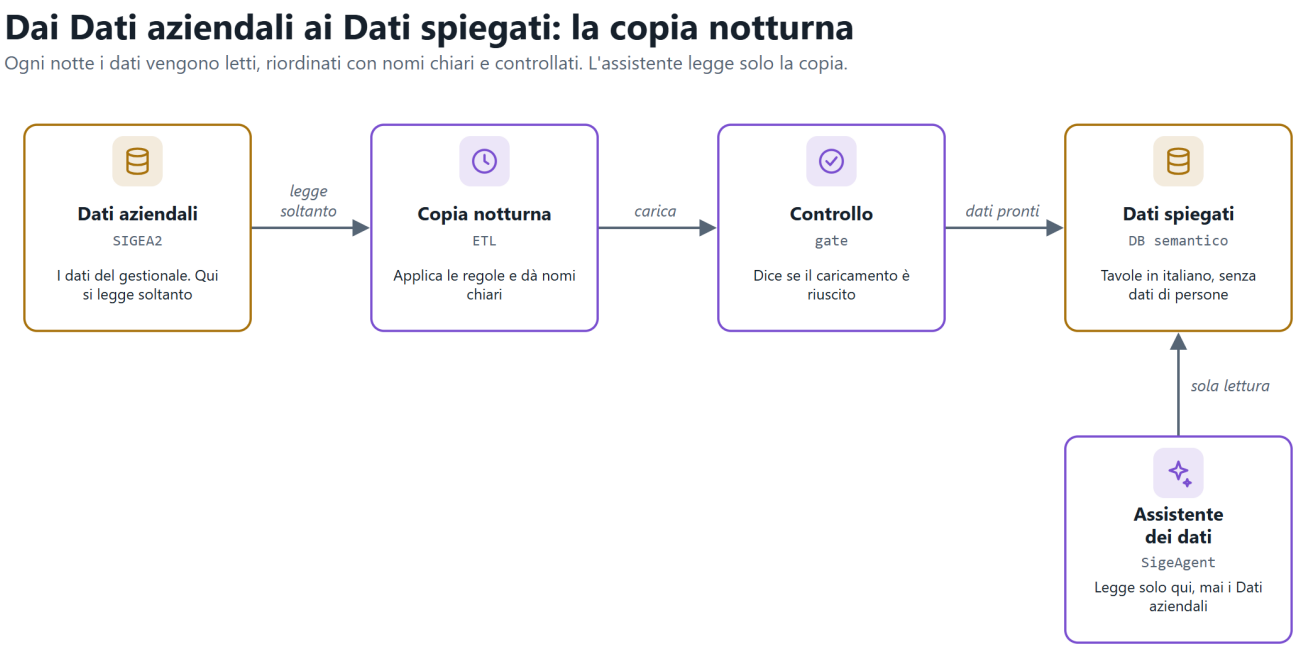
JJ2 - Come ragiona

- JJ2.1 - I dati spiegati: la copia notturna
- JJ2.2 - Cosa sa e cosa non sa
- JJ2.3 - Perché non vede i dati grezzi

JJ2.1 - I dati spiegati: la copia notturna

In una frase. Ogni notte i dati del gestionale vengono copiati in un archivio ordinato, con nomi in italiano: l'assistente legge solo quello.

Le tabelle del gestionale non si spiegano da sole. Hanno nomi in sigla, centinaia di colonne e nessun collegamento dichiarato fra una tabella e l'altra. Chi ci fa una ricerca senza conoscerle sbaglia senza accorgersene. Per questo l'assistente non le tocca: lavora su una copia già riordinata.

