JJ5 - Per chi amministra

- JJ5.1 - Accendere la chat
- JJ5.2 - Tetto di spesa e riservatezza
- JJ5.3 - Se qualcosa non va

JJ5.1 - Accendere la chat

In una frase. L'assistente si accende in due tempi: prima il servizio con i suoi dati e il catalogo, poi, con un interruttore sul Gateway, la chat.

L'assistente è spento finché non lo accendi tu. I due tempi servono a non mostrare a nessuno una chat che risponde su dati vuoti: prima si preparano e si controllano i dati, poi si apre la porta.

> **Attenzione.** Nell'ambiente di prova interno il servizio, i Dati spiegati e il catalogo sono accesi. La chat è ancora da accendere: l'ultimo passo qui sotto.

1. **Prepara i segreti.** Nel file dei segreti del server metti il blocco dell'assistente: le password dei suoi archivi e la chiave del servizio di intelligenza artificiale. L'interruttore della chat resta spento.
2. **Accendi il servizio.** Con il rilascio del server partono due programmi: l'assistente e il suo archivio. Nessuno dei due è raggiungibile da fuori. Se manca una password, il rilascio si ferma subito e dice quale.
3. **Crea gli archivi vuoti.** Una volta sola: le tavole dei Dati spiegati e quelle del catalogo.
4. **Carica i dati.** Il primo caricamento si lancia a mano. In fondo scrive un riepilogo: controlla che dica «ok». Da lì in poi parte da solo ogni notte.
5. **Riempi il catalogo e fai un rilascio.** Il catalogo rilegge i Dati spiegati e le schede; poi si crea il primo rilascio numerato. Apri il catalogo e controlla che le tessere ci siano.
6. **Accendi la chat.** Metti l'interruttore `SIGEAGENT_ENABLED` a `true` e rilancia il rilascio del server. Il Gateway riparte e manda la chat all'assistente.

Come controlli. Nel registro del Gateway compare «SigeAgent acceso». I due programmi risultano in salute. Nella chat l'etichetta del servizio di intelligenza artificiale è quella dell'assistente.

Per spegnere. Rimetti l'interruttore a `false` e rilancia: la chat torna com'era prima. I dati e il catalogo restano.

> **Nota.** Il catalogo contiene il lavoro delle persone: va salvato con una copia di sicurezza. I Dati spiegati no: si ricostruiscono con un nuovo caricamento.

> **Per chi sviluppa.** Profilo `sigeagent` del compose di deploy, servizi `sigeagent` (porta interna 8765) e `sigeagent-mysql`. Con l'interruttore acceso il Gateway instrada `/api/sigea-cop/ai/agent/stream`, `/ai/agent/fornitore` e `/api/sigeagent/**`. Con l'interruttore acceso ma il profilo spento, la chat dice «AI non abilitata». Comandi, uno per uno: `devops/deploy/README.md`, paragrafo 3.6. Variabili: `docs/devops/ENVIRONMENT.md`.

>

Per approfondire.

- Deploy, paragrafo 3.6 «SigeAgent: l'agente dei dati»: `devops/deploy/README.md`

JJ5.2 - Tetto di spesa e riservatezza

In una frase. L'assistente ha un tetto di spesa al giorno e regole precise su quali dati possono arrivare al servizio di intelligenza artificiale.

Ogni domanda costa qualcosa, perché usa un servizio esterno a pagamento. E ogni domanda manda fuori un po' di testo. Chi amministra deve poter rispondere a due domande: quanto si spende al massimo, e cosa esce.

Il tetto di spesa. L'assistente registra per ogni domanda modello, consumo, costo e durata. Prima di ogni nuova domanda somma la spesa del giorno: se ha raggiunto il tetto, non parte e la chat dice «Per oggi è stato raggiunto il tetto di spesa: riprova domani o chiedi di alzarlo». Il tetto predefinito è 5 dollari al giorno. C'è anche un limite di lavoro per singola domanda e un numero massimo di domande insieme.

Cosa esce verso il modello.

- La domanda, così come l'hai scritta.
- I significati del catalogo: nomi di tavole e colonne, regole.
- Un'anteprima dei risultati delle ricerche sui Dati spiegati: le prime righe.

Cosa non esce.

- I Dati aziendali: il modello non li vede e l'assistente non li legge.
- Dati di persone: i Dati spiegati sono pensati senza email, telefoni e indirizzi.

Dove sta il modello. La regola di ADSWnext per i dati dei clienti è: intelligenza artificiale spenta in partenza, e solo fornitori nell'Unione europea. L'assistente dichiara alla chat, prima di ogni invio, quale servizio usa e in quale regione: se non risponde, la chat non invia.

