



UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI ECONOMIA “GIORGIO FUÀ”

Corso di Laurea Magistrale o Specialistica in

DATA SCIENCE PER L'ECONOMIA E LE IMPRESE

**Progettazione di un modello di integrazione dati per la gestione dei rientri
dei componenti**

Design of a data integration model for managing component returns

Relatore: Chiar.mo

Prof. Domenico Potena

Tesi di Laurea di:

Simone Blaso

Correlatore:

Ing. Lucrezia Antenucci

Anno Accademico 2025 – 2026

Sommario

1 Introduzione	3
1.1 Contesto aziendale	4
1.2 Obiettivi.....	8
1.3 Struttura della tesi	9
2 Analisi delle sorgenti dati	11
2.1 La gestione dei rientri in garanzia.....	12
2.2 Il processo diagnostico.....	13
2.3 MEERA.....	15
2.4 Struttura dati MEERA.....	16
2.4.1 Dettaglio tabelle principali	19
2.5 RemoteMaster	22
2.6 Struttura dati di RemoteMaster	24
2.6.1 Dettaglio tabelle principali	28
3 Integrazione	32
3.1 Elementi comuni tra le piattaforme.....	33
3.2 Architettura proposta per l'integrazione.....	35
4 Sperimentazione del modello di integrazione.....	39
4.1 Costruzione del dataset	40
4.2 Analisi esplorativa.....	44
4.3 Standardizzazione delle problematiche.....	48
4.4 Validazione del modello	51
4.4.1 Bronze Layer	53
4.4.2 Silver Layer	53
4.3.3 Gold Layer.....	59
4.5 Commento dei risultati e limiti del modello	63
5 Conclusioni e sviluppi futuri.....	66
5.2 Sviluppi futuri	68
8 Bibliografia	71

Capitolo 1

1 Introduzione

Il presente lavoro di tesi nasce da una collaborazione tra l'Università Politecnica delle Marche e FPT Industrial S.p.A., con l'obiettivo di coniugare esigenze accademiche e necessità operative di un contesto industriale di rilevanza internazionale.

Negli ultimi decenni, le imprese hanno progressivamente affiancato alla centralità del prodotto una crescente attenzione verso il cliente. Il cliente è diventato un riferimento fondamentale non solo nello sviluppo di nuovi prodotti, ma anche nel miglioramento di quelli già esistenti. Questo cambiamento risulta particolarmente evidente nei servizi di assistenza post-vendita, la cui qualità incide direttamente sulla soddisfazione del cliente e sulla sua fedeltà al marchio.

Parallelamente, la crescente digitalizzazione dei processi aziendali ha reso disponibile una quantità sempre maggiore di dati. Se opportunamente raccolti, integrati e analizzati, tali dati rappresentano una risorsa preziosa per le aziende, poiché permettono di comprendere meglio i processi interni, individuare

criticità ricorrenti e supportare decisioni orientate al miglioramento delle attività operative.

Questi due fenomeni trovano una naturale convergenza nel settore dei motori e delle powertrain, in cui la gestione dell'assistenza post-vendita riveste un ruolo strategico. La capacità di raccogliere, integrare e analizzare i dati generati lungo l'intero ciclo di vita del prodotto, dalla segnalazione di un malfunzionamento fino alla sua risoluzione, rappresenta oggi un elemento importante sia per la soddisfazione del cliente sia per il miglioramento continuo della qualità dei prodotti.

In questo scenario si colloca il lavoro di ricerca, sviluppato nell'ambito di un percorso di tirocinio presso FPT Industrial. L'elaborato analizza il processo di gestione dei rientri dei componenti dell'azienda, attualmente supportato da due piattaforme digitali, MEERA e RemoteMaster, che operano su fasi diverse dello stesso processo, con lo scopo di valutare come una migliore integrazione dei dati possa contribuire a rendere il processo più tracciabile, efficiente e analizzabile.

1.1 Contesto aziendale

La FPT Industrial S.p.A., società appartenente a Iveco Group, occupa una posizione di rilievo a livello globale nel campo dei sistemi di propulsione e delle soluzioni energetiche destinate a impieghi industriali. L'attività

dell'azienda non si limita alla semplice produzione di motori, ma comprende l'intero ciclo di vita del prodotto: dalla fase di ricerca e progettazione, fino alla realizzazione, alla distribuzione e al supporto post-vendita. In questo senso, FPT Industrial si configura come un operatore integrato, capace di rispondere a esigenze differenti provenienti dai settori on-road, off-road, marino e della generazione di energia.

La dimensione internazionale dell'azienda è testimoniata da una rete produttiva e di ricerca particolarmente estesa: oltre 8.000 dipendenti lavorano in 10 stabilimenti e 10 Centri di Ricerca e Sviluppo distribuiti in diverse aree del mondo: Italia, Francia, Svizzera, Regno Unito, Stati Uniti, Brasile, Argentina, Cina, India e Corea del Sud. Tale struttura consente non solo di presidiare mercati eterogenei, ma anche di mantenere un costante flusso di innovazione tecnologica.

A questo si aggiunge una rete commerciale e di assistenza tecnica attiva in circa 100 Paesi, che garantisce un supporto capillare ai clienti e rafforza la competitività del marchio a livello globale.

In questo quadro si inserisce anche il cosiddetto sistema di gestione "round-the-clock", che mette

l'eccellenza tecnologica dell'azienda a disposizione dei clienti in ogni momento della giornata, assicurando continuità operativa e rapidità di risposta.

L'offerta di FPT Industrial è estremamente diversificata e comprende sei famiglie di motori con potenze che spaziano da 42 a oltre 1.000 cavalli, trasmissioni con coppia fino a 500 Nm e assali anteriori e posteriori con

capacità da 2,45 a 32 tonnellate. Questa varietà testimonia la capacità dell'azienda di adattarsi a molteplici contesti applicativi, dal trasporto leggero fino alle esigenze più impegnative del settore pesante. Particolare attenzione è rivolta ai motori a gas naturale, che rappresentano una delle aree in cui FPT Industrial ha consolidato una leadership riconosciuta. Con una gamma di potenza compresa tra 50 e 520 cavalli, tali soluzioni costituiscono oggi la proposta più completa disponibile sul mercato. L'impiego del gas naturale non risponde solo a logiche di

efficienza economica, legate al minor costo d'esercizio, ma si inserisce in una più ampia strategia di sostenibilità, grazie a emissioni ridotte e a un impatto acustico significativamente inferiore rispetto ai combustibili tradizionali.

All'interno della gamma, tre linee di prodotto assumono un ruolo centrale. La famiglia F1, destinata principalmente ai veicoli commerciali leggeri, è riconosciuta per l'elevata efficienza e i bassi consumi. Le cilindrata spaziano da 2,3 a 3,0 litri, con potenze comprese tra 85 e 170 kW (116–230 CV) e coppia massima fino a 500 Nm. Questi motori sono disponibili in versioni diesel, a gas naturale e ibride, e si caratterizzano per intervalli di manutenzione prolungati, con conseguente riduzione dei costi di esercizio.

La famiglia NEF rappresenta invece una soluzione versatile, ideale per camion medi e autobus fino a 18 metri. È composta da motori a 4 e 6 cilindri con cilindrata comprese tra 4,5 e 6,7 litri, potenze che vanno da 100 a 220 kW (136–300 CV) e coppie massime fino a 1.200 Nm. Grazie a compattezza e robustezza, i propulsori NEF sono apprezzati per l'affidabilità e per l'adattabilità a diverse configurazioni applicative, sia diesel sia a gas naturale.

La famiglia CURSOR, infine, è stata sviluppata per applicazioni ad alte prestazioni e per veicoli pesanti, destinati sia al trasporto a lungo raggio che agli autobus urbani e interurbani. Con cilindrata comprese tra 8,7 e 12,9 litri, offre potenze da 220 a 442 kW (300–600 CV) e coppie fino a 2.850 Nm. La gamma CURSOR comprende anche il motore a gas naturale più potente e silenzioso attualmente disponibile sul mercato, confermandosi come una delle soluzioni di riferimento per chi cerca prestazioni elevate unite a sostenibilità ambientale.

Un ulteriore elemento che caratterizza FPT Industrial è lo sviluppo di soluzioni legate alla mobilità elettrica attraverso la divisione ePowertrain. Con la realizzazione di trasmissioni elettriche, pacchi batteria e sistemi avanzati di gestione dell'energia, l'azienda contribuisce in maniera concreta alla transizione verso una mobilità a zero emissioni nette di carbonio. Questo impegno non è episodico, ma rappresenta una direttrice strategica che si affianca al lavoro svolto sui carburanti alternativi, confermando la volontà di posizionarsi come punto di riferimento nella propulsione sostenibile.

Nel complesso, FPT Industrial si presenta come un attore chiave nel panorama della propulsione industriale, capace di coniugare dimensione globale, innovazione tecnologica e attenzione alle sfide ambientali. Tale approccio la rende non solo un leader di mercato, ma anche un laboratorio di sperimentazione e sviluppo, il cui contributo risulta particolarmente rilevante

in un contesto in cui le dinamiche energetiche e ambientali assumono un ruolo centrale nelle politiche industriali e nei processi di innovazione.

1.2 Obiettivi

Il lavoro di tesi nasce dall'esigenza di comprendere come le informazioni generate durante il processo di gestione dei rientri in garanzia possano essere valorizzate attraverso una maggiore integrazione delle sorgenti dati aziendali. In particolare, l'attenzione è rivolta alle piattaforme MEERA e RemoteMaster, che intervengono in momenti differenti ma strettamente collegati dello stesso processo operativo.

La tesi valuta le possibilità di integrazione tra le informazioni gestite dalle due piattaforme, individuando gli elementi comuni presenti nei rispettivi modelli dati e definendo una struttura in grado di collegare le informazioni amministrative, logistiche e tecniche associate al ciclo di vita del componente.

Un secondo obiettivo riguarda la progettazione di un modello di integrazione che consenta di superare la frammentazione informativa attualmente presente tra i due sistemi, favorendo la costruzione di una base dati unificata a supporto delle attività di analisi e monitoraggio.

Infine, il lavoro si propone di valutare le potenzialità analitiche derivanti dall'integrazione delle informazioni. Attraverso una sperimentazione realizzata su dati rappresentativi delle due piattaforme, viene mostrato come la disponibilità di una base dati condivisa possa consentire la costruzione di indicatori e viste informative non ottenibili attraverso l'analisi separata delle singole sorgenti.

1.3 Struttura della tesi

L'elaborato è stato organizzato secondo un percorso che parte dall'analisi del contesto aziendale e delle sorgenti informative disponibili, per arrivare alla definizione e alla sperimentazione di un modello di integrazione dei dati a supporto del processo di gestione dei rientri in garanzia.

Il primo capitolo introduce il contesto aziendale in cui si inserisce il lavoro di tesi e ne definisce gli obiettivi, fornendo una panoramica delle tematiche affrontate nel corso dell'elaborato.

Il secondo capitolo è dedicato all'analisi delle sorgenti dati considerate nello studio. Dopo una descrizione del processo di gestione dei rientri in garanzia e del relativo flusso diagnostico, vengono presentate le piattaforme MEERA e RemoteMaster e ne vengono analizzate le strutture dati, con l'obiettivo di comprenderne il funzionamento e individuare i principali elementi informativi coinvolti nel processo.

Il terzo capitolo affronta il tema dell'integrazione tra le due piattaforme. In particolare, vengono individuati gli attributi comuni presenti nei rispettivi modelli informativi e viene proposta un'architettura di integrazione finalizzata alla costruzione di una base dati unificata.

Il quarto capitolo presenta la sperimentazione del modello proposto. In una prima fase viene descritta la costruzione del dataset sintetico utilizzato per le attività di analisi, seguita da un'analisi esplorativa delle informazioni disponibili e da una fase di standardizzazione delle problematiche registrate. Successivamente viene illustrata la simulazione del processo di integrazione attraverso i livelli Bronze, Silver e Gold della Medallion Architecture e vengono discussi i risultati ottenuti e le principali limitazioni della sperimentazione.

Infine, il quinto capitolo riporta le conclusioni del lavoro svolto e i possibili sviluppi futuri, evidenziando le opportunità di estensione del modello proposto e le potenziali evoluzioni delle attività di integrazione e analisi dei dati.

Capitolo 2

2 Analisi delle sorgenti dati

Il capitolo è dedicato all'analisi delle sorgenti dati considerate nel corso del lavoro. In una prima fase viene descritto il processo di gestione dei rientri in garanzia e il relativo flusso diagnostico, al fine di contestualizzare le attività svolte dai diversi attori coinvolti e le informazioni generate durante le varie fasi del processo.

