Βελτιστοποίηση Κόστους & Απόδοσης

- ▶ Μοτίβα δρομολόγησης — απλά ερωτήματα → ελαφριά μοντέλα· σύνθετα ερωτήματα → premium μοντέλα (μείωση κόστους 60-80%)
- ▶ Async & Batching — παράλληλη επεξεργασία, Anthropic Batch API (έκπτωση 50% έναντι πραγματικού χρόνου)
- ▶ Κβαντοποίηση (GGUF, GPTQ, AWQ) — εκτέλεση μεγάλων μοντέλων σε τυπικό υλικό
- ▶ Structured Output / Pydantic — αυτόματα επικυρωμένες απαντήσεις JSON, χωρίς εύθραυστη ανάλυση

## 5. Ασφάλεια & Συμμόρφωση

- ▶ Ανίχνευση & Απόκρυψη PII (Presidio, AWS Comprehend) — αυτόματη απόκρυψη προσωπικών δεδομένων πριν από την αποστολή σε cloud LLM
- ▶ Guardrails (NeMo Guardrails, LlamaGuard) — διαμορφώσιμοι κανόνες συμπεριφοράς: απαγορευμένα θέματα, ανίχνευση τοξικού περιεχομένου
- ▶ Προστασία από Prompt Injection — εξυγίανση εισόδου, επικύρωση εξόδου, ισχυρό system prompt
- ▶ Rate Limiting & Auth — πιστοποίηση ανά χρήστη, περιορισμός RPM/TPM, πλήρες αρχείο ελέγχου
- ▶ Διαδικασίες επικύρωσης εξόδου — βαθμολόγηση εμπιστοσύνης, ανίχνευση παραισθήσεων, επαλήθευση πηγών
- ▶ Συμμόρφωση με τον GDPR — τοπική επεξεργασία δεδομένων (Ollama on-premise), χωρίς διαρροή δεδομένων

## 6. Επιλογές Εγκατάστασης

|                 |                                                                                                      |
|-----------------|------------------------------------------------------------------------------------------------------|
| Cloud (SaaS)    | Διαχειριζόμενη εγκατάσταση σε υποδομή cloud· αυτόματη κλιμάκωση, μηδενική διαχείριση διακομιστών     |
| On-Premise      | Εγκατάσταση σε ίδια υποδομή· τα δεδομένα δεν εξέρχονται από τον οργανισμό· ιδανικό για αυστηρό GDPR  |
| Υβριδικό        | Τοπικό LLM (Ollama) + ενορχήστρωση στο cloud· ισορροπία κόστους-συμμόρφωσης                          |
| SageMaker (AWS) | Εγκατάσταση fine-tuned μοντέλων στο AWS· αυτόματη κλιμάκωση, Spot Instances, βελτιστοποιημένο κόστος |
| Docker / Lambda | Αρχιτεκτονική serverless· Lambda + endpoint SageMaker· χωρίς διακομιστές προς διαχείριση             |

## 7. Υποστηριζόμενες Ενσωματώσεις

- ▶ VOGO Portal — βοηθός AI διαθέσιμος απευθείας στην πύλη του οργανισμού
- ▶ VOGO Forms — AI Live Assistant για τη δημιουργία και συμπλήρωση φορμών

- ▶ VOGO Services — AI Live Support για ροές ψηφιακών υπηρεσιών
- ▶ DMS (Document Management System) — πρόσβαση σε έγγραφα μέσω web services
- ▶ CRM, ERP, Ticketing — ενσωμάτωση μέσω MCP ή REST API
- ▶ SAP, Oracle eBS, Microsoft D365 — επιχειρησιακοί connectors για λειτουργικά δεδομένα
- ▶ E-mail, SMS, Social Media — πολυκαναλικά κανάλια επικοινωνίας

## 8. Πρόσθετες Λειτουργίες

- ▶ Η ενότητα περιλαμβάνει σύστημα αυτοματοποίησης της επικοινωνίας με τους πολίτες — έναν βοηθό AI που μπορεί να εμφανίζεται τόσο απευθείας στην πύλη του οργανισμού όσο και στην αντίστοιχη εφαρμογή για κινητά (Android ή iOS), η οποία μπορεί να δημοσιευθεί στο Google Play και στο Apple Store
- ▶ Chatbot βασισμένο στην τεχνητή νοημοσύνη, προγραμματισμένο να διεξάγει συνομιλίες όσο το δυνατόν πλησιέστερες στην ανθρώπινη συμπεριφορά, μέσω αναγνώρισης κειμένου και φωνής.
- ▶ Μπορεί να λειτουργεί εξ ολοκλήρου on-premise, χωρίς εξάρτηση από εξωτερικές υπηρεσίες.
- ▶ Παρέχει τη δυνατότητα ενσωμάτωσης με πολλαπλά κανάλια επικοινωνίας.
- ▶ Για το κανάλι web chat, παρέχει web εφαρμογή ενσωματώσιμη σε οποιαδήποτε σελίδα HTML.
- ▶ Επιτρέπει στον διαχειριστή του έργου να ενεργοποιεί και να διαμορφώνει ενσωματώσεις με εξωτερικά συστήματα και κανάλια κοινωνικής δικτύωσης.
- ▶ Επιτρέπει την εύκολη διαμόρφωση και διαχείριση του περιεχομένου του βοηθού, χωρίς κώδικα.
- ▶ Παρέχει πλήρη εργαλεία δοκιμής του chatbot: δοκιμή ροής συνομιλίας, δοκιμή αναγνώρισης προθέσεων και οντοτήτων.
- ▶ Επιτρέπει την αλληλεπίδραση με τη λύση DMS μέσω των web services της.
- ▶ Επιτρέπει την ενσωμάτωση με λύσεις e-mail και SMS, με δυνατότητα επέκτασης σε πρόσθετα κανάλια.
- ▶ Επιτρέπει διαφορετικούς τρόπους αλληλεπίδρασης για εσωτερικούς και εξωτερικούς χρήστες.
- ▶ Επιτρέπει ροές εργασίας Human-in-the-Loop (HITL).
- ▶ Επιτρέπει τον συγχρονισμό με το σύστημα SSO, καταργώντας την ανάγκη διαχείρισης πολλών λογαριασμών για τους εσωτερικούς χρήστες.
- ▶ Περιλαμβάνει ειδική ενότητα LLM.
- ▶ Περιέχει ιδιόκτητη ενότητα Natural Language Understanding (NLU), βασισμένη στην τεχνητή νοημοσύνη, που αναγνωρίζει προθέσεις και οντότητες που εκφράζονται ελεύθερα από τους χρήστες.
- ▶ Επιτρέπει τη διαμόρφωση μηνύματος καλωσορίσματος που αποστέλλεται αυτόματα με το άνοιγμα του παραθύρου συνομιλίας. Επιτρέπει τη διαμόρφωση προκαθορισμένων, ιεραρχικών ερωτήσεων και απαντήσεων, σε μοντέλο σεναρίου συνομιλίας.
- ▶ Επιτρέπει την ενσωμάτωση με καθιερωμένες λύσεις LLM, ενδεικτικά και όχι περιοριστικά: LLaMA, OpenAI, Gemini.
- ▶ Επιτρέπει σε κάθε εξωτερικό χρήστη να δημιουργεί ανώνυμο λογαριασμό (βάσει cookie του browser) ή λογαριασμό με πιστοποίηση μέσω δικού του ονόματος χρήστη και κωδικού.