> **Attenzione.** Oggi l'assistente gira solo nell'ambiente di prova interno, con dati fittizi, e lì usa un instradamento non garantito nell'Unione europea: la chat lo dichiara. Prima di usarlo con i dati di un cliente vanno chiuse due cose: passare all'instradamento europeo e decidere come trattare i nomi scritti nelle domande, che oggi partono senza essere sostituiti.

> **Nota.** L'assistente che nel catalogo aiuta a preparare le proposte ha un tetto di spesa suo, separato.

>

Per chi sviluppa. Tetto: `SIGEAGENT_TETTO_SPESA_GIORNALIERO_USD`. Instradamento europeo: l'impostazione `OPENROUTER_BASE_URL` del servizio va portata a `https://eu.openrouter.ai/api/v1`; il compose di deploy oggi non la passa. La dichiarazione del fornitore è l'endpoint `/ai/agent/fornitore` in `backend/middleware/sigeagent/src/sigeagent/web.py`. Regole generali e l'eccezione dell'assistente: `docs/ai/AI-CONFORMITA.md`.

Per approfondire.

• Informativa sull'intelligenza artificiale per il cliente: `docs/ai/INFORMATIVA-AI-CLIENTE.md`

JJ5.3 - Se qualcosa non va

In una frase. I problemi più comuni con l'assistente e con il catalogo, come «vedo questo, faccio questo», e le parole da sapere.

Quasi sempre l'assistente dice già cosa non va, in italiano. Qui trovi cosa fare dopo averlo letto.

Nella chat

- **La chat non invia, o dice «AI non abilitata».** L'assistente è spento o acceso a metà. Avvisa chi amministra: vedi Accendere la chat.
- **«Per oggi è stato raggiunto il tetto di spesa».** Riprova domani, o chiedi a chi amministra di alzarlo: vedi Tetto di spesa e riservatezza.
- **«Ci sono già troppe domande in corso».** Aspetta qualche secondo e riprova.
- **«Ho raggiunto il limite di lavoro per una domanda».** Quello che vedi è parziale. Fai una domanda più stretta: un periodo, un fornitore, un gruppo di articoli.
- **Dice che i dati non bastano.** Il dato non è nei Dati spiegati. Chiedilo: vedi Chiedere un dato nuovo.
- **Il numero non mi torna.** Leggi nel testo la definizione, il periodo e il deposito usati. Se la definizione è sbagliata, si corregge la regola nel catalogo.
- **I dati sono vecchi.** La risposta dice a che data sono aggiornati. Se è indietro di più giorni, la copia notturna non è riuscita: avvisa chi amministra.

Nel catalogo

- **Ho corretto una scheda ma l'assistente risponde come prima.** Manca il rilascio. La riga in alto dice quante correzioni aspettano: premi **Rilascia** e crea il rilascio.
- **«Chiedi a un collega» non parte.** Scrivi il perché, in una riga.
- **La mia proposta è ferma «in assestamento».** È un dato nuovo: aspetta il lavoro di chi sviluppa.
- **La pagina «Assistente» dice «non configurato».** Manca la sua chiave: avvisa chi amministra.

Per chi amministra

- **Il servizio non parte.** Di solito manca una password del suo archivio: il registro dell'archivio dice quale.
- **Il riepilogo della copia notturna segna un gruppo «saltato».** Nel gestionale mancano quelle tabelle. Gli altri dati sono caricati.

Parole da sapere

- **Assistente dei dati.** Il servizio che risponde alle domande. A schermo, nel catalogo, è «l'agente».
- **Dati spiegati.** La copia ordinata dei dati aziendali, con nomi in italiano. L'assistente legge solo questa.
- **Copia notturna.** Il lavoro che ogni notte riempie i Dati spiegati.
- **Catalogo.** Le schede che dicono cosa significa ogni tavola e ogni colonna.
- **Tavola.** Un argomento dei Dati spiegati: articolo, fornitore, giacenza.
- **Regola.** La definizione scritta di una parola del lavoro, come «acquisto» o «copertura».
- **Stato.** Quanto è sicuro un significato: ignoto, ipotesi, verificato, smentito.
- **Proposta.** La richiesta di cambiare qualcosa nel catalogo. La decide una persona.
- **Rilascio.** Una versione numerata del catalogo. L'assistente usa l'ultima.
- **Tetto di spesa.** Il massimo che l'assistente può spendere in un giorno.

> **Per chi sviluppa.** Esito dei caricamenti notturni: cartella dei log del servizio e tavola `caricamento_etl`. Messaggi della chat: `backend/middleware/sigeagent/src/sigeagent/flusso.py` e `backend/middleware/sigeagent/src/sigeagent/agente.py`. Stati a metà dell'accensione: `devops/deploy/README.md`, paragrafo 3.6.