Successivamente vengono presentate le due piattaforme aziendali che supportano tale processo: MEERA, utilizzata per la gestione amministrativa e logistica delle richieste di rientro, e RemoteMaster, dedicata alla registrazione e alla documentazione delle attività di analisi tecnica eseguite sui componenti rientrati.

Le due piattaforme svolgono funzioni differenti ma complementari all'interno del processo aziendale e rappresentano le principali sorgenti informative considerate nel presente studio. È opportuno precisare che entrambe si collocano in un contesto in continua evoluzione, poiché alcune funzionalità risultano ancora in fase di sviluppo, revisione o progressiva implementazione operativa.

Infine, vengono illustrate le rispettive strutture dati, con particolare attenzione alle principali entità, alle relazioni esistenti tra i dati e ai possibili punti di contatto tra i due sistemi, che costituiranno la base per le successive attività di integrazione.

2.1 La gestione dei rientri in garanzia

Il processo di gestione dei rientri in garanzia di FPT Industrial si articola in una sequenza di fasi che coinvolgono diversi attori e strumenti digitali, dalla segnalazione iniziale del problema da parte del cliente fino alla definizione della destinazione finale del componente ispezionato.

Il processo ha inizio con la creazione di una richiesta di rientro da parte del reparto Qualità, a cui vengono associati i componenti da ispezionare e le informazioni relative alla spedizione. All'arrivo in magazzino, il componente viene identificato e collocato in una posizione specifica in attesa dell'analisi tecnica al banco, durante la quale ogni passaggio viene documentato fotograficamente e ogni controllo riceve un esito. Al termine dell'analisi viene generato un report PDF condiviso con il dealer, che documenta l'intero processo ispettivo. Successivamente viene compilata una Survey tramite il sistema Qualtrics, un questionario strutturato in cui vengono registrate le informazioni generali sull'analisi, come il cliente, la tipologia di componente e la data.

Qualora l'analisi richiedesse competenze specialistiche non disponibili internamente, è prevista una fase di analisi esterna in cui il componente viene inviato a un fornitore specializzato. Il processo si conclude con la definizione della destinazione finale del componente — remanufacturing, smaltimento, magazzino ricambi usati, Training Center o restituzione al Buffer Warehouse — e con la registrazione dell'uscita dal magazzino.

2.2 Il processo diagnostico

La diagnostica è una procedura di analisi delle informazioni derivanti dalla misura di parametri e dalla raccolta di dati relativi a una macchina o un sistema, finalizzata a trarre informazioni sui guasti effettivi o incipienti della macchina stessa. In altri termini, la diagnostica è il complesso delle attività di analisi e sintesi che, utilizzando i rilievi di determinate grandezze fisiche caratteristiche della macchina sorvegliata, consentono di valutare le condizioni della macchina e la loro tendenza nel tempo, per previsioni sulla sua affidabilità a breve e lungo termine.

Gli obiettivi della diagnostica sono tre: l'individuazione del guasto ("detection"), la localizzazione del guasto ("isolation") e l'identificazione del guasto ("identification"). Un programma di diagnostica si sviluppa secondo i seguenti passi: acquisizione dati ("data acquisition"), elaborazione dati ("data processing") e presa di decisione ("decision-making"). I metodi per l'acquisizione e l'elaborazione dei dati e per la scelta della soglia che differenzia

la situazione di guasto dalla situazione di sano per un componente possono influenzare pesantemente le probabilità di commettere un errore in fase di presa di decisione.

Nel contesto dei rientri di FPT Industrial, questi tre passi si traducono concretamente nel processo che va dalla ricezione del componente in magazzino fino alla redazione del report finale. La fase di acquisizione dati riguarda la raccolta delle informazioni iniziali sul componente rientrato, come la descrizione del problema dichiarato dal cliente, gli identificativi del componente e i dati relativi alla spedizione. Le fasi di elaborazione dei dati e di presa di decisione corrispondono invece all'ispezione tecnica al banco, durante la quale il componente viene analizzato, i controlli vengono documentati e l'esito dell'ispezione viene formalizzato. Le informazioni prodotte durante queste attività vengono registrate all'interno delle piattaforme aziendali secondo strutture dati differenti, la cui analisi costituisce il punto di partenza del presente lavoro.

Nei paragrafi successivi vengono quindi descritte le funzionalità di MEERA e RemoteMaster e analizzati i rispettivi modelli informativi, con l'obiettivo di evidenziarne le principali caratteristiche e i possibili punti di contatto.

2.3 MEERA

MEERA è la piattaforma digitale dedicata alla gestione amministrativa e logistica dei rientri in garanzia. La piattaforma copre il flusso operativo che precede l'analisi tecnica dei componenti, dalla creazione della richiesta di rientro fino alla presa in carico del componente da parte del reparto di Analisi Componenti.

Il processo si avvia con la creazione di una Teardown Return Request, denominata TRR, che rappresenta l'entità centrale attorno alla quale si organizza il flusso. La TRR può essere aperta dall'operatore FPT Quality/Warranty Recovery oppure direttamente dall'OEM o dal dealer, e contiene le informazioni principali relative alla richiesta, tra cui il problema segnalato, la priorità e l'urgenza. In base alla tipologia di rientro, la richiesta può essere associata a un numero THD, nel caso dei motori completi, oppure a un numero di Claim, nel caso dei singoli componenti oggetto di pratica di garanzia.

A ciascuna TRR vengono associati uno o più item, ovvero i componenti fisici oggetto del rientro. Gli item possono essere rappresentati da motori completi oppure da componenti singoli, come ad esempio un turbocompressore. In entrambi i casi, l'identificazione dell'item avviene attraverso informazioni quali Part Number e Serial Number, che permettono di distinguere il tipo di componente e lo specifico esemplare rientrato. Nel caso dei motori, è inoltre possibile creare componenti figli associati al motore padre, riflettendo la gerarchia fisica del prodotto.

Una volta creata la TRR e associati gli item, il componente viene spedito verso il magazzino FPT. La spedizione può transitare attraverso un Buffer Warehouse intermedio, gestito da un operatore esterno, oppure essere inviata direttamente al magazzino FPT nel caso di spedizioni urgenti. In entrambi i casi, la piattaforma traccia lo stato della spedizione attraverso tre fasi: da spedire, in consegna e consegnato.

All'arrivo del componente in magazzino, la data di arrivo viene registrata automaticamente e la scatola contenente il componente viene identificata tramite un QR Code. Tale codice consente la localizzazione fisica del componente all'interno del magazzino, definita attraverso scaffale, livello e posizione. In questo modo, MEERA permette di seguire il rientro dal punto di vista amministrativo, logistico e fisico fino al momento in cui il componente è disponibile per l'analisi tecnica.

2.4 Struttura dati MEERA

Per analizzare la struttura dati di MEERA sono stati utilizzati lo schema ER della piattaforma e una prima estrazione di test. Lo schema ER offre una rappresentazione complessiva delle entità presenti nel sistema e dei collegamenti tra le diverse aree informative, mentre l'estrazione di test permette di osservare più concretamente la struttura delle tabelle, i campi disponibili e il modo in cui le informazioni vengono registrate.

Il modello dati di MEERA comprende diverse entità legate al processo di gestione dei rientri. Oltre alla richiesta di rientro, sono presenti informazioni relative agli item associati, alla spedizione, alla localizzazione fisica in magazzino, alle attività svolte sul componente e alla classificazione del prodotto. Di seguito viene quindi presentato lo schema ER della piattaforma, seguito da una panoramica delle principali tabelle individuate, con l'obiettivo di chiarire il ruolo di ciascuna entità all'interno del modello informativo.

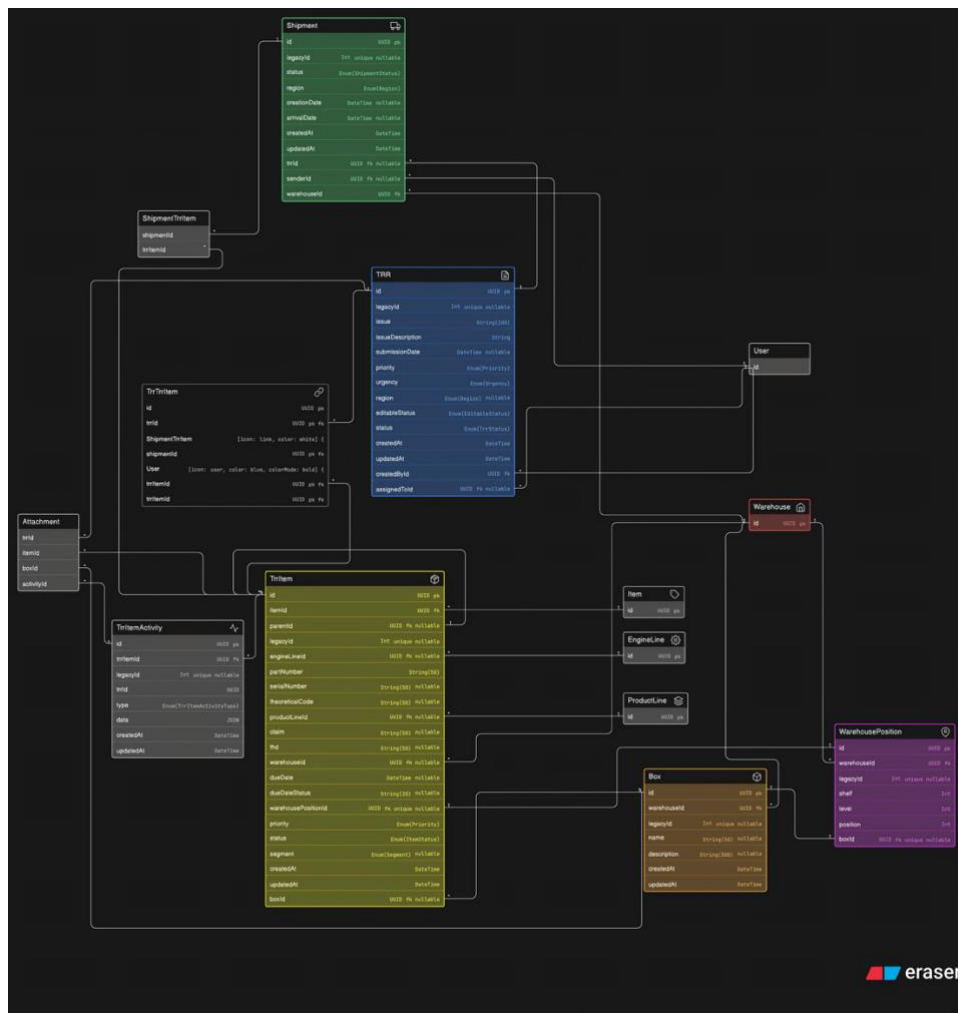


Figure 1 - Modello ER della piattaforma MEERA

Tabella	Descrizione
Trr	Contiene le informazioni principali della richiesta di rientro, come stato, priorità, urgenza e problema segnalato.
TrrTrrItem	Collega le richieste di rientro agli item associati, permettendo di rappresentare il legame tra una TRR e uno o più componenti o motori.
TrrItem	Rappresenta il componente fisico o il motore completo oggetto del rientro, includendo informazioni operative e identificative utili alla gestione del processo.
Item	Contiene le informazioni anagrafiche e descrittive degli item, come nome, classe, gruppo, sottogruppo e tipologia.
Shipment	Raccoglie le informazioni relative alla spedizione del componente verso il magazzino FPT o verso eventuali punti intermedi del flusso logistico.
ShipmentTr rItem	Collega gli item associati alle richieste di rientro con le rispettive spedizioni, consentendo di seguire il componente lungo il flusso logistico.
Box	Identifica la scatola o il contenitore associato al componente, utile per la gestione fisica e la tracciabilità in magazzino.
Warehouse	Contiene le informazioni relative ai magazzini coinvolti nel processo, distinguendo ad esempio magazzini centrali e buffer warehouse.
Warehouse Position	Descrive la collocazione fisica del componente all'interno del magazzino, attraverso informazioni come scaffale, livello e posizione.
TrrItemAct ivity	Registra le attività o gli eventi collegati all'item durante il processo, contribuendo alla ricostruzione dello stato di avanzamento.
EngineLine	Contiene informazioni di classificazione relative alla famiglia o linea motore associata al prodotto.
ProductLin e	Raccoglie informazioni relative alla linea o segmento di prodotto, utili per classificare gli item all'interno della gamma aziendale.