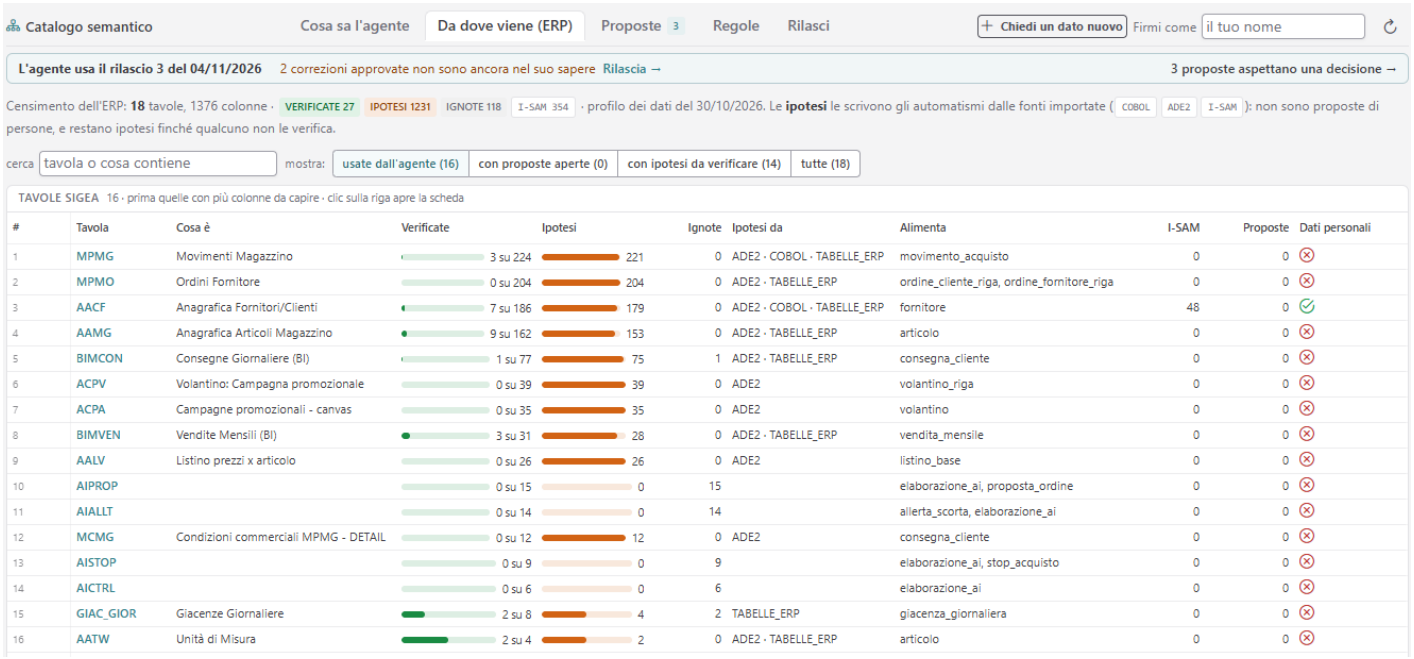


Dai Dati aziendali ai Dati spiegati: la copia notturna

Cosa fa la copia. Legge i Dati aziendali, senza scrivere niente. Tiene solo le righe che in azienda si chiamano davvero con quel nome: per esempio, fra tutti i movimenti di magazzino, solo i carichi da acquisto. Dà a tavole e colonne nomi chiari (`articolo`, `fornitore`, `giacenza_giornaliera`) e scrive i collegamenti fra loro. Le regole sono applicate qui, una volta, e non a ogni domanda.

Quando. La copia parte da sola ogni notte, alle 02:30 se non si sceglie un altro orario. Alla fine scrive un riepilogo: per ogni gruppo di dati dice se è andato bene o se è stato saltato perché nel gestionale quelle tabelle non ci sono.

Come sai se i dati sono freschi. Per ogni tavola resta scritta la data fino a cui i dati sono caricati. L'assistente la conosce e la dichiara nella risposta.



La pagina «Da dove viene (ERP)» del catalogo: per ogni tabella del gestionale, quale tavola dell'assistente alimenta

> **Nota.** I Dati spiegati sono indietro al massimo di un giorno. Per una domanda sul «proprio adesso» vale ancora la schermata del gestionale.

> **Per chi sviluppa.** La copia è l'ETL in `backend/middleware/sigeagent/etl` (`python -m etl tutto`). Legge SIGEA2 con la stessa connessione di SIGEAnext, in transazioni di sola lettura. L'orario è `SIGEAGENT_ETL_ORA_NOTTURNA`. I caricamenti sono registrati nella tavola `caricamento_etl`.

> **Per approfondire.**
> • Lo schema dei Dati spiegati: `backend/middleware/sigeagent/docs/SCHEMA.md`

