

LA TECHNOLOGY TRANSFER PRESENTA

# MIKE FERGUSON

**CATALOGHI DATI**

**governare e fornire dati in una azienda data-driven**

**ONLINE LIVE STREAMING**

**7 GIUGNO 2024**



[info@technologytransfer.it](mailto:info@technologytransfer.it)  
[www.technologytransfer.it](http://www.technologytransfer.it)

## DESCRIZIONE

Questa classe di un giorno esamina in dettaglio cos'è un Catalogo dati e quali sono i motivi principali per averne bisogno. In particolare, esamineremo la sfida che le aziende devono affrontare rispetto alla complessità dei dati e alla necessità di implementare sia un programma di Governance dei dati aziendali sia di fornire tecnologie per aiutare le persone a trovare, progettare, fornire e consumare dati "business ready" sia per l'Analitica che per casi d'uso operativi.

### Cosa Imparerete

I partecipanti impareranno come lavorano i Cataloghi dati e quali sono le loro funzionalità. Impareranno inoltre come usare i Cataloghi dati per scoprire, classificare e catalogare i dati in più archivi di dati sia on-premise e su più Cloud.

Le persone coinvolte nella Governance dei dati impareranno come utilizzare i Cataloghi di dati in un programma di governance dei dati aziendali per classificare sistematicamente i dati e impostare criteri per gestire il proprio patrimonio di dati distribuiti da un unico luogo. Questo include il rilevamento automatico e la classificazione dei dati sensibili, la governance della sicurezza dell'accesso ai dati, la privacy dei dati, la prevenzione della perdita dei dati, la condivisione dei dati, l'utilizzo dei dati, la conservazione dei dati e la qualità dei dati.

Inoltre, i data engineer e i data scientist impareranno come utilizzare i Cataloghi dati per il discovery dei dati che possono essere usati nelle pipeline di integrazione dei dati per produrre prodotti dati.

### PARTECIPANTI

Questo seminario è rivolto a professionisti aziendali e IT responsabili dell'ingegneria dei dati, del provisioning dei dati e della governance dei dati aziendali, tra cui la sicurezza dell'accesso ai dati, la privacy dei dati, la condivisione dei dati, l'utilizzo dei dati, la conservazione dei dati, la qualità sia dei dati strutturati che dei contenuti.

- Chief Data Officer
- Data Engineer
- Data Architect
- Data Scientist
- Responsabili di Data Governance
- Data Steward
- Solution Architect
- Enterprise Architect

## 1. Introduzione

Questa sessione esamina i tipici setup che oggi esistono nelle aziende e le sfide che queste devono affrontare relativamente ai dati e spiega perché i Cataloghi dati sono necessari.

- Il complesso panorama dei dati distribuiti in continua crescita: on-premise, più Cloud, applicazioni SaaS
- La crescita delle nuove sorgenti dati
- I diversi sistemi transazionali: applicazioni SaaS, on-premise e Cloud
- Sistemi analitici a silos esistenti: Data Warehouse, Data Lake, Data Lakehouse
  - o Problemi di Data Engineering in un ambiente a silos
- L'emergenza del Data Mesh e il suo impatto nei confronti di Data Engineering e Data Architecture
- L'impatto dei dati non governati
  - o Operazioni di Business, Decision Making e rischio
  - o La redditività del Business e la capacità di rispondere alla pressione della competizione
  - o L'abilità di essere compliant con la legislazione sulla Data Privacy in una o più giurisdizioni
  - o Sicurezza dei dati, la possibilità di Data Breaches e il loro impatto sul Business
- Principali requisiti che le aziende richiedono rispetto ai dati
  - o La necessità di una governance dei dati end-to-end per:
    - Semplificare la ricerca e la comprensione dei dati ovunque si trovino
    - Sapere dove sono i dati sensibili per poterli governare
    - Essere compliant con le diverse leggi e regolamenti sulla Data Privacy
    - Evitare Data Breaches in un ambiente distribuito complesso
    - Capire dove la qualità dei dati è scadente
    - Governare l'accesso ai dati, la privacy, la condizione, l'uso, la retention e la qualità
  - o La necessità di accelerare il Data Engineering e fornire dati pronti per il business da condividere e riutilizzare all'interno dell'azienda
  - o Principali tecnologie necessarie: il Catalogo dati e Data Fabric

## 2. Cosa è il Catalogo dati e perché dobbiamo averlo?

Questa sessione esamina cos'è un Catalogo dati, perché ne avete bisogno, i fornitori di software sul mercato e delinea le sue funzionalità.