Table 1 - Descrizione delle tabelle del database di MEERA

2.4.1 Dettaglio tabelle principali

La panoramica precedente mostra come il modello dati di MEERA comprenda entità relative a diverse fasi del processo di gestione dei rientri. Tuttavia, ai fini del presente lavoro, l'attenzione viene concentrata sulle tabelle che permettono di ricostruire il legame tra la richiesta di rientro e il componente oggetto di analisi. In particolare, risultano centrali le tabelle *Trr*, *TrrItem* e *Item*.

Per ciascuna tabella viene riportata una riga esemplificativa, costruita con valori realistici e coerenti con la struttura effettiva dei campi presenti nell'estrazione di test, accompagnata da una descrizione degli stessi. Tali valori hanno finalità esclusivamente illustrativa e servono a mostrare la tipologia di informazioni contenute nelle tabelle.

Trr

Campo	Valore esemplificativo	Descrizione
issue	low oil pressure	Sintesi del problema segnalato nella richiesta di rientro.
issueDescription	low oil pressure during engine operation	Descrizione più estesa della problematica riportata dal cliente o dal dealer.
assignedToId	45	Identificativo dell'utente o operatore a cui è assegnata la richiesta.
createdById	12	Identificativo dell'utente che ha creato la richiesta.

Campo	Valore esemplificativo	Descrizione
submissionDate	2025-10-30 09:15:00	Data e ora di invio o formalizzazione della richiesta.
id	1	Identificativo univoco della TRR.
editableStatus	SUBMITTED	Indica lo stato di modificabilità della richiesta.
status	OPEN	Stato di avanzamento della richiesta nel processo.
priority	HIGH	Priorità assegnata alla richiesta.
region	EMEA	Area geografica di riferimento della richiesta.
urgency	URGENT	Livello di urgenza associato alla richiesta.
createdAt	2025-10-30 09:10:24.315	Data e ora di creazione del record nel sistema.
updatedAt	2025-10-31 14:25:10.023	Data e ora dell'ultimo aggiornamento del record.

Table 2 - Descrizione dei campi della tabella "Trr"

TrrItem

Campo	Valore esemplificativo	Descrizione
id	91	Identificativo univoco dell'item associato alla richiesta di rientro.
boxId	29	Identificativo della scatola o del contenitore associato all'item.
itemId	34	Riferimento all'anagrafica dell'item presente nella tabella Item.

Campo	Valore esemplificativo	Descrizione
dueDate	2026-02-20	Data entro cui l'attività o la gestione dell'item dovrebbe essere completata, se prevista.
dueDateStatus	ON_TIME	Stato relativo al rispetto della scadenza associata all'item, se presente.
warehousePositionId	15	Riferimento alla posizione fisica dell'item all'interno del magazzino.
parentId	NULL	Riferimento a un eventuale item padre, utile nel caso di componenti collegati a un motore completo.
serialNumber	SN1979541	Numero seriale che identifica lo specifico esemplare rientrato.
partNumber	5801234567	Codice identificativo del tipo di componente o motore.
theoreticalCode	F4AFE601A*S002	Codice teorico identificativo del motore.
claim	CLM202600145	Numero di Claim associato alla pratica di garanzia, utilizzato soprattutto per i singoli componenti.
thd	THD202600089	Numero THD associato alla richiesta di supporto, utilizzato soprattutto per i motori completi.
status	FIRST_INSPECTION	Stato di avanzamento dell'item nel processo.
engineLineId	9	Riferimento alla linea motore associata all'item.
productLineId	9	Riferimento alla linea prodotto associata all'item.
segment	ON_ROAD	Segmento applicativo o commerciale del prodotto.
warehouseId	1	Identificativo del magazzino in cui si trova l'item.
priority	LOW	Priorità associata alla gestione dell'item.
createdAt	2026-02-09 06:41:27.261	Data e ora di creazione del record nel sistema.

Campo	Valore esemplificativo	Descrizione
updatedAt	2026-02-10 07:19:51.025	Data e ora dell'ultimo aggiornamento del record.

Table 3 - Descrizione dei campi della tabella "TrrItem"

Item

Campo	Valore esemplificativo	Descrizione
itemName	TURBOCHARGER	Nome dell'item, cioè del componente o motore presente nell'anagrafica.
id	101	Identificativo univoco dell'item.
class	54	Codice numerico della classe di appartenenza dell'item.
group	5400	Codice numerico del gruppo associato all'item.
subgroup	540012	Codice numerico del sottogruppo associato all'item.
type	Component	Tipologia dell'item, utile a distinguere tra componente singolo e motore completo.

Table 4 - Descrizione dei campi della tabella "Item"

2.5 RemoteMaster

RemoteMaster è la piattaforma digitale utilizzata dal reparto di Analisi Componenti per gestire e documentare il processo di ispezione dei componenti rientrati in garanzia. La piattaforma persegue due obiettivi fondamentali. Il primo è di natura gestionale: verificare se il componente presenta

effettivamente un difetto, fornendo evidenza documentata a supporto della decisione di riconoscere o meno la garanzia richiesta dal cliente. Il secondo è di natura analitica: alimentare una base di conoscenza strutturata sui difetti riscontrati, che permetta di identificare nel tempo problematiche ricorrenti su specifici componenti o famiglie di motori, supportando così il miglioramento continuo della qualità dei prodotti.

La piattaforma è articolata in tre versioni distinte:

- **RemoteMaster Analisi Componenti (Nexus)**, dedicata all'ispezione fisica dei componenti al banco;
- **RemoteMaster Training**, utilizzata per la creazione di videolezioni destinate ai dealer, relative alle operazioni da effettuare sui motori;
- **RemoteMaster App**, versione mobile utilizzata principalmente dai meccanici sul campo per fotografare e segnalare direttamente un problema, generando un report semplificato.

Il presente lavoro si concentra su RemoteMaster Analisi Componenti (Nexus), in quanto è la versione direttamente coinvolta nel processo di gestione dei rientri in garanzia.

Il processo di analisi si svolge su un banco appositamente attrezzato, dotato di camere ad alta risoluzione posizionate in diverse angolazioni, che permettono

di documentare fotograficamente il componente da più prospettive. Ogni componente viene identificato tramite il numero di Claim associato alla pratica di garanzia, che costituisce il riferimento principale per tracciare l'intero processo ispettivo.

L'analisi segue una checklist preimpostata in base alla tipologia di componente da esaminare. La checklist è composta da una sequenza di step, ciascuno dei quali indica all'operatore le operazioni da eseguire, ad esempio ruotare il componente di novanta gradi e scattare fotografie da angolazioni specifiche, garantendo che ogni ispezione segua un protocollo standardizzato e ripetibile. Al termine di ogni step, l'operatore registra un esito: OK se il controllo non ha evidenziato anomalie, KO in caso contrario. Al completamento dell'intera checklist, la piattaforma genera automaticamente un report PDF che raccoglie la documentazione fotografica di ogni step, gli esiti dei controlli e le note conclusive dell'analisi. Il report viene condiviso internamente all'azienda e successivamente con il dealer, fornendo una documentazione trasparente e strutturata dell'ispezione effettuata e dell'esito della valutazione in garanzia.

2.6 Struttura dati di RemoteMaster

A differenza di MEERA, per RemoteMaster non è attualmente disponibile un'estrazione di dati associata al database, poiché la piattaforma si trova ancora in fase di revisione e costruzione. L'analisi della struttura dati si basa quindi esclusivamente sul modello ER preliminare, utilizzato per comprendere le

principali entità previste e il modo in cui la piattaforma dovrebbe organizzare le informazioni relative all'analisi tecnica dei componenti.

Il modello dati di RemoteMaster è orientato alla documentazione dell'ispezione al banco. Le entità previste riguardano principalmente l'anagrafica del componente, la checklist associata alla tipologia di componente, i moduli e gli step dell'analisi e i relativi esiti. Rispetto a MEERA, che gestisce la fase amministrativa e logistica del rientro, RemoteMaster si concentra quindi sulla fase tecnica del processo, registrando le informazioni generate durante l'analisi.

Di seguito viene presentato il modello ER preliminare di RemoteMaster, utilizzato come riferimento per comprendere la struttura informativa prevista dalla piattaforma. Successivamente viene proposta una rielaborazione dello schema, costruita sulla base delle considerazioni emerse durante l'analisi, con l'obiettivo di rappresentare in modo più esplicito le relazioni tra le principali entità coinvolte nel processo di analisi dei componenti.

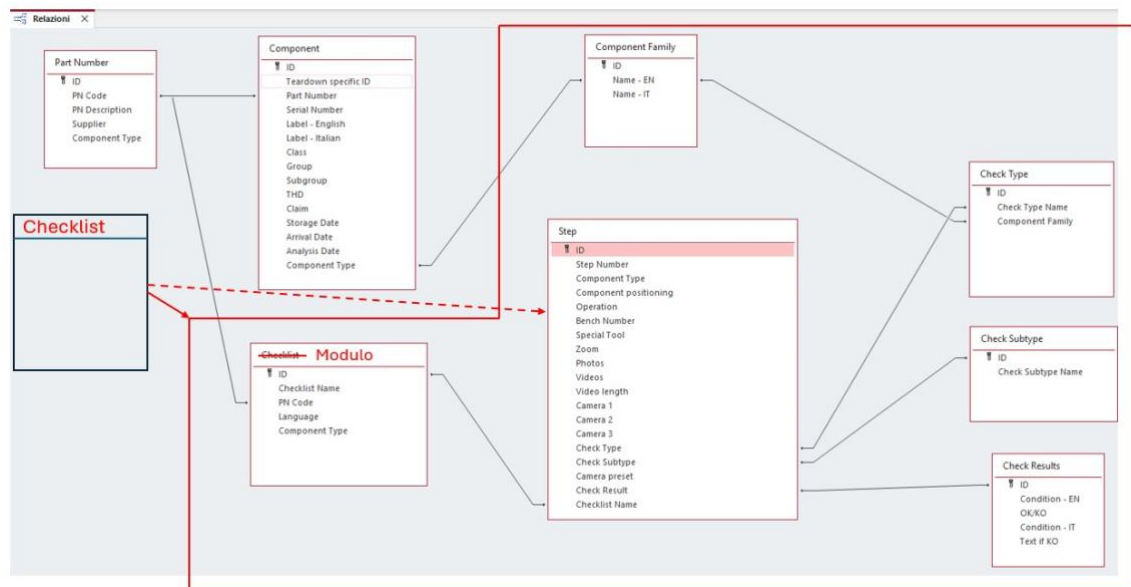


Figure 2 - Modello ER preliminare della piattaforma RemoteMaster

Sulla base del modello preliminare è stata proposta una rielaborazione finalizzata a rendere più chiara la struttura informativa della piattaforma:

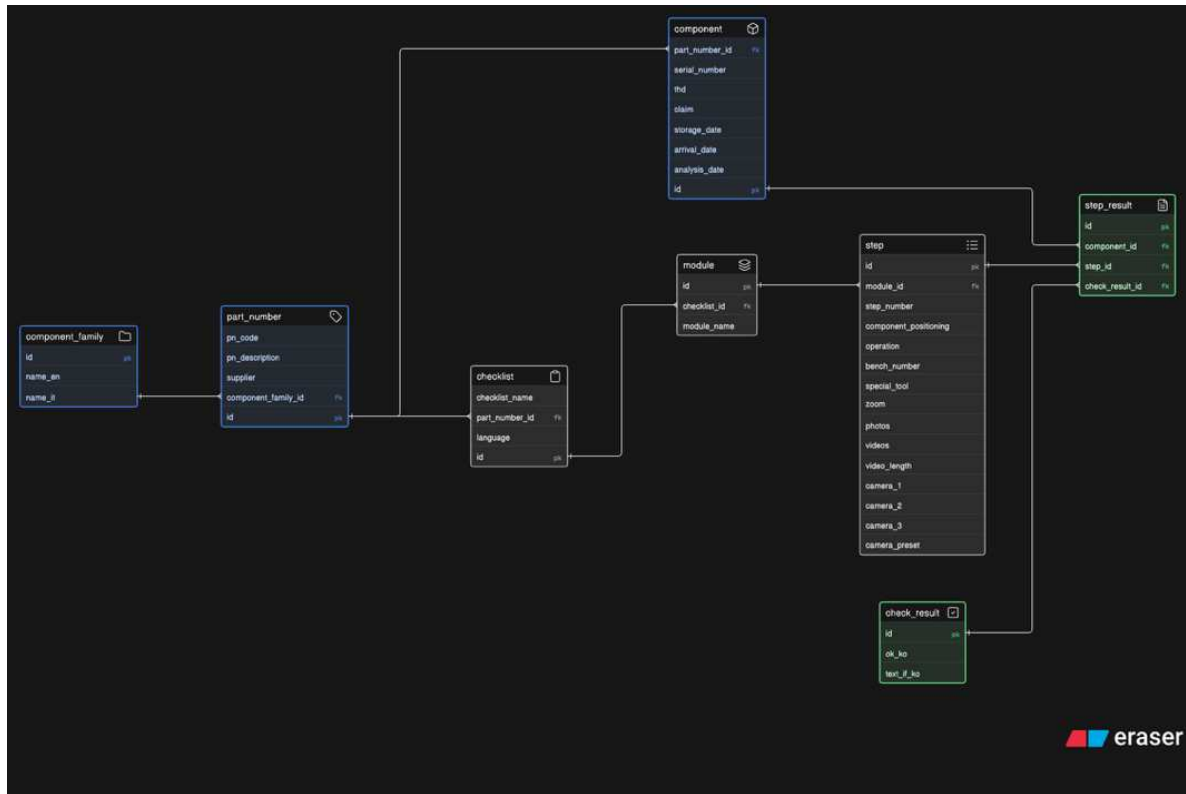


Figure 3 - Modello ER rivisitato della piattaforma RemoteMaster

In primo luogo, è stata introdotta una distinzione esplicita tra checklist, moduli e step di analisi, al fine di rappresentare in modo più aderente la struttura operativa della piattaforma. In questa configurazione, la checklist identifica il protocollo di analisi associato a uno specifico componente, il modulo rappresenta una sezione della checklist e lo step descrive la singola operazione eseguita dall'operatore durante l'ispezione.