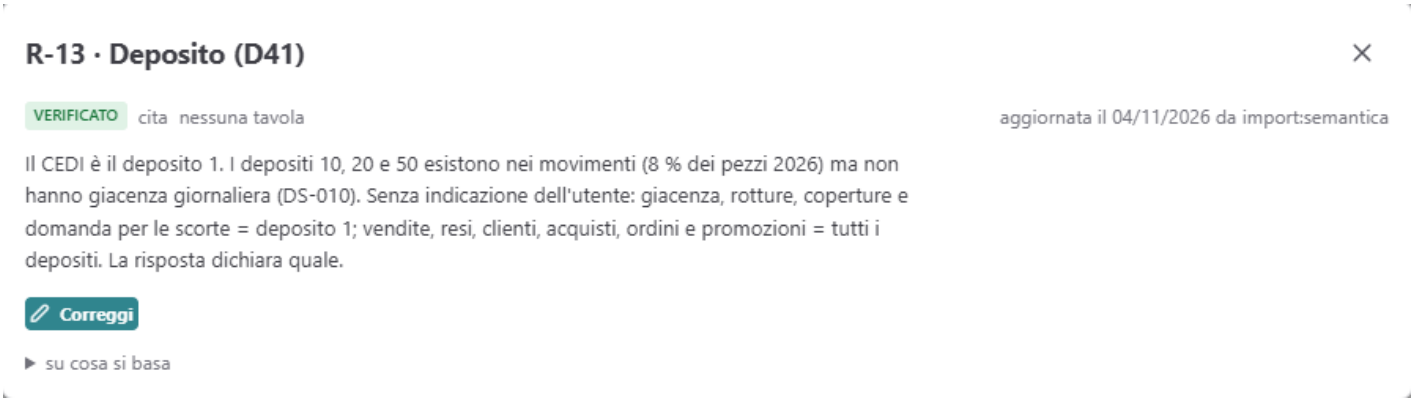
JJ2.2 - Cosa sa e cosa non sa

In una frase. L'assistente sa quello che è scritto nel catalogo: tavole, significati e regole. Quello che non è scritto, non lo inventa.

Un numero sbagliato detto con sicurezza fa più danni di un «non lo so». Per questo l'assistente ha confini chiari, e li dichiara.

Cosa sa. Il primo argomento è il riordino: articoli, fornitori e clienti; acquisti; ordini a fornitori e ordini dei clienti; consegne ai clienti; vendite mensili; giacenze giorno per giorno; volantini e prezzi di listino. Per ognuno il catalogo dice a cosa serve, cosa rappresenta una riga e cosa **non** contiene.

Le regole. Le parole del lavoro hanno una definizione scritta: cosa conta come «acquisto», quando un ordine a fornitore è «aperto», come si misura la «copertura» o la «rottura di stock». Se la tua domanda usa una di queste parole, l'assistente applica la regola e la dichiara in una frase.



Una regola del catalogo aperta: il testo che l'assistente applica, con lo stato «verificato»

Cosa non sa.

- Quello che non è nei Dati spiegati. Se una tavola o una colonna non c'è, l'assistente spiega cosa manca.
- I significati non verificati. Nel catalogo un significato può essere ignoto, un'ipotesi, verificato o smentito. All'assistente arriva solo ciò che è verificato.
- Il presente esatto. I dati sono quelli dell'ultima copia notturna.
- I dati delle persone. I Dati spiegati sono pensati senza email, telefoni e indirizzi.

Quanto può cercare. Per ogni domanda fa poche ricerche mirate, ognuna con un tempo massimo e un numero massimo di righe. Se la domanda è troppo larga, conviene restringerla.

> **Nota.** Un dato che manca si può chiedere: vedi Chiedere un dato nuovo.

Per chi sviluppa. Le regole sono in `backend/middleware/sigeagent/semantica/regole.md` (R-01 e seguenti), le schede delle tavole nella stessa cartella. Limiti di una ricerca: una sola `SELECT`, al massimo 1000 righe, 30 secondi; al massimo 20 passi per domanda (`backend/middleware/sigeagent/src/sigeagent/config.py`).

JJ2.3 - Perché non vede i dati grezzi

In una frase. L'assistente e l'intelligenza artificiale non entrano mai nei Dati aziendali: lavorano solo sui Dati spiegati, in sola lettura.

È una scelta di sicurezza e di qualità insieme. Di sicurezza, perché fra l'assistente e il gestionale non c'è nessuna strada. Di qualità, perché sui dati grezzi le risposte sarebbero spesso sbagliate.

Tre protezioni, una sopra l'altra.

1. **Un archivio a parte.** I Dati spiegati sono separati dai Dati aziendali. L'assistente conosce solo quelli.
2. **Un utente che può solo leggere.** L'assistente entra nei Dati spiegati con un utente a cui scrivere non è permesso.
3. **Un controllo su ogni ricerca.** Prima di eseguirla, l'assistente verifica che sia una sola ricerca di lettura, con un limite di righe. Tutto il resto viene rifiutato.

Cosa vede l'intelligenza artificiale. Riceve la tua domanda e i significati del catalogo, e prepara la ricerca. Poi, per scrivere la risposta, riceve un'anteprima del risultato: le prime righe. La tabella completa arriva alla tua pagina direttamente dall'assistente. Le tabelle del gestionale il modello non le vede mai.

Perché i dati grezzi darebbero risposte peggiori. Nel gestionale il nome di una colonna può ingannare. Una che sembra «marca» è compilata solo su una piccola parte degli articoli: una classifica delle vendite per marca costruita lì sarebbe sbagliata, e nessuno se ne accorgerebbe. Nei Dati spiegati una colonna entra solo quando una persona ne ha verificato il significato.

> **Nota.** La domanda che scrivi parte verso il modello così com'è. Non metterci nomi di persone o altri dati riservati che non servono.

> **Per chi sviluppa.** Utente di database di sola lettura; controllo delle ricerche in `backend/middleware/sigeagent/src/sigeagent/sql_sicuro.py`; il modello vede le prime 50 righe di ogni risultato (`backend/middleware/sigeagent/src/sigeagent/strumenti.py`). Fornitore del modello e spesa: Tetto di spesa e riservatezza.