- Cosa è un Catalogo dati
- Perché averlo?
  - o Casi d'uso per il Catalogo dati
- Il mercato del Catalogo dati
  - o Alation, Ataccama, AWS Glue Data Catalog, BigID, Cambridge Semantics Anzo Data Catalog, Collibra Data Catalog, data.world, Google Data Catalog, Hitachi Vantara Lumada, IBM Watson Knowledge Catalog, Informatica IDMC Data Governance and Catalog, Microsoft Purview, Oracle, SAP, Talend Data Catalog, Top Quadrant TopBraid, Truist Zaloni Data Catalog
- Fondamentali funzionalità del Catalogo dati
  - o Glossario del Business
  - o Data discovery automatizzato
  - o Mapping al glossario di Business sia manuale che automatico
  - o Classificazione di Data Governance manuale o automatica
  - o Profiling di Data Quality automatizzato
  - o Gestione della politica di Data Governance
- Quali tipi di metadati e di asset sono memorizzati in un Catalogo dati?
- Supportare diverse tipologie di persone
- Competenze necessarie per usarne uno
- Manutenzione del Catalogo dati: manuale o automatica

## 3. Catalogo Dati: l'importanza di un Glossario di Business

Questa sessione spiega perché è necessario capire il panorama dei dati da una prospettiva Business. La chiave per far sì che ciò accada è stabilire un vocabolario di Business comune nel Glossario di Business di un Catalogo dati per creare nomi e definizioni di dati comuni per i vostri dati. Ciò vi consente di cercare e gestire i dati da una prospettiva di Business.

- Standardizzazione dei dati usando un vocabolario di Business condiviso
- Lo scopo di un vocabolario comune nella Data Governance
- Software di Glossario di Business - ora come una capability del Catalogo dati
  - o Alation, Amazon Glue, Collibra, Informatica Axon Business Glossary, IBM Watson Knowledge Catalog, Microsoft Azure Purview, Talend Business Glossary and Data Catalog, SAS Business Data Network, TopQuadrant TopBraid EDG Business Glossary
- Pianificare per un Glossario del Business
- Ruoli e responsabilità per il Glossario

- Processo di sottomissione dei termini del Glossario, approvazione con voto e risoluzione delle controversie
- Approcci per creare un vocabolario comune
- Organizzare le definizioni dei dati in un Glossario di Business
- Il ruolo di un Data Concept Model
- Utilizzare un vocabolario comune nel Data Modeling, ETL, BI, ESB, APIs, MDM

#### 4. Catalogo Dati: comprendere il vostro panorama di dati utilizzando il discovery automatico dei dati, la catalogazione e la mappatura a un Glossario di Business

Dopo aver definito i vostri dati, questa sessione esamina quali dati avete, dove si trovano e come si associano al vostro Glossario di Business per fornire una comprensione aziendale del vostro panorama di dati.

- Capire il vostro panorama dati: il ruolo critico del software di Data Catalog
- Il processo di Data Discovery
- Registrare le sorgenti dati per il Discovery
- Data discovery e prolifing automatizzato usando un Catalogo dati
- Fare il mapping di asset di dati al Glossario di Business

#### 5. Catalogo dati: classificare dati e contenuti per sapere come governarli

Questa sessione esamina l'etichettatura manuale e automatica dei dati utilizzando un Catalogo dati per sapere come governarli utilizzando classificatori predefiniti, schemi di classificazione definiti dall'utente e classificatori addestrabili. Esamina quindi il modo in cui i dati classificati vengono visualizzati in un Catalogo dati e in che modo i criteri possono essere assegnati ai dati etichettati per governarli in tutto il patrimonio dati.

- Cosa è una classificazione dei dati?
- Individuazione e classificazione automatiche di dati sensibili utilizzando classificatori addestrati predefiniti
- Creare i propri schemi di classificazione dei dati per la riservatezza e la conservazione dei dati
- Classificazione manuale dei contenuti utilizzando il proprio schema di classificazione, per esempio: Documenti Office, SharePoint, Email, Chat, Microsoft Teams o Zoom Meetings
- Classificatori per etichettare automaticamente i contenuti
- Utilizzo di classificatori addestrati per etichettare au-

- tomaticamente i contenuti nel Cloud e in locale
- Utilizzo dei propri schemi di classificazione con un Catalogo dati
- Classificazione automatica di oggetti e dati strutturati sensibili utilizzando un Catalogo dati
- Utilizzo di informazioni dettagliate sulla classificazione per comprendere la proliferazione e la ridondanza dei dati sensibili all'interno del tuo patrimonio dati
- Impostazione di criteri in un Catalogo dati per governare tutti i dati del vostro patrimonio

#### 6. Caso d'uso del Catalogo dati: cosa è necessario per governare i dati in un ambiente di dati distribuiti?

Questa sessione esamina quali sono i requisiti per governare i dati in un'azienda moderna e come il requisito può essere soddisfatto utilizzando un Catalogo dati.