Sono state inoltre rimosse alcune entità presenti nello schema preliminare, quali Check Subtype e Check Type, poiché le informazioni da esse rappresentate risultavano ridondanti rispetto alla struttura basata su checklist, moduli e step. Analogamente, alcuni attributi relativi alla classificazione del componente sono

stati esclusi dalla rappresentazione, privilegiando una struttura più compatta e focalizzata sugli elementi direttamente coinvolti nel processo di analisi.

Un'ulteriore modifica ha riguardato la gestione degli esiti delle verifiche. Nel modello preliminare il risultato del controllo risultava associato direttamente allo step. Nello schema rielaborato è stata invece introdotta l'entità Step Result, che collega il componente analizzato, lo step eseguito e il relativo esito. Questa soluzione favorisce una distinzione più chiara tra la definizione della procedura di analisi e i risultati prodotti durante la sua esecuzione, permettendo la registrazione dello stesso protocollo di controllo per componenti differenti e la costruzione di uno storico delle analisi effettuate.

Le modifiche proposte non hanno l'obiettivo di definire una struttura definitiva della piattaforma, ma di ottenere un modello dati più coerente con il flusso operativo descritto e maggiormente adatto alle successive attività di integrazione.

2.6.1 Dettaglio tabelle principali

A differenza di quanto fatto per MEERA, per RemoteMaster non vengono riportate righe esemplificative, poiché non è disponibile un'estrazione di dati associata al database. Il dettaglio delle tabelle si concentra quindi sulla funzione delle principali entità previste dal modello ER e sui campi più rilevanti dal punto di vista informativo.

Le tabelle Part Number, Component e Component Family non vengono approfondite singolarmente, poiché contengono principalmente informazioni anagrafiche e identificative del componente, in larga parte già descritte nell'analisi della struttura dati di MEERA. L'attenzione viene invece concentrata sulle tabelle che descrivono la struttura della checklist di analisi, ossia Checklist, Module, Step, Step Result e Check Result.

Checklist

Campo	Descrizione
ID	Identificativo univoco della checklist.
Checklist Name	Nome della checklist.
Part Number ID	Riferimento al Part Number associato alla checklist.
Language	Lingua della checklist.

Table 5 - Descrizione dei campi della tabella "Checklist"

Module

Campo	Descrizione
ID	Identificativo univoco del modulo.
Checklist ID	Riferimento alla checklist di appartenenza.
Module Name	Nome del modulo, corrispondente a una sezione della checklist.

Table 6 - Descrizione dei campi della tabella "Module"

Step

Campo	Descrizione
ID	Identificativo univoco dello step.
Module ID	Riferimento al modulo di appartenenza.
Step Number	Numero progressivo dello step all'interno del modulo.
Component positioning	Posizionamento richiesto del componente durante l'operazione.
Operation	Operazione che l'operatore deve eseguire.
Bench Number	Banco utilizzato per l'analisi.
Special Tool	Eventuale strumento speciale richiesto per lo step.
Zoom	Indicazione relativa allo zoom da utilizzare.
Photos	Indicazione relativa alle fotografie richieste.
Videos	Indicazione relativa ai video richiesti.
Video length	Durata prevista del video, se richiesto.
Camera 1	Prima camera da utilizzare.
Camera 2	Seconda camera da utilizzare.
Camera 3	Terza camera da utilizzare.
Camera preset	Preset della camera associato allo step.

Table 7 - Descrizione dei campi della tabella "Step"

Check Result

Campo	Descrizione
ID	Identificativo univoco dell'esito del controllo.
OK / KO	Indicazione sintetica dell'esito del controllo.
Text if KO	Testo o nota associata al caso in cui l'esito sia negativo.

Table 8 - Descrizione dei campi della tabella "Check Result"

Step Result

Campo	Descrizione
ID	Identificativo univoco del record.
Component ID	Riferimento al componente analizzato. Collega il risultato dell'analisi al componente specifico sul quale è stato eseguito il controllo.
Step ID	Riferimento allo step della checklist eseguito durante l'ispezione.
Check Result ID	Riferimento all'esito registrato per lo step eseguito. Collega il record alla tabella <i>Check Result</i> , nella quale sono definiti i possibili risultati dell'analisi.

Table 9 - Descrizione dei campi della tabella "Step Result"

Capitolo 3

3 Integrazione

La descrizione delle due piattaforme evidenzia come MEERA e RemoteMaster intervengano su fasi consecutive dello stesso processo. MEERA raccoglie e organizza le informazioni necessarie alla gestione del rientro, mentre RemoteMaster documenta l'analisi tecnica del componente una volta arrivato al banco. Nonostante questa continuità operativa, allo stato attuale non vi è una connessione logica tra le piattaforme e le relative basi dati.

L'assenza di una struttura informativa condivisa limita la possibilità di ottenere una visione unitaria del processo e di valorizzare congiuntamente le informazioni amministrative, logistiche e tecniche generate durante la gestione del rientro. Una maggiore integrazione dei dati consentirebbe infatti di migliorare la tracciabilità del componente, ridurre la frammentazione informativa e creare le condizioni per lo sviluppo di analisi basate sull'intero ciclo di gestione della garanzia.

A partire da queste considerazioni, il presente capitolo individua gli elementi comuni presenti nei modelli informativi delle due piattaforme e presenta una

possibile architettura di integrazione finalizzata alla costruzione di una base dati unificata.

3.1 Elementi comuni tra le piattaforme

L'analisi delle strutture dati delle piattaforme ha permesso di identificare alcuni campi presenti in entrambe le piattaforme, riconducibili principalmente all'identificazione del componente oggetto di rientro e di analisi. Tali attributi assumono un ruolo centrale poiché permettono di seguire lo stesso componente lungo le diverse fasi del processo, dalla registrazione della richiesta di rientro fino alla conclusione delle attività di analisi tecnica.

L'identificazione del componente avviene innanzitutto attraverso due codici principali: il Part Number e il Serial Number. Il Part Number identifica la tipologia del componente, ovvero il codice prodotto a cui esso appartiene. Componenti differenti possono quindi condividere lo stesso Part Number se appartenenti alla medesima categoria. Il Serial Number, invece, identifica il singolo componente fisico e consente di distinguere unità diverse appartenenti allo stesso codice prodotto. La combinazione tra Part Number e Serial Number permette pertanto di riconoscere in modo univoco il componente analizzato.

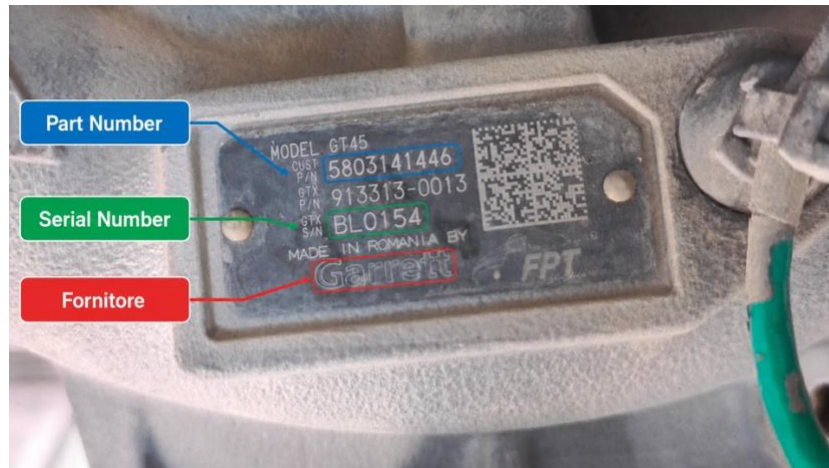


Figure 4 - Esempio di targhetta identificativa di un turbocompressore

Accanto a questi codici identificativi assumono rilevanza anche il numero di Claim e il numero THD. Il Claim rappresenta il riferimento alla pratica di garanzia associata al rientro del componente e costituisce il principale collegamento tra le informazioni amministrative e le successive attività di analisi. Il THD, acronimo di Ticket Help Desk, identifica invece la richiesta di supporto tecnico aperta dal cliente ed è associato principalmente ai casi relativi ai motori. Pur non descrivendo direttamente il componente dal punto di vista tecnico, tali riferimenti consentono di collegarlo al contesto operativo da cui ha origine il processo di analisi.

In questo senso, il riconoscimento del componente non dipende da un singolo attributo, ma dalla combinazione di informazioni tecniche e operative. Part Number e Serial Number garantiscono l'identificazione del componente, mentre Claim e THD permettono di collegarlo alla pratica di garanzia o alla richiesta di supporto tecnico associata. Tali elementi costituiscono il principale punto di contatto tra MEERA e RemoteMaster e rappresentano la base informativa sulla quale può essere costruito il modello di integrazione.

3.2 Architettura proposta per l'integrazione

Per la definizione della proposta di integrazione è stata adottata un'architettura Data Lakehouse, utilizzando Databricks come riferimento tecnologico. Databricks è una piattaforma cloud progettata per la gestione, l'elaborazione e l'analisi dei dati all'interno di un ambiente unificato. La piattaforma consente di lavorare su dati provenienti da sorgenti eterogenee e rappresenta uno dei principali esempi di implementazione dell'architettura Data Lakehouse.

Il Data Lakehouse nasce con l'obiettivo di combinare alcune caratteristiche tipiche dei Data Lake e dei Data Warehouse. Essi rendono possibile la raccolta di grandi quantità di dati, anche eterogenei e non necessariamente strutturati, mantenendoli in una forma vicina a quella originaria. I Data Warehouse, invece, sono tradizionalmente orientati alla gestione di dati strutturati, organizzati e ottimizzati per attività di reporting e analisi. Il Data Lakehouse cerca di unificare questi due approcci, offrendo un ambiente in grado di conservare dati provenienti da sorgenti differenti e, al tempo stesso, supportare interrogazioni, trasformazioni e analisi su dati organizzati e governati.

In questo senso, un'architettura Data Lakehouse può essere utilizzata come livello centrale di raccolta e trasformazione dei dati aziendali. Le informazioni provenienti da diversi sistemi possono essere acquisite all'interno di una struttura comune, mantenute inizialmente nella loro forma originale e

successivamente elaborate per ottenere dataset più coerenti e adatti alle analisi. Questo approccio risulta particolarmente adatto a contesti nei quali le informazioni sono distribuite su più piattaforme e devono essere integrate per costruire una visione più completa dei processi aziendali.

Coerentemente con il paradigma Data Lakehouse, il processo di integrazione può essere descritto secondo una logica di tipo ELT. L'ELT, acronimo di Extract, Load, Transform, è un approccio moderno all'integrazione dei dati progettato per piattaforme analitiche cloud-native. In una pipeline ELT, i dati vengono estratti dai sistemi sorgente, caricati direttamente in un repository dati centrale e infine trasformati all'interno di quel sistema di destinazione. Questa sequenza è la caratteristica distintiva dell'ELT e uno dei motivi principali per cui questo modello è diventato fondamentale nelle moderne architetture dati.

Tale logica viene implementata all'interno di Databricks attraverso la cosiddetta Medallion Architecture, un modello di gestione dei dati articolato in tre livelli progressivi, comunemente identificati come Bronze, Silver e Gold. Ciascun livello rappresenta uno stadio differente del processo di elaborazione, consentendo di mantenere una chiara separazione tra dati originali, dati trasformati e dati destinati al consumo finale.

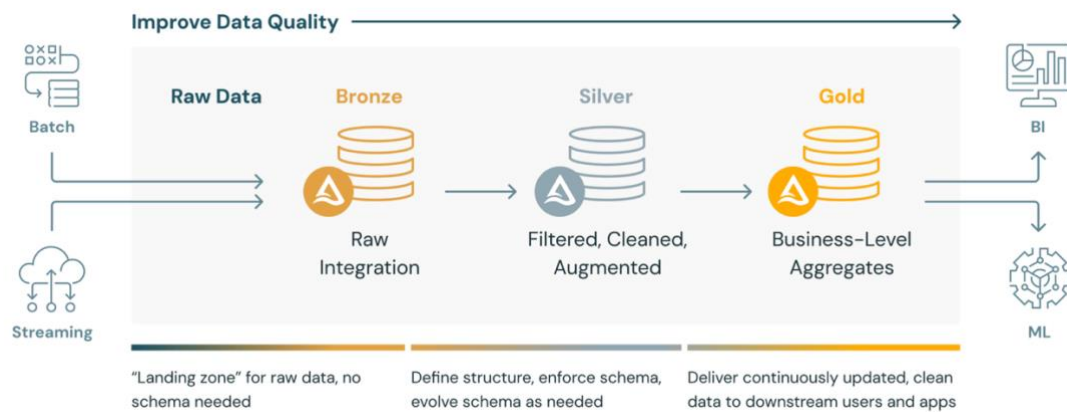


Figure 5 - Architettura "Medallion" implementata da Databricks

Il Bronze Layer costituisce il punto di ingresso delle informazioni nel sistema. In questa fase i dati provenienti dalle diverse sorgenti vengono acquisiti e memorizzati senza modificarne la struttura originaria, preservando integralmente il contenuto dei database operativi.

Il Silver Layer rappresenta il livello dedicato alla preparazione e alla normalizzazione dei dati. In questa fase vengono effettuate le operazioni necessarie per rendere confrontabili le informazioni provenienti dalle diverse sorgenti, quali il controllo della coerenza dei formati, la standardizzazione degli attributi condivisi e l'applicazione delle regole di integrazione.

Il Gold Layer rappresenta il livello finale del processo di elaborazione e raccoglie i dati ormai consolidati e integrati all'interno di una struttura unificata. In questo livello le informazioni provenienti dalle diverse sorgenti confluiscono in una base dati storica comune sulla quale possono essere

sviluppate successive elaborazioni, indicatori di monitoraggio e strumenti di Business Intelligence.

L'architettura descritta costituisce il modello di riferimento per la proposta di integrazione sviluppata nel presente lavoro. Nel capitolo successivo tale modello verrà applicato nell'ambito di una sperimentazione finalizzata alla simulazione del processo di integrazione tra MEERA e RemoteMaster.

Capitolo 4

4 Sperimentazione del modello di integrazione

Come anticipato nei capitoli precedenti, MEERA e RemoteMaster si collocano in un contesto ancora in fase di sviluppo e progressivo consolidamento. Di conseguenza, il presente lavoro non può basarsi su una base dati storica completa e consolidata: per MEERA è disponibile esclusivamente un'estrazione di test, mentre per RemoteMaster l'analisi si fonda sul modello ER preliminare della piattaforma.

In assenza di una base dati reale sufficientemente strutturata, la sperimentazione ha richiesto la costruzione di un dataset sintetico capace di riprodurre le principali caratteristiche informative osservate nelle sorgenti disponibili. La costruzione del dataset prende avvio da MEERA, unica piattaforma per la quale sono disponibili dati effettivamente popolati, utilizzati come riferimento per definire relazioni, distribuzioni e attributi coerenti con il contesto operativo analizzato.

La sperimentazione si articola quindi in più fasi successive. In primo luogo, viene descritto il processo di costruzione del dataset sintetico e le scelte metodologiche adottate per il suo popolamento. Successivamente vengono

presentate alcune analisi esplorative finalizzate a evidenziare criticità e potenziali aree di miglioramento dell'attuale modello dati di MEERA. Le evidenze emerse costituiscono la base per la definizione di un modello informativo più strutturato e per l'introduzione dei dati rappresentativi di RemoteMaster.

Nella parte conclusiva del capitolo viene quindi proposta e sperimentata una soluzione di integrazione tra le due piattaforme attraverso una struttura multilivello basata sulla Medallion Architecture. La base dati integrata ottenuta viene infine utilizzata per costruire indicatori esemplificativi, con l'obiettivo di mostrare le potenzialità analitiche derivanti dalla centralizzazione e dall'integrazione delle informazioni generate durante il processo di gestione dei rientri in garanzia.

4.1 Costruzione del dataset

La costruzione del dataset è stata realizzata con il supporto di un Large Language Model (LLM), impiegato come strumento per la generazione dei dati. La scelta di adottare questo approccio è stata motivata dalla necessità di ottenere un numero sufficiente di osservazioni mantenendo al tempo stesso una distribuzione plausibile delle informazioni e delle relazioni presenti nel processo reale. L'utilizzo della LLM non ha però comportato una generazione libera dei dati: la struttura del dataset, i vincoli tra i campi e le principali regole

di compilazione sono stati definiti sulla base dell'estrazione disponibile e della conoscenza operativa del processo.

Per la costruzione del dataset sono state ricreate le quattro tabelle principali presenti nell'estrazione di test di MEERA: Item, TrrItem, Trr e TrrTrrItem. La scelta di limitare la generazione a queste sole tabelle è stata dettata dalla necessità di riprodurre gli elementi essenziali del processo di gestione dei rientri, mantenendo una struttura coerente con quella osservata nell'estrazione disponibile.

La tabella Item rappresenta l'anagrafica di riferimento degli item gestiti nel dataset. Per la sua costruzione è stata utilizzata una lista ufficiale di componenti analizzabili, comprensiva dei relativi Part Number e fornitori. A partire da tale elenco è stata definita la base identificativa degli item, composta da 37 componenti univoci e dal motore completo.

La tabella TrrItem è stata popolata con 500 record, generati in modo da riprodurre una distribuzione plausibile delle richieste di rientro gestite dalla piattaforma. In particolare, il 70% dei record è stato associato a componenti singoli, mentre il restante 30% a motori completi. Tale scelta è stata effettuata per rappresentare in modo realistico le principali tipologie di richieste osservate nel processo, mantenendo al tempo stesso una prevalenza dei componenti, che costituiscono l'oggetto principale delle attività di analisi.

I valori dei campi identificativi sono stati generati seguendo la struttura osservata nell'estrazione di test. I Part Number utilizzati corrispondono a codici reali presenti nella lista ufficiale dei componenti, mentre i campi Claim e THD sono stati costruiti rispettando la logica operativa del processo. In particolare, le richieste relative ai componenti singoli presentano un numero di Claim valorizzato e un campo THD vuoto, viceversa per i motori completi.

Anche i campi temporali sono stati generati seguendo criteri di coerenza con il processo reale. Le date di creazione, aggiornamento e scadenza sono state costruite in modo compatibile con la priorità e con lo stato della richiesta. In particolare, la data di scadenza è stata determinata a partire dalla priorità assegnata alla pratica, mentre la data di aggiornamento è stata resa coerente con il relativo stato di avanzamento. Alcuni campi presenti ma non valorizzati nell'estrazione originale sono stati mantenuti vuoti anche nel dataset sintetico, al fine di preservare la coerenza con i dati di partenza.

Infine, la distribuzione dei componenti non è stata definita in modo uniforme. Alcuni componenti sono stati fatti comparire con maggiore frequenza rispetto ad altri sulla base di confronti effettuati con il personale tecnico, così da riflettere in modo più realistico la tipologia di richieste che caratterizza il processo di rientro. La Figura X mostra la distribuzione dei componenti più frequenti presenti nel dataset sintetico.

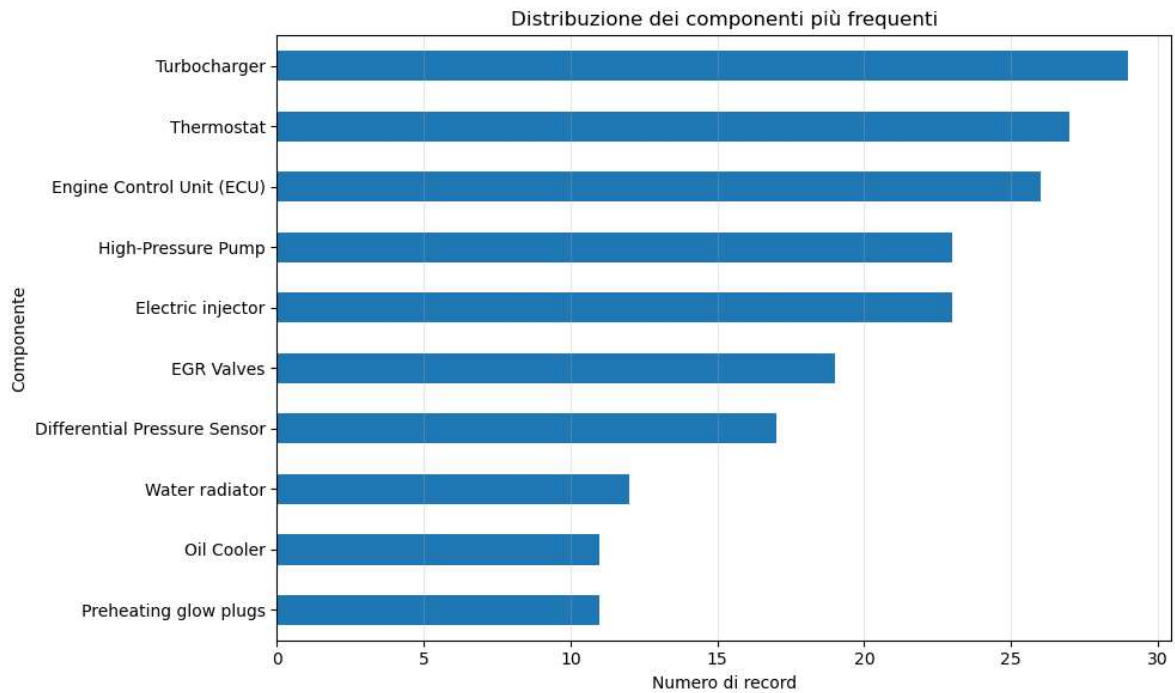


Figure 6 - Grafico a barre della distribuzione dei 10 componenti più frequenti

La tabella Trr contiene le informazioni associate alla richiesta di rientro e rappresenta il punto in cui vengono registrate le problematiche segnalate dal cliente. In particolare, i campi Issue e IssueDescription descrivono il problema riscontrato e costituiscono la principale fonte informativa per comprendere le cause che hanno portato all'apertura della richiesta.

Per la generazione di questi campi è stata predisposta una lista iniziale di problematiche compatibili con ciascun componente presente nel dataset. Per ogni tipologia di componente sono state individuate circa quattro o cinque possibili issue, selezionate in modo da rappresentare situazioni realistiche di guasto o malfunzionamento. I valori dei campi Issue e IssueDescription sono stati successivamente assegnati a partire da tali elenchi, mantenendo una coerenza logica con il componente associato alla richiesta.

Per rendere il dataset maggiormente rappresentativo di una situazione reale, una quota compresa tra il 5% e il 10% delle osservazioni è stata inoltre modificata introducendo errori di digitazione, differenze nell'utilizzo di maiuscole e minuscole, variazioni nella punteggiatura o valori mancanti. Questa scelta ha permesso di simulare alcune delle criticità tipiche dei campi a testo libero, la cui compilazione dipende direttamente dall'utente e non segue necessariamente criteri uniformi.

La presenza di informazioni descrittive non strutturate assume particolare rilevanza dal punto di vista analitico. Per questo motivo, una volta completata la costruzione del dataset, è stata effettuata una fase preliminare di preparazione ed esplorazione dei dati con l'obiettivo di valutare il grado di variabilità delle problematiche registrate e l'impatto che tale variabilità può avere sulle successive attività di analisi.