- Principali requisiti per governare dati e contenuti in un ambiente di dati distribuiti
- Di cosa avete bisogno per sapere come governare i dati?
- Introdurre un framework di Data Governance per aiutare ad affrontare la sfida
- Persone
  - o Principali ruoli e responsabilità
  - o Ottenere l'organizzazione e il modello operativo corretti
  - o Proprietari dei dati, amministratori dei dati, comitato di controllo della governance dei dati e gruppi di lavoro
- Processi fondamentali necessari per stabilire e governare i dati comunemente compresi
- Tipi di politiche e regole necessarie per governare:
  - o Data quality
  - o Data access security
  - o Data privacy
  - o Data retention
  - o Data loss prevention
  - o Data sharing
  - o Data use and maintenance
- Il ruolo del Catalogo dati
  - o Discovery automatico dei dati sensibili
  - o Trainable Classifiers
  - o Creazione di una policy per il Catalogo dati
  - o Integrazione con le tecnologie di applicazione delle policy
    - Dynamic data masking
    - Data loss prevention
- Le principali e necessarie funzionalità della Data Governance
- Task coinvolti nel governare un ambiente di dati distribuiti

## 6. Caso d'uso del Catalogo dati: trovare, ingegnerizzare e fornire dati

Questa sessione esamina quali sono i requisiti per accelerare la Data Engineering in un'azienda moderna e come soddisfare i requisiti utilizzando un software di Catalogo dati e di Data Fabric.

- Definire prodotti dati in un Glossario di Business
- Discovering, mappatura e classificazione automatici dei dati in un Catalogo dati
- Costruire pipelines per produrre prodotti dati
- Creazione di un mercato dei dati come applicazione di Catalogo dati per condividere dati pronti per il Business
- Pubblicazione e consumo di prodotti dati utilizzando un mercato dei dati e i metadati del Catalogo dati

## DOCENTE

**Mike Ferguson** è Managing Director di Intelligent Business Strategies Limited. Come analista e consulente è specializzato in Business Intelligence e Enterprise Business Integration. Ha più di 41 anni di esperienza di IT, ha operato consulenze per importanti aziende ed è stato frequentemente speaker in molti eventi e seminari in tutto il mondo. E' stato Principal e co-fondatore della Codd & Date UK e Chief Architect alla Teradata su DBMS Teradata e Managing Director Europeo di Database Associates. Insegna delle master classes che hanno molto successo su Analitica, Big Data, Data Governance & MDM, Data Warehouse Modernization e Data Lake Operations.

# INFORMAZIONI

QUOTA DI PARTECIPAZIONE	MODALITÀ D'ISCRIZIONE	CONDIZIONI GENERALI	TUTELA DATI PERSONALI
€ 800 (+iva)	Il pagamento della quota, IVA inclusa, dovrà essere effettuato tramite bonifico, codice IBAN: IT 03 W 06230 03202 000057031348 Banca: Cariparma Agenzia 1 di Roma intestato alla Technology Transfer S.r.l. e la ricevuta di versamento inviata insieme alla scheda di iscrizione a: info@technologytransfer.it	In caso di rinuncia con preavviso inferiore a 15 giorni verrà addebitato il 50% della quota di partecipazione, in caso di rinuncia con preavviso inferiore ad una settimana verrà addebitata l'intera quota. In caso di cancellazione del seminario, per qualsiasi causa, la responsabilità della Technology Transfer si intende limitata al rimborso delle quote di iscrizione già pervenute.	Ai sensi dell'art. 13 della legge n. 196/2003, il partecipante è informato che i suoi dati personali acquisiti tramite la scheda di partecipazione al seminario saranno trattati da Technology Transfer anche con l'ausilio di mezzi elettronici, con finalità riguardanti l'esecuzione degli obblighi derivati dalla Sua partecipazione al seminario, per finalità statistiche e per l'invio di materiale promozionale dell'attività di Technology Transfer. Il conferimento dei dati è facoltativo ma necessario per la partecipazione al seminario. Il titolare del trattamento dei dati è Technology Transfer, Piazza Cavour, 3 - 00193 Roma, nei cui confronti il partecipante può esercitare i diritti di cui all'art. 13 della legge n. 196/2003.
<b>DURATA ED ORARIO</b> 1 giorno: 9.30-13.00 14.00-17.00	<b>TECHNOLOGY TRANSFER S.r.l.</b> Piazza Cavour, 3 00193 ROMA (Tel. 06-6832227)	<b>SCONTI</b> I partecipanti che si iscriveranno al seminario 30 giorni prima avranno uno sconto del 5%.  Se un'azienda iscrive allo stesso evento 5 partecipanti, pagherà solo 4 partecipazioni. Gli sconti per lo stesso evento non sono cumulabili fra di loro.	
<b>Il seminario sarà tenuto in lingua inglese con il servizio di Traduzione Simultanea incluso.</b>	Vi consigliamo di far precedere la scheda d'iscrizione da una prenotazione telefonica.		

## MIKE FERGUSON

CATALOGHI DATI

7 GIUGNO 2024

Quota di iscrizione:

€ 800 (+iva)

nome .....

cognome .....

funzione aziendale .....

azienda .....

partita iva .....

codice fiscale .....

indirizzo .....

città .....

cap .....

provincia .....

telefono .....

fax .....

e-mail .....



Timbro e firma

In caso di rinuncia o di cancellazione dei seminari valgono le condizioni generali riportate all'interno.

Da restituire compilato a:  
**Technology Transfer S.r.l.**  
Piazza Cavour, 3 - 00193 Roma  
Tel. 06-6832227 - Fax 06-6871102  
info@technologytransfer.it  
www.technologytransfer.it