4.2 Analisi esplorativa

Come evidenziato nel paragrafo precedente, i campi Issue e IssueDescription sono stati progettati come campi a testo libero, permettendo agli utenti di descrivere la problematica riscontrata senza particolari vincoli di compilazione. Sebbene questa soluzione garantisca una maggiore flessibilità nella registrazione delle informazioni, può introdurre difficoltà nelle successive

attività di analisi, poiché problematiche simili possono essere descritte utilizzando terminologie differenti.

Per valutare l'entità di questo fenomeno è stata condotta un'analisi esplorativa delle problematiche presenti nel dataset. Poiché il dataset comprende sia motori completi sia componenti singoli appartenenti a numerose tipologie differenti, è stata effettuata una distinzione tra queste due categorie. La Tabella 10 riporta il numero totale di richieste analizzate, il numero di problematiche univoche rilevate e il relativo tasso di variabilità, calcolato come rapporto tra problematiche univoche e numero totale di osservazioni.

Tipologia di item	Record totali	Issue univoche	Tasso di variabilità
Motore completo	145	37	25,5%
Componenti	355	155	43,7%

Table 10 - Variabilità delle problematiche per singoli componenti e motore completo

I risultati mostrano come le richieste relative ai componenti presentino una variabilità significativamente superiore rispetto a quelle associate ai motori completi. Tale differenza è in parte spiegata dalla maggiore eterogeneità dei componenti presenti nel dataset: mentre tutte le richieste relative ai motori completi fanno riferimento a un'unica tipologia di item, le richieste relative ai componenti risultano distribuite su numerose categorie differenti, ciascuna caratterizzata da problematiche specifiche.

Per approfondire ulteriormente questo aspetto sono state analizzate le distribuzioni delle problematiche più frequenti. La Figura 7 mostra le dieci issue più ricorrenti associate ai motori completi, mentre la Figura 8 riporta le dieci issue più frequenti per il componente maggiormente rappresentato nel dataset, il turbocompressore.

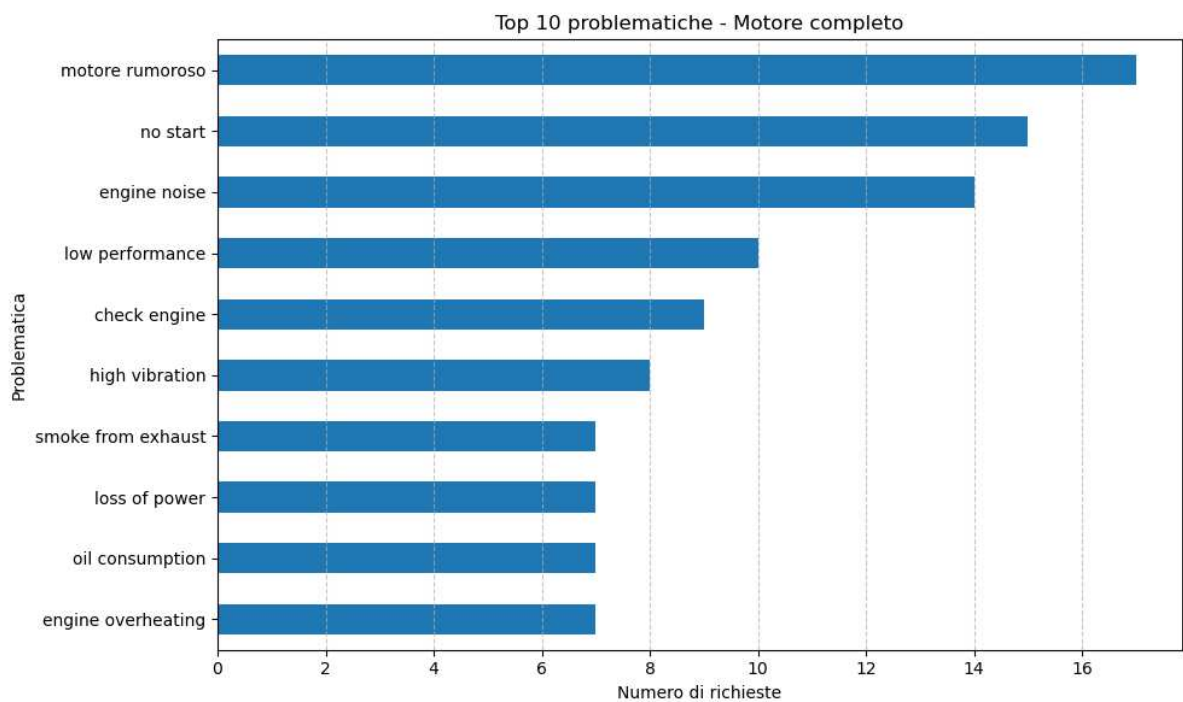


Figure 7 - Top 10 problematiche più frequenti per motore completo

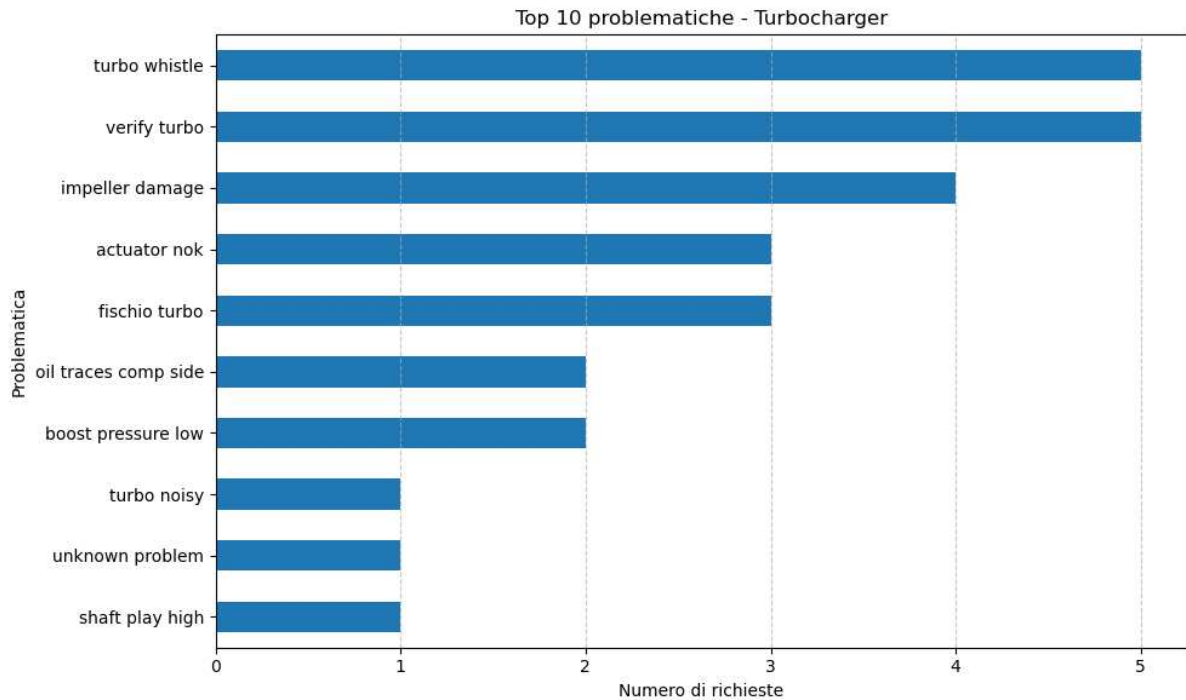


Figure 8 - Top 10 problematiche più frequenti per il turbocompressore

I risultati evidenziano come problematiche riconducibili a fenomeni simili possano essere descritte attraverso formulazioni differenti. Nel caso del turbocompressore, ad esempio, sono presenti descrizioni quali “turbo whistle”, “turbo noisy” e “fischio turbo”, che fanno riferimento a sintomi strettamente correlati ma vengono registrate come valori distinti. Analogamente, alcune descrizioni risultano particolarmente generiche, come “verify turbo” o “unknown problem”, rendendo più complessa una classificazione automatica delle informazioni.

Queste evidenze mostrano come la semplice normalizzazione del testo non sia sufficiente a risolvere il problema della frammentazione delle problematiche registrate. Pur eliminando differenze puramente formali, infatti, rimangono numerose descrizioni semanticamente simili ma registrate come valori

differenti. Tale situazione limita la possibilità di costruire indicatori affidabili e rende più complessa l'identificazione di problematiche ricorrenti all'interno del processo di rientro.

4.3 Standardizzazione delle problematiche

L'analisi esplorativa ha evidenziato come l'utilizzo di campi a testo libero per la registrazione delle problematiche possa generare un elevato livello di variabilità nelle informazioni memorizzate. Sebbene tale approccio garantisca flessibilità operativa agli utenti della piattaforma, esso introduce criticità significative dal punto di vista analitico, rendendo più complessa l'identificazione delle problematiche più frequenti e la costruzione di indicatori affidabili a supporto delle attività decisionali.

Per ridurre tali criticità, si propone l'introduzione di una struttura dati dedicata alla gestione delle problematiche standardizzate. In particolare, viene introdotta una nuova tabella denominata *TrrIssue*, contenente l'elenco delle problematiche codificate utilizzabili all'interno della piattaforma. Ciascuna problematica è identificata da un codice univoco e può essere associata a uno o più componenti attraverso una tabella ponte denominata *ItemTrrIssue*.

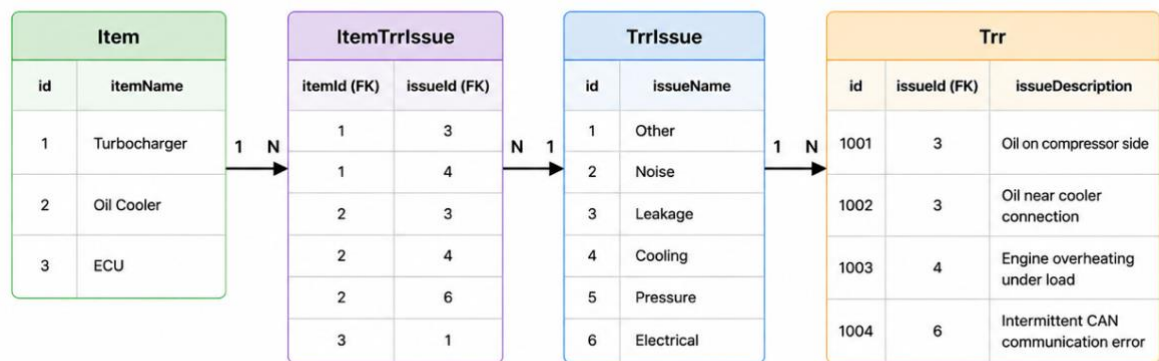


Figure 9 - Struttura delle tabelle per la standardizzazione delle problematiche

La Figura 9 mostra la struttura proposta e le relazioni tra le tabelle coinvolte. In questo modello una problematica può essere associata a più componenti differenti e, allo stesso tempo, ciascun componente può essere collegato a più problematiche. La relazione multi-a-molti viene gestita attraverso la tabella ItemTrrIssue. Inoltre, ogni richiesta di rientro memorizzata nella tabella Trr viene associata a una sola problematica standardizzata mediante il relativo identificativo, mentre una stessa problematica può comparire in più richieste di rientro.

Dal punto di vista operativo, tale soluzione potrebbe essere implementata all'interno della piattaforma tramite un menu a tendina dinamico. Una volta selezionato il componente oggetto della richiesta, l'utente visualizzerebbe esclusivamente l'elenco delle problematiche compatibili con quel componente, eliminando la necessità di inserire manualmente il testo della segnalazione e riducendo il rischio di errori, ambiguità o duplicazioni.

Al fine di rendere il modello dati maggiormente coerente con la struttura proposta, nel presente lavoro è stata adottata una versione semplificata della classificazione delle problematiche. Le segnalazioni originariamente presenti nel dataset sono state ricondotte a quindici categorie generiche, definite con l'obiettivo di rappresentare in modo sintetico le principali tipologie di malfunzionamento osservabili all'interno del processo analizzato.

L'assegnazione delle categorie è stata effettuata mantenendo una coerenza logica con il componente associato alla richiesta, preservando il significato informativo delle segnalazioni originali e riducendo al contempo la frammentazione delle problematiche registrate. È importante sottolineare che l'intervento ha riguardato esclusivamente il campo Issue, mentre il campo IssueDescription è stato mantenuto invariato. Quest'ultimo continua infatti a svolgere una funzione descrittiva utile ai tecnici per l'analisi dettagliata del singolo caso, mentre il campo Issue risulta sufficiente per le finalità statistiche e analitiche considerate nel presente lavoro.

La struttura così ottenuta è stata quindi assunta come riferimento per la costruzione del dataset utilizzato nelle successive attività di integrazione e analisi, rappresentando la configurazione finale del modello informativo considerato nello studio.

4.4 Validazione del modello

Per validare il modello di integrazione scelto per l'analisi, è stata realizzata una sperimentazione della pipeline all'interno dell'ambiente Databricks. L'obiettivo dell'attività non è stato quello di riprodurre integralmente l'infrastruttura aziendale, ma di verificare concretamente la fattibilità delle logiche di integrazione proposte e la possibilità di ottenere informazioni analitiche a partire dai dati provenienti dalle due piattaforme.

La sperimentazione è stata sviluppata utilizzando un sottoinsieme rappresentativo del dataset sintetico di MEERA costruito nel Capitolo 4, costituito da cento richieste di rientro associate a tre differenti tipologie di componenti. A tali informazioni sono stati affiancati dati esemplificativi relativi a RemoteMaster, generati in modo coerente con il modello dati descritto nel Capitolo 3 e con il flusso operativo previsto dalla piattaforma. In particolare, per ciascun componente sono stati definiti checklist, moduli, step di analisi ed esiti associati, così da riprodurre il processo di ispezione tecnica documentato da RemoteMaster.

All'interno di Databricks sono stati quindi creati tre schemi distinti, corrispondenti ai livelli Bronze, Silver e Gold della Medallion Architecture. Le tabelle di partenza sono state caricate nel livello Bronze, successivamente integrate e arricchite nel livello Silver mediante operazioni di join e trasformazione, fino alla costruzione nel livello Gold di una vista aggregata finalizzata alla sintesi delle informazioni provenienti dalle due piattaforme.

La Figura 10 riporta la struttura generale della pipeline implementata. Nei paragrafi successivi vengono descritti nel dettaglio i tre livelli dell'architettura, illustrando le tabelle realizzate, le trasformazioni applicate e i risultati ottenuti dalla simulazione.

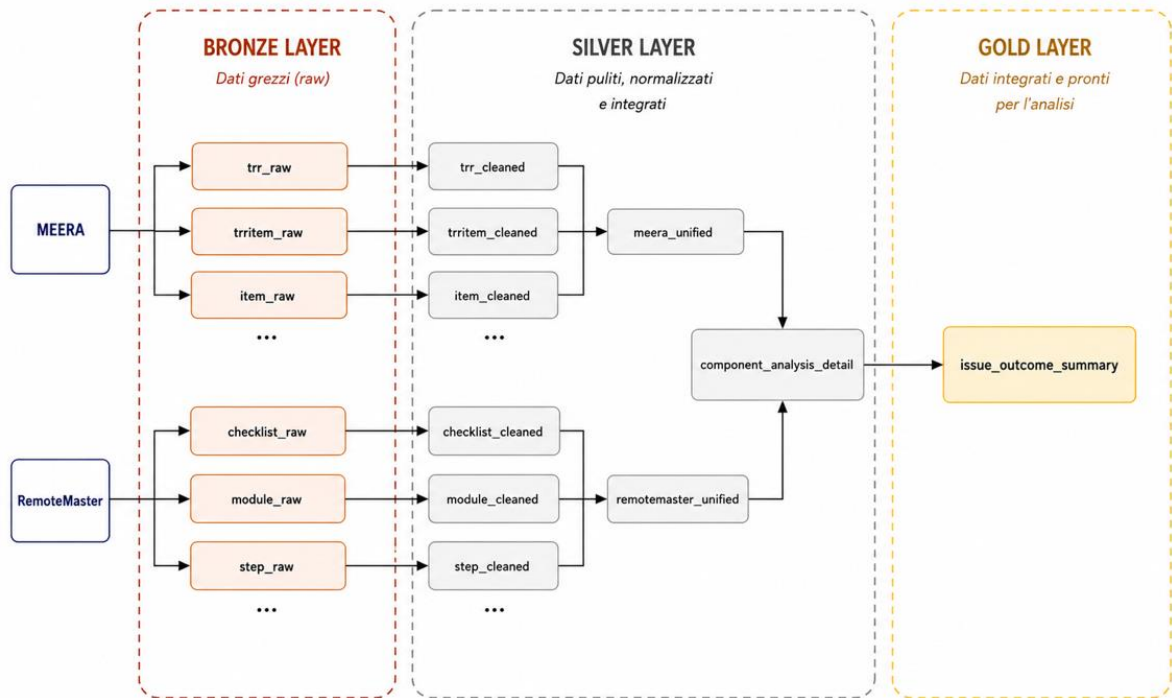


Figure 10 - Struttura del modello di integrazione proposto

Nei paragrafi seguenti i tre livelli verranno analizzati singolarmente, descrivendone la struttura, le logiche di costruzione e il ruolo assunto all'interno dell'architettura proposta.

4.4.1 Bronze Layer

Il Bronze Layer rappresenta il punto di ingresso delle informazioni all'interno della pipeline di integrazione. In questo livello vengono acquisite le principali tabelle appartenenti ai sistemi sorgente, mantenendone inalterata la struttura originaria e senza applicare alcuna logica di trasformazione o normalizzazione.

Nell'ambito della sperimentazione realizzata, il livello Bronze è stato implementato all'interno di Databricks mediante uno schema dedicato contenente le principali tabelle associate alle due piattaforme. Per MEERA sono state considerate le tabelle Trr, TrrItem, TrrTrrItem, Item, ItemTrrIssue e TrrIssue, mentre per RemoteMaster sono state utilizzate le tabelle PartNumber, Component, Checklist, Module, Step, CheckResult e StepResult.

4.4.2 Silver Layer

Nel livello Silver vengono costruite le strutture logiche necessarie alla successiva integrazione dei dati. A partire dalle tabelle presenti nel livello Bronze, sono state ricostruite le relazioni tra le diverse entità appartenenti a ciascuna piattaforma, creando viste unificate in grado di raccogliere in un'unica struttura le informazioni rilevanti per l'analisi.

Per la piattaforma MEERA è stata realizzata la vista meera_unified, ottenuta ricostruendo il collegamento tra la richiesta di rientro e il componente associato. La vista parte dalla tabella Trr, che contiene le informazioni generali

della richiesta, e la collega alla tabella ponte TrrTrrItem, utilizzata per associare ciascuna richiesta al relativo record presente in TrrItem. Attraverso TrrItem vengono recuperati gli attributi identificativi del componente, quali Claim, Part Number e Serial Number. Successivamente, il collegamento con la tabella Item consente di associare il componente alla relativa descrizione, mentre il collegamento con TrrIssue permette di sostituire l'identificativo della problematica con la corrispondente categoria standardizzata.

Il risultato è una vista unificata nella quale ogni richiesta di rientro viene rappresentata insieme al componente analizzato, alla problematica dichiarata e ad alcuni attributi amministrativi della richiesta, come data di invio, stato e priorità.

Column	Type
trr_id	bigint
claim	string
part_number	string
serial_number	string
item_name	string
issue_name	string
issue_description	string
submission_date	timestamp
status	string
priority	string

Figure 11 - Elenco dei campi presenti nella tabella "meera_unified" catturato da Databricks

Per la piattaforma RemoteMaster è stata realizzata la vista remotemaster_unified, costruita con l'obiettivo di rappresentare in un'unica struttura il dettaglio delle analisi tecniche eseguite sui componenti. La vista parte dalla tabella Component, che contiene gli identificativi del componente analizzato, tra cui Claim e Serial Number, e la collega alla tabella PartNumber, dalla quale vengono recuperati il codice identificativo del componente, la relativa descrizione e il fornitore.

A partire dal Part Number viene poi individuata la checklist associata al componente, tramite la tabella Checklist. Tramite il collegamento con la tabella Module è possibile ricostruire le sezioni in cui è articolata la checklist, mentre la tabella Step permette di recuperare le singole operazioni previste all'interno di ciascun modulo. Infine, la tabella StepResult collega ogni componente agli step effettivamente eseguiti e ai relativi esiti, recuperati attraverso la tabella CheckResult.

Il risultato è una vista unificata nella quale ciascun record rappresenta uno step eseguito durante l'analisi tecnica di un componente, comprensivo delle informazioni identificative del componente, della checklist utilizzata, del modulo, dell'operazione svolta e dell'esito registrato.

Column	Type
component_id	int
claim	string
part_number	string
component_name	string
storage_date	date
arrival_date	date
analysis_date	date
checklist_name	string
module_name	string
step_id	int
step_number	int
operation	string
ok_ko	string
text_if_ko	string

Figure 12 - Elenco dei campi presenti nella tabella "remotemaster_unified" catturato da Databricks

Una volta costruite le due viste intermedie, queste vengono integrate attraverso un'operazione di join utilizzando gli attributi Claim e Part Number, individuati come chiave logica di collegamento tra i due sistemi. Tale scelta consente di associare a ciascuna richiesta di rientro registrata in MEERA le informazioni prodotte durante le successive attività di analisi tecnica simulate in RemoteMaster.

```
CREATE OR REPLACE VIEW component_analysis_detail AS
```



```
SELECT
    m.claim,
    m.part_number,
    m.item_name,
    m.issue_name,
    r.analysis_date,
    r.module_name,
    r.step_number,
    r.ok_ko

FROM meera_unified AS m

INNER JOIN remotemaster_unified AS r
    ON m.claim = r.claim
    AND m.part_number = r.part_number;
```

L'operazione produce la vista `component_analysis_detail`, che rappresenta la prima struttura dati realmente integrata dell'architettura proposta. La vista combina infatti le informazioni relative al componente e alla problematica dichiarata con gli esiti delle verifiche effettuate durante il processo di analisi tecnica.

Column	Type
claim	string
part_number	string
item_name	string
issue_name	string
analysis_date	date
module_name	string
step_number	int
ok_ko	string

Figure 13 - Elenco dei campi presenti nella tabella "component_analysis_detail" catturato da Databricks

In particolare, per ogni richiesta di rientro vengono riportati il Claim, il Part Number, il componente coinvolto, la problematica associata, la data di analisi, il modulo della checklist eseguito, il numero dello step e il relativo esito. Con la struttura ottenuta è possibile, quindi, collegare direttamente le informazioni amministrative e operative provenienti da MEERA con i risultati generati durante l'ispezione tecnica del componente.

La Tabella 11 mostra un estratto della vista component_analysis_detail ottenuta dalla simulazione.

#	claim	part_number	item_name	issue_name	analysis_date	module_name	step_number	ok_ko
1	202590533201	5803141446	Turbocharger	Oil Leakage	2025-09-10	Visual Check	1	OK
2	202590533201	5803141446	Turbocharger	Oil Leakage	2025-09-10	Visual Check	2	OK
3	202590533201	5803141446	Turbocharger	Oil Leakage	2025-09-10	Visual Check	1	OK
4	202590533201	5803141446	Turbocharger	Oil Leakage	2025-09-10	Visual Check	2	KO

Table 11 - Estratto della tabella "component_analysis_detail"

Come evidenziato nella Tabella 11, ciascun record della vista rappresenta un singolo step di analisi eseguito su uno specifico componente. Questa scelta riflette la struttura del modello dati di RemoteMaster, nella quale gli esiti vengono registrati a livello di singolo controllo e non come risultato complessivo dell'intera analisi. Il mantenimento di tale livello di dettaglio rende possibile preservare integralmente le informazioni prodotte durante il processo di ispezione e costituisce la base per le successive operazioni di aggregazione e sintesi realizzate nel livello Gold.

4.3.3 Gold Layer

Nel livello Gold vengono realizzate le viste destinate alla sintesi delle informazioni integrate e al supporto delle successive attività di analisi. A differenza del livello Silver, che conserva il dettaglio operativo necessario a mantenere la tracciabilità delle informazioni provenienti dai sistemi sorgente, il livello Gold ha l'obiettivo di fornire strutture dati aggregate e immediatamente interpretabili.

A partire dalla vista integrata `component_analysis_detail` è stata costruita la vista `issue_outcome_summary`, progettata per mettere in relazione le problematiche dichiarate all'apertura del rientro con gli esiti delle verifiche tecniche effettuate durante l'analisi del componente. La vista consente di sintetizzare il comportamento delle diverse combinazioni tra Part Number e problematica segnalata, permettendo di osservare quante richieste di rientro siano state effettivamente confermate dall'analisi tecnica.

La costruzione della vista richiede innanzitutto la determinazione dell'esito complessivo di ciascun rientro. Poiché nel modello dati di RemoteMaster l'esito viene registrato a livello di singolo step, si assume che una richiesta di rientro sia confermata qualora almeno uno degli step eseguiti presenti esito KO. In caso contrario, il rientro viene considerato non confermato. Una volta determinato l'esito di ciascun rientro, le informazioni vengono aggregate per Part Number, componente e problematica dichiarata.

```
CREATE OR REPLACE VIEW gold.issue_outcome_summary AS

WITH claim_outcome AS (
    SELECT
        claim,
        part_number,
        item_name,
        issue_name,
        CASE
            WHEN MAX(CASE WHEN ok_ko = 'KO' THEN 1 ELSE 0 END) = 1
            THEN 1
            ELSE 0
        END AS confirmed_return
    FROM silver.component_analysis_detail
    GROUP BY
        claim,
        part_number,
        item_name,
        issue_name
)

SELECT
    part_number,
    item_name,
    issue_name,
```

```

COUNT(*) AS total_returns_issue,
SUM(confirmed_return) AS confirmed_returns,
COUNT(*) - SUM(confirmed_return) AS not_confirmed_returns,
ROUND(SUM(confirmed_return) * 100.0 / COUNT(*), 2) AS confirmation_rate,
ROUND((COUNT(*) - SUM(confirmed_return)) * 100.0 / COUNT(*), 2) AS false_return_rate

FROM claim_outcome

GROUP BY
    part_number,
    item_name,
    issue_name

ORDER BY
    item_name,
    confirmation_rate DESC;

```

La query mostrata costruisce una vista aggregata a partire da component_analysis_detail. Nella prima fase viene determinato l'esito complessivo di ciascun rientro attraverso la variabile confirmed_return, che assume valore pari a 1 nel caso in cui l'analisi presenti almeno uno step con esito KO e 0 in caso contrario. Successivamente, i risultati vengono aggregati per Part Number, componente e problematica dichiarata, calcolando il numero totale di rientri associati alla combinazione considerata, il numero di rientri confermati, il numero di rientri non confermati e i relativi indicatori percentuali.

Column	Type
part_number	string
item_name	string
issue_name	string
total_returns_issue	bigint
confirmed_returns	bigint
not_confirmed_returns	bigint
confirmation_rate	decimal(27,2)
false_return_rate	decimal(27,2)

Figure 14 - Elenco dei campi presenti nella tabella " issue_outcome_summary " catturato da Databricks

Selezionando l'intero contenuto della vista otteniamo il seguente risultato:

#	part_number	item_name	issue_name	total_returns_issue	confirmed_returns	not_confirmed_returns	confirmation_rate	false_return_rate
1	5040867812	EGR Valve	Electrical Failure	8	6	2	75.00	25.00
2	5040867812	EGR Valve	Contamination	12	6	6	50.00	50.00
3	5040867812	EGR Valve	Mechanical Damage	5	2	3	40.00	60.00
4	5809231745	High Pressure Pump	Calibration Error	8	7	1	87.50	12.50
5	5809231745	High Pressure Pump	Low Pressure	10	8	2	80.00	20.00
6	5809231745	High Pressure Pump	Oil Leakage	5	3	2	60.00	40.00
7	5809231745	High Pressure Pump	Mechanical Damage	2	1	1	50.00	50.00
8	5803141446	Turbocharger	Oil Leakage	20	18	2	90.00	10.00
9	5803141446	Turbocharger	Low Pressure	10	6	4	60.00	40.00
10	5803141446	Turbocharger	Mechanical Damage	5	3	2	60.00	40.00
11	5803141446	Turbocharger	Excessive Noise	15	4	11	26.67	73.33

Table 12 - Estrazione completa del contenuto della tabella "issue_outcome_summary"

La vista issue_outcome_summary rappresenta il risultato finale del processo di integrazione implementato nella sperimentazione. Grazie all'aggregazione delle informazioni provenienti dai livelli precedenti, è possibile sintetizzare in un'unica struttura i dati relativi ai rientri, alle problematiche segnalate e agli

esiti delle analisi tecniche, fornendo una base informativa adatta alle successive attività di interpretazione e valutazione.

4.5 Commento dei risultati e limiti del modello

I risultati ottenuti dalla vista `issue_outcome_summary` consentono di osservare come l'integrazione tra le informazioni provenienti da MEERA e RemoteMaster permetta di mettere in relazione le problematiche dichiarate all'apertura del rientro con gli esiti delle successive verifiche tecniche. La Tabella 13 riporta una sintesi delle principali combinazioni tra componente e problematica registrate nel dataset utilizzato per la sperimentazione.

Componente	Problematica	Totale rientri	Rientri confermati	Confirmation Rate (%)	False Return Rate (%)
Turbocharger	Oil Leakage	20	18	90,00	10,00
Turbocharger	Excessive Noise	15	4	26,67	73,33
EGR Valve	Contamination	12	6	50,00	50,00
High Pressure Pump	Calibration Error	8	7	87,50	12,50
High Pressure Pump	Low Pressure	10	8	80,00	20,00

Table 13 - Sintesi dei risultati della tabella "issue_outcome_summary"

Alcuni risultati mostrano comportamenti particolarmente significativi. Ad esempio, la combinazione tra Turbocharger e Oil Leakage presenta un numero elevato di rientri associato a un valore di `confirmation_rate` pari al 90%. Questo suggerisce che la problematica segnalata trova riscontro nella maggior parte delle analisi tecniche effettuate, configurandosi come un possibile difetto ricorrente meritevole di ulteriori approfondimenti.

Al contrario, la combinazione tra Turbocharger ed Excessive Noise presenta un numero significativo di segnalazioni ma un valore di `confirmation_rate` pari al 26,67%, accompagnato da un `false_return_rate` del 73,33%. Tale comportamento evidenzia una problematica frequentemente dichiarata ma raramente confermata durante l'analisi tecnica, suggerendo possibili criticità nella fase di diagnosi preliminare o nella classificazione delle anomalie.

Altri casi mostrano situazioni intermedie. Ad esempio, la combinazione tra EGR Valve e Contamination presenta valori equilibrati tra rientri confermati e non confermati, indicando una situazione in cui non è possibile trarre una conclusione univoca e che potrebbe richiedere ulteriori osservazioni su una base dati più ampia. Analogamente, combinazioni caratterizzate da pochi rientri ma valori elevati di conferma possono indicare problematiche effettivamente riscontrabili, ma ancora non sufficientemente frequenti da rappresentare una priorità immediata.

È opportuno sottolineare che gli indicatori ottenuti devono essere interpretati nel contesto della sperimentazione realizzata. Il dataset utilizzato è infatti di natura sintetica ed è stato costruito con l'obiettivo di validare le logiche di integrazione e le potenzialità analitiche del modello proposto, piuttosto che fornire evidenze statistiche sul processo reale. In un contesto operativo, l'analisi dovrebbe essere supportata da una base dati più ampia e storicizzata, comprendente un numero maggiore di componenti, problematiche e risultati di analisi.

Nonostante tali limitazioni, la sperimentazione dimostra come l'integrazione delle informazioni provenienti da MEERA e RemoteMaster renda possibile la costruzione di indicatori non ottenibili attraverso l'analisi separata delle due piattaforme. La disponibilità di una base dati condivisa permette infatti di collegare le problematiche dichiarate agli esiti delle verifiche tecniche, creando le condizioni per future attività di monitoraggio e supporto alle decisioni basate su informazioni più complete e strutturate.

Capitolo 5

5 Conclusioni e sviluppi futuri

Il presente lavoro ha avuto come obiettivo l'analisi delle informazioni generate durante il processo di gestione dei rientri in garanzia di FPT Industrial, con particolare attenzione alle piattaforme MEERA e RemoteMaster. Lo studio è nato dall'esigenza di comprendere come dati appartenenti a sistemi differenti possano essere valorizzati attraverso una maggiore integrazione e utilizzati come supporto alle attività operative e decisionali.

Una prima fase del lavoro è stata dedicata alla comprensione del processo aziendale e all'analisi delle strutture dati delle due piattaforme, che ha evidenziato come entrambe operino su fasi consecutive dello stesso processo. In particolare, MEERA supporta la gestione amministrativa e logistica delle richieste di rientro, mentre RemoteMaster è destinata alla registrazione e documentazione delle attività di analisi tecnica svolte sui componenti rientrati.

L'assenza di una base dati storica sufficientemente ampia e consolidata ha reso necessario sviluppare un dataset sintetico costruito a partire dalle strutture dati disponibili e dalle conoscenze operative raccolte durante il tirocinio. Questo

passaggio ha consentito di creare un ambiente di sperimentazione coerente con il contesto aziendale e di svolgere successive attività di analisi e integrazione.

Sulla base delle conoscenze acquisite è stato definito e successivamente implementato un modello di integrazione dei dati ispirato al paradigma Data Lakehouse e organizzato secondo la Medallion Architecture. La simulazione realizzata in ambiente Databricks ha permesso di riprodurre il flusso logico di acquisizione, trasformazione e integrazione delle informazioni, dalla costruzione delle tabelle sorgente fino alla realizzazione di viste aggregate dedicate all'analisi dei rientri.

Il principale contributo del lavoro consiste quindi nell'aver mostrato come una gestione più strutturata e integrata delle informazioni possa costituire un valido supporto ai processi di monitoraggio e alle attività decisionali. Pur trattandosi di una simulazione sviluppata in assenza di dati storici completi, i risultati ottenuti evidenziano il potenziale di un approccio integrato alla gestione delle informazioni generate durante il processo di garanzia. La disponibilità di una base dati condivisa consente infatti di collegare dati amministrativi, logistici e tecnici all'interno di un unico contesto informativo, creando le condizioni per attività di monitoraggio più efficaci e per future applicazioni di analisi avanzata. In questa prospettiva, il modello proposto non rappresenta soltanto un esercizio di integrazione dei dati, ma un possibile punto di partenza per l'evoluzione dei processi decisionali e delle pratiche di gestione della conoscenza all'interno dell'organizzazione.

5.2 Sviluppi futuri

Il lavoro svolto rappresenta un punto di partenza per possibili sviluppi futuri sia sul piano dell'integrazione dei dati sia su quello dell'evoluzione dei processi aziendali. L'integrazione proposta è stata costruita selezionando un sottoinsieme delle tabelle disponibili nelle due piattaforme, privilegiando quelle maggiormente rilevanti per la gestione delle garanzie e per le attività di analisi componenti. Una possibile evoluzione del lavoro potrebbe prevedere l'estensione del modello ad ulteriori entità presenti nei database, quali informazioni relative ai motori, ai veicoli, alle attività di magazzino e alle spedizioni. L'inclusione di tali dati consentirebbe di sviluppare analisi più approfondite e di esplorare ulteriori relazioni tra i diversi processi aziendali.

Un ulteriore sviluppo riguarda la validazione operativa del modello attraverso l'utilizzo di dati storici reali. La disponibilità di informazioni più estese e consolidate consentirebbe lo sviluppo di strumenti di Business Intelligence e dashboard dedicate al monitoraggio delle attività di rientro, permettendo di affiancare agli indicatori proposti ulteriori analisi e visualizzazioni a supporto delle decisioni aziendali.

Infine, sebbene il presente lavoro si sia concentrato esclusivamente sull'integrazione dei dati e non sull'integrazione diretta delle piattaforme, la

disponibilità di una base dati condivisa potrebbe rappresentare il presupposto per future forme di comunicazione automatica tra MEERA e RemoteMaster. Un'evoluzione di questo tipo permetterebbe di ridurre le attività manuali di trasferimento delle informazioni, migliorare la continuità del flusso operativo e garantire una maggiore tracciabilità lungo l'intero ciclo di gestione del componente, dalla richiesta di rientro fino alla definizione della sua destinazione finale.

8 Bibliografia

Armbrust, M., Ghodsi, A., Xin, R., Zaharia, M., Franklin, M., Stoica, I., & Zaharia, M. (2021). *Lakehouse: A New Generation of Open Platforms that Unify Data Warehousing and Advanced Analytics*. Proceedings of CIDR 2021.

https://www.cidrdb.org/cidr2021/papers/cidr2021_paper17.pdf

Databricks Documentation. (2026). *Medallion Architecture*.

<https://docs.databricks.com/aws/en/lakehouse/medallion>

FPT Industrial. (2023). *Sito ufficiale FPT Industrial*.

<https://www.fptindustrial.com>

FPT Industrial. (2025). *FPT Warehouse - User Stories Functional Requirements v.4* (documentazione aziendale).

Ferrante, A. (2022). *Analisi dei dati provenienti da veicoli connessi con telematica, al fine di correlare malfunzionamenti e possibili sostituzioni di componenti* (Tesi di